

Metallomics

Accepted Manuscript



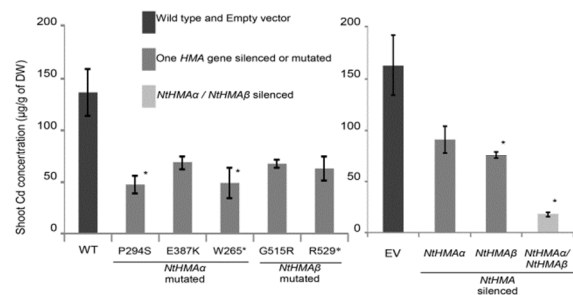
This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

Using *NtHMA* mutants and *NtHMA*-silenced lines we successfully developed tobacco plants with reduced cadmium concentration in shoots.



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Title: Inactivation of two newly identified tobacco heavy metal ATPases leads to reduced Zn and Cd accumulation in shoots and reduced pollen germination

Victor Hermand¹, Emilie Julio², François Dorlhac de Borne², Tracy Punshon³, Felipe K Ricachenevsky⁴, Arnaud Bellec⁵, Françoise Gosti¹, Pierre Berthomieu¹

(1) Institut National de la Recherche Agronomique, Montpellier SupAgro, Centre National de la Recherche Scientifique, Université Montpellier 2, UMR Biochimie et Physiologie Moléculaire des Plantes, Place Viala, 34060 Montpellier, France ; (2) Institut du Tabac, Domaine de la Tour, LBCM, 24100 Bergerac, France ; (3) Department of Biological Sciences, Dartmouth College, Hanover, New Hampshire, United States of America ; (4) Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Avenida Bento Gonçalves 9500, Porto Alegre, Brazil (5) Institut National de la Recherche Agronomique, Centre National de Ressources Génomiques Végétales, Chemin de borde Rouge, BP 52627, 31326 Castanet Toulousan, France.

Keywords: amiRNA, biotechnology, Cadmium, EMS-mutagenized lines, GUS staining, Heavy metal, Heavy Metal ATPases, metal repartition, NtHMA α , NtHMA β , pollen, pollen germination, seed, seed metal concentration, silencing, SXRF, tolerance, toxicity, zinc.

Abstract

Cadmium (Cd) is a non-essential heavy metal, which is classified as a “known human carcinogen” by the International Agency for Research on Cancer (IARC). Understanding the mechanisms controlling Cd distribution *in planta* is essential to develop phytoremediation approaches as well as for food safety. Unlike most other plants, tobacco (*Nicotiana tabacum*) plants translocate most of the Cd taken up from the soil, out of the roots and into the shoots, leading to high Cd accumulation in tobacco shoots. Two orthologs to the *Arabidopsis thaliana* *HMA2* and *HMA4* Zn and Cd ATPases that are responsible for zinc (Zn) and Cd translocation from roots to shoots were identified in tobacco and sequenced. These genes, named *NtHMA α* and *NtHMA β* , were more highly expressed in roots than in shoots. *NtHMA α* was expressed in the vascular tissues of both roots and leaves as well as in anthers. No visual difference was observed between wild-type plants and plants in which the *NtHMA α* and *NtHMA β* genes were either mutated or silenced. These mutants showed reduced Zn and Cd accumulation in shoots as well as increased Cd tolerance. When both *NtHMA* genes were silenced, plant development was altered and pollen germination was severely impaired due to Zn deficiency. Interestingly, seeds from these lines also showed decreased Zn concentration but increased iron (Fe) concentration.

Introduction

Many metals and metalloids are essential for plant and animal development. For example, zinc (Zn) is a cofactor in more than 300 plant enzymes, including carbonic anhydrase, RNA polymerase, Zn-superoxide dismutase and alcohol dehydrogenase, all of which are key enzymes for plant survival.¹ However, both metals and metalloids are toxic in excess.

Cadmium (Cd) is one of the few metals for which there is no reported beneficial effect for plant or human development. It is classified in Group 1 (known human carcinogen) by the International Agency for Research on Cancer (IARC). It has been hypothesized that Cd circulates in plants by co-opting transporters of other divalent metals.² Accumulated in excess, Cd triggers many deleterious effects on plants such as a reduction in leaf surface, dry weight, water content, chlorophyll content, and carotenoid content, as well as an increase in protease activity.³ Cd toxicity is also due to the similarity (in terms of size and charge) between this element and other metals such as Zn, iron (Fe) and calcium (Ca). Cd causes toxicity by mimicking these elements, and replacing them in compartments that normally do not contain Cd. Another explanation for Cd toxicity could be the association of Cd with proteins which are essential for plant survival.⁴ Cd can be accumulated in plants at levels that do not trigger any symptom of toxicity in the plant but can be toxic upon human consumption.

One way Cd enters the human body is by ingestion of Cd-contaminated food. However, inhalation of smoke from cigarettes is a far more effective uptake route than absorption through the gastrointestinal tract. Smoker's blood-Cd content is 4 to 5 times higher than non-smokers.⁵ Exposure to Cd induces damage in bones, lungs and liver. Cd has also been reported to deregulate hormone balance leading to spontaneous abortion.⁶ It is thus of interest that tobacco plants accumulate less cadmium in the leaves so that less cadmium would be found in the smoke of cigarettes. Tobacco is a plant species characterized by an ability to accumulate four times higher levels of Cd in the shoots than in the roots.⁷ By comparison, Cd concentration in *Arabidopsis thaliana* is approximately two fold higher in roots compared to

1
2
3 shoots⁸ and Cd concentration in the cadmium hyper accumulator species
4 *Arabidopsis halleri* is similar in roots and in shoots.⁹
5
6

7
8 Previous attempts to reduce the Cd concentration of tobacco leaves were
9 based on the expression of transgenes. For example, reducing the
10 accumulation of Cd in shoots has been achieved by sequestering Cd in root
11 vacuoles. This was accomplished by over-expressing the *A. thaliana* CAX2
12 and CAX4 Ca and manganese (Mn) vacuolar transporters in tobacco roots¹⁰.
13 The same result had previously been reported with transgenic tobacco
14 plants expressing a human metallothionein.^{11,12} The role of plant
15 metallothioneins is still unclear but animal metallothioneins were shown to
16 chelate Cd and modify its distribution within cells and organs.¹³
17
18
19
20
21
22
23

24
25 The recent knowledge gained about the transport of metals between organs,
26 cells or subcellular compartments in model plants opens an avenue to
27 develop a non-GMO approach to reduce Cd accumulation in tobacco leaves.
28 In *A. thaliana* only one transporter, IRT1, has to-date been identified as
29 responsible for Cd uptake into roots. IRT1 is a Fe transporter also
30 responsible for the uptake of many other essential metals such as Zn or
31 Mn.¹⁴ In the roots, Cd diffuses through the symplasm to reach the vascular
32 tissue. There, it can be loaded into the xylem by HMA2 and HMA4, two Zn
33 and Cd ATPases.^{15,16,8} Transpiration then drives the translocation of Cd from
34 the roots to the shoots. In shoot and root cells, Cd can be transported to the
35 vacuole by the CAX2 and CAX4 Mn and Ca transporters¹⁷ or the HMA3 Cd
36 and lead (Pb) transporter.¹⁸ In addition to Fe and Mn, NRAMP3, NRAMP4
37 and NRAMP6 can transport Cd out of the vacuole.^{19,20} Phytochelatin also
38 play an important role in Cd distribution. They are capable of chelating Cd
39 and facilitating its transport into cell vacuoles.²¹
40
41
42
43
44
45
46
47
48
49
50

51
52 Several of the above-mentioned transporters could be good targets to
53 develop tobacco lines with reduced Cd accumulation in the shoots through
54 gene inactivation. However, the genome of tobacco is still unavailable and
55 tobacco orthologues to these transporters cannot be easily identified from
56 the tobacco raw sequences available in public databases.
57
58
59
60

1
2
3
4 Working on orthologues of *AtHMA2* and *AtHMA4* seems particularly wise
5 since the inactivation of these genes led to an almost complete suppression
6 of Cd translocation from the roots to the shoot in *Arabidopsis thaliana*⁸. The
7 *athma2 athma4* double mutant also showed enhanced sensitivity to
8 cadmium whereas lines overexpressing *AtHMA4* were more resistant to
9 cadmium and zinc and translocated both metals more efficiently from roots
10 to shoots.¹⁶ Interestingly, *Arabidopsis* plants mutated in either *AtHMA2* or
11 *AtHMA4* present normal development, but when both genes are mutated
12 development is impaired and fertility is reduced.¹⁵

13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
Orthologs of *AtHMA2* and *AtHMA4* have already been identified and
characterized in a crop, namely rice. In rice, *OsHMA2* has been shown to be
particularly important for Zn and Cd transport to the shoot and for Zn
distribution to the reproductive tissues.²² Reducing the accumulation of HMA
protein has recently been proposed as a strategy to reduce Cd content in
rice grain.²³ No similar information has to-date been available concerning the
tobacco orthologs of *AtHMA2* or *AtHMA4*.

The goal of the present work was thus to investigate the effects of altering
the expression of the tobacco orthologs of *AtHMA2* and *AtHMA4* and test
the possibility of using this approach to obtain tobacco lines with reduced Cd
levels in the leaves. The genome of tobacco has not been fully sequenced
yet. Tobacco is an allopolyploid ($2n = 4x = 48$) whose genome originates
from the fusion of the ancestral genomes of *Nicotiana tomentosiformis* ($2n =$
 $2x = 24$) and *Nicotiana glauca* ($2n = 2x = 24$).²⁴ The genome of tobacco is
0.2 Myr old; most genes are still present in 2 copies, originating each from
one of the two ancestors.²⁵ Two *NtHMA* genes were identified from the
screening of a bacterial artificial chromosome (BAC) library that we
constructed from *N. tabacum* genomic DNA. The role of the two newly
identified *NtHMA* genes was investigated using plants harboring stop or
missense mutations in these genes, as well as using plants expressing
artificial microRNA (amiRNA) constructs. Contributions of the *NtHMA* genes
to Cd and Zn accumulation in the shoots, to Cd and Zn tolerance and to
plant development and fertility were examined.

Results

Identification of tobacco homologs to *AtHMA2* and *AtHMA4*

Using Southern hybridization Tobacco homologs of *AtHMA2* and *AtHMA4* were identified by screening a tobacco genomic BAC library that we constructed for that purpose. This BAC clone library represents 4 times the *N. tabacum* genome. Positive BAC clones were sequenced. Only two different members of the heavy metal ATPase family were identified and named *NtHMA α* and *NtHMA β* . Their cDNAs were also cloned, indicating that the two genes are expressed. The sequences that we obtained were used as baits to screen for similar sequences among the available tobacco database. No other similar sequence was found. We thus consider that *NtHMA α* and *NtHMA β* are the only two *N. tabacum* orthologs to *AtHMA2* and *AtHMA4*.

The two genes are very long, the ATG initiation codon and the stop codon being separated by approximately 16.5 kb in the genome. The encoded proteins have 1402 amino acids (a.a.) for *NtHMA α* and 1444 a.a. for *NtHMA β* , compared to 1172 a.a. for *AtHMA4*. They are very similar to each other, showing 93% a.a. identity over the whole protein. The promoter regions of *NtHMA α* and *NtHMA β* were also very similar (~90% sequence identity on 1.75kb), except for the insertion/deletion of a defective 700bp transposable element 200bp upstream of the ATG initiation codon of *NtHMA β* . Major differences were in contrast observed within the introns mainly due to numerous large insertions/deletions.

The genome sequences of *N. sylvestris*, *N. tomentosiformis* and *N. benthamiana* have very recently been released in public databases.^{26,27} Comparisons between the *NtHMA α* and *NtHMA β* genomic sequences and the genomic sequences of these Nicotianeae clearly revealed that there is only one ortholog to the *NtHMA* genes in each of their genomes. Focussing on *N. sylvestris* and *N. tomentosiformis*, which are the two ancestors of tobacco, we observed that *NtHMA α* is more closely related to the *N. sylvestris* copy (99% nucleic acid identity, including in non-coding

1
2
3 sequences) and *NtHMA β* to the *N. tomentosiformis* one (98% nucleic acid
4 identity, including in non-coding sequences).
5
6

7
8 *NtHMA α* and *NtHMA β* are not highly similar to their *A. thaliana* homologs.
9 The intron-exon organization is similar in the first part of the gene. However,
10 the last exon of the *AtHMAs* is split into three separate exons in the tobacco
11 *HMA*s (Sup. fig. 1A). The *NtHMA α* and *NtHMA β* proteins only show
12 approximately 65% amino acid similarity with their *A. thaliana* homologs, and
13 this similarity is only found in the first part of the protein, which correspond
14 the transmembrane domain (Sup. fig. 1B). In the second part of the protein,
15 which encodes a cytosolic C-terminal tail, the similarity between the tobacco
16 and the *AtHMAs* decreases to less than 40%. Previous publications have
17 shown that this cytoplasmic tail plays an important role for the activity of the
18 protein.^{8,28} Interestingly, the C-terminal tails of *NtHMA α* and *NtHMA β* are
19 much longer (704 and 746 a.a., respectively) than those of *AtHMA2* and
20 *AtHMA4* (267 and 470 a.a., respectively) and possess 11 and 13 cysteines
21 present in tandem, respectively. The stretch of histidines located at the C-
22 terminal end of the *AtHMA2* and *AtHMA4* proteins (3 and 11 histidines,
23 respectively) is lacking in *NtHMA α* and *NtHMA β* . The long C-terminal tails of
24 *NtHMA α* and *NtHMA β* are also found, with >99.5% amino acid identity, in
25 the *N. sylvestris* and *N. tomentosiformis* respective homologs (711 and 728
26 a.a., respectively). Interestingly, a long C-terminal tail was also found in the
27 *N. benthamiana* (788 a.a. and 13 cystein tandems) ortholog of *NtHMAs* and
28 to a lesser extent as well as with much greater sequence divergence in the
29 single *Solanum lycopersicum* ortholog of *NtHMAs* (601 a.a., 9 cysteine
30 tandems).
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48

49 Zn and Cd HMA transporters are also characterized by specific amino acid
50 motifs, which include the GICCTSE metal binding domain present on the N-
51 terminal tail of the protein, the DKTGT P-domain motif located in the third
52 cytosolic loop and the CPx/SPC motif located in the sixth transmembrane
53 domain.^{29,30} All these motifs are present in *AtHMA2*, *AtHMA4*, *NtHMA α* and
54 *NtHMA β* (Sup. fig. 2).
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

A phylogenetic analysis performed with several HMA proteins, considering their entire N-terminal moiety which comprises all the transmembrane domains (Fig. 1), clearly showed that NtHMA α and NtHMA β belong to the P_{1B} ATPase clade gathering Zn, Cd and Pb ATPases. Within this clade, they belong to a sub-clade gathering Solanaceae sequences, which is distinct from subclades gathering Brassicaceae or Poaceae sequences. Thus no clear and direct orthologous relationship could at this stage be established between NtHMA α and NtHMA β and the *A. thaliana* or rice HMAs for which functions have already been ascribed. Within the Solanaceae sub-clade only tobacco has two members of the Zn, Cd and Pb ATPases family. Analyses of sequence similarities revealed that the *S. lycopersicum* copy was closer to NtHMA α and NsylHMA4 than to NtHMA β and NtomHMA4 while the *N. benthamiana* copy was closer to NtHMA β and NtomHMA4 than to NtHMA α and NsylHMA4.

***NtHMA α* and *NtHMA β* are mainly expressed in roots and their level of expression is stable over a wide range of Cd and Zn concentrations in the culture medium**

Using quantitative PCR analysis, transcripts of both *NtHMA α* and *NtHMA β* were shown to be mainly accumulated in roots, although they could also be detected in shoots (Fig. 2A). This was a similar situation as for *AtHMA2* and *AtHMA4*¹⁶ or *AhHMA4-2*.⁹ The *NtHMA α* and *NtHMA β* transcript levels were comparable to each other, both in roots and shoots.

When plants were exposed to moderate concentrations of Cd or Zn (0 to 10 μ M) no visible impact on the growth of the plants and no modification of *NtHMA α* and *NtHMA β* transcript accumulation could be observed (Fig. 2B, 2C). The accumulation of the transcripts of both genes was increased only when much higher concentrations of Cd or Zn (100 μ M of Cd or 30-100 μ M of Zn) were applied. Applying these high Zn or Cd treatments in hydroponic conditions was toxic for the plants. The 100 μ M Cd treatment induced necroses in the leaf lamina while the 30-100 μ M Zn treatments induced chlorosis around the main vein (data not shown).

***NtHMA α* is expressed in root vascular cells, leaf veins and anthers**

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

To further analyze the expression pattern of *NtHMA α* , a promoter::GUS fusion was constructed. A 3-kb-long region upstream the *NtHMA α* ATG initiation codon was cloned immediately upstream of the GUS coding sequence and introduced in wild-type tobacco. *NtHMA α* appeared to be expressed in the vascular tissues in roots, more precisely in xylem parenchyma cells (Fig. 3A to 3D). Interestingly, expression of *NtHMA α* could be detected in young roots but not in older ones. Expression of *NtHMA α* could also be detected around the primary, secondary and tertiary veins in leaves (Fig. 3E). This pattern was however only observed in young leaves (about 4 cm in length). Two- to four-week-old leaves only showed *NtHMA α* expression in tertiary veins (Fig. 3F) and even older leaves did not show any detectable level of *NtHMA α* expression (data not shown). Finally, expression of *NtHMA α* could also be observed in anthers (Fig. 3G). According to the analysis of transversal sections, the coloration was specific to the anther itself and was not found in the pollen (Fig. 3H).

This localization of the expression of *NtHMA α* is similar to what has been obtained for *AtHMA4*¹⁶ and *AhHMA4*⁹, at least as far as roots and leaves are concerned.

Assessing the role of *NtHMA α* and *NtHMA β* in planta

In order to assess the role of the *NtHMA* genes in tobacco, two strategies were employed to either silence or reduce the activity of *NtHMAs*. The first strategy aimed at identifying tobacco lines harboring a mutation in either one of the *NtHMA* genes. This was achieved by screening an EMS (Ethyl methanesulfonate)-mutagenized mutant collection using CE-SSCP (Capillary Electrophoresis-Single Strand Conformation Polymorphism)³¹, a PCR-based technique. Several missense mutations as well as one nonsense mutation were detected in each of the two genes (Supplementary Table S1). All these mutations altered amino acids present in the third cytosolic loop of the HMA protein. The second strategy aimed at silencing the *NtHMA* genes by expressing artificial micro RNAs (amiRNAs)³² in transgenic tobacco plants. The constructs were designed either to specifically inactivate one of the *NtHMA* transcripts, or to simultaneously

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

inactivate the two transcripts (Sup Table S1). Up to 10 independent transformants were regenerated for each of the amiRNA constructs. For each construct, the three transgenic lines showing the highest and most specific reduction of *NtHMA* transcript level in roots were selected using quantitative PCR. None of the transgenic lines showed a complete silencing of the *NtHMA* genes. The transgenic plants expressing the amiRNA construct designed to target both *NtHMA α* and *NtHMA β* displayed a high degree of silencing of both genes (Sup. Fig. S3). With the amiRNA construct designed to specifically target one of the *NtHMA* (e.g. *NtHMA α*), it was possible to reduce the level of the targeted transcripts without significantly altering the level of the other transcripts (in the case of *NtHMA β* , $P=0.35$).

The growth and development of all the mutant and silenced lines in which only one of the *NtHMA* genes was targeted was comparable to that of the wild type. In contrast, the lines in which the two *NtHMA* genes were silenced displayed top leaves that were slightly distorted and capsules that were significantly smaller (Sup. fig. S4A and S4B). Lines with a nonsense mutation in both *NtHMA α* and *NtHMA β* were obtained by crossing. In the F2 progeny, a shift in the segregation was observed with fewer double homozygous mutants than expected (data not shown). In addition, plants harboring a nonsense mutation in both *NtHMA* genes displayed a dwarf phenotype and could not complete their life cycle (Sup. fig. S4C). This shows that a minimum level of *NtHMA* activity is needed for plant survival and that *NtHMA α* and *NtHMA β* are functionally redundant.

***NtHMA α* and *NtHMA β* are redundantly involved in the translocation of Cd from the roots to the shoots**

When grown on MS plates in the presence of 1 μ M Cd, both the *HMA*-silenced lines and the mutant lines exhibited ~50% reduction in Cd accumulation in the shoots (Fig. 4A) suggesting that the missense and nonsense mutations had a similar impact on the function of the protein. The greatest reduction of shoots Cd concentration (~80%) was obtained when the two *NtHMA* genes were silenced. These results clearly indicate that the

1
2
3 two *NtHMA* genes play a role in Cd accumulation in the shoots and that they
4 are at least partially redundant for this function.
5
6

7 ***NtHMA*-silenced lines and *Nthma* mutant lines are more tolerant to Cd** 8 **than the wild type** 9

10
11 The contribution of the *NtHMA* genes to Cd and Zn tolerance was assessed
12 by growing the silenced lines as well as the EMS-induced mutant lines on
13 plates in MS media supplemented by up to 400 μ M of Cd or Zn. The silenced
14 and mutant lines responded as the wild type to the presence of high
15 concentrations of Zn in the medium (data not shown). In contrast, they were
16 more tolerant to Cd than the wild type, as deduced from the analysis of root
17 growth (Fig. 5). Interestingly, lines in which both of the *NtHMA* genes were
18 silenced displayed even longer roots than lines in which only one of the
19 *NtHMA* genes was inactivated (Fig. 5A). This again suggests a partial
20 functional redundancy between *NtHMA α* and *NtHMA β* . In response to Cd,
21 leaves of the silenced and mutant plants were less chlorotic than leaves of
22 wild-type plants. This was particularly clear for the line in which the two
23 *NtHMA* genes were silenced (Fig. 5B) thus suggesting that protecting the
24 shoots from Cd over-accumulation increased Cd tolerance in tobacco.
25
26
27
28
29
30
31
32
33
34
35
36
37

38 **Contribution of *NtHMA α* and *NtHMA β* to the accumulation of Zn and Fe** 39

40 Zn concentration was analyzed in the *NtHMA*-silenced plants. Silencing
41 each of the *NtHMA* genes as well as silencing both copies resulted in
42 decreased Zn concentration in the shoots (Fig. 4B).
43
44
45

46 The *NtHMA α* /*NtHMA β* -silenced lines exhibited a reduced number of seeds
47 per capsule and a reduced ability of the seeds to germinate (Fig. 6A and
48 6B). To check whether this difference was correlated with metal
49 concentration and/or metal compartmentalization changes in the seed,
50 elemental composition was analyzed using synchrotron X-ray fluorescence
51 mapping (SXRF) and Inductively coupled plasma-mass spectroscopy (ICP-
52 MS) (Sup. fig. S5).
53
54
55
56
57
58
59

60 For the SXRF analysis, two sets of seeds originating from two different
capsules were analysed. Only the first set of seeds appeared to contain

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

mature embryos (Sup. fig. S5A) suggesting that part of the seeds expressing the amiRNA construct targeting both *NtHMA* genes are unable to mature properly. From the SXRF analysis of mature seeds, the distribution of Zn and Fe within the seed seemed to be comparable for the *NtHMA α /NtHMA β* -silenced seeds and the wild type. In particular, Fe was distributed mainly in the cotyledons (66% of total embryo Fe concentration) compared to in the radicle (34% of total embryo Fe concentration). This distribution differs from what had been observed for *Arabidopsis* seeds in which approximately 60% of the seed Fe is found within organelles of the endodermal cells of the radicle and around the vasculature of the cotyledons.³³

Quantification of the fluorescence signal from the SXRF analysis indicated that zinc concentration is reduced in the silenced line compared to wild type whereas iron concentration is increased (Sup. fig. S5B). We were able to confirm the increase in iron concentration using ICP-MS analysis (Sup. fig. S5C) but results were inconsistent for zinc as ICP-MS analysis suggested an increase in zinc concentration (Sup. fig. S5D). One possible explanation could be that SXRF allowed us to selectively analyse mature seeds expressing the amiRNA construct whereas ICP-MS analysis was performed on approximately 100mg of a mix of mature and aborted seeds, some of the latter being able to accumulate very high amounts of zinc with respect to their reduced mass (Sup. fig. S5A, S5B).

Pollen germination is reduced in *NtHMA*-silenced lines

As mentioned above, plants harboring nonsense mutations in both *NtHMA α* and *NtHMA β* were dwarf when grown on soil (Sup. fig. S4A) and were unable to produce seed. In contrast, transgenic plants expressing amiRNA constructs leading to a reduction of the expression of both *NtHMA* genes were still able to grow and produce seed, although the seed production was very low and occurred only when pistils were hand-pollinated (Fig. 6A). The phenotypic difference between the *Nthma α /Nthma β* double mutant lines and the *NtHMA α /NtHMA β* -silenced lines is likely due to the fact that the function of the *NtHMA* genes is most likely entirely lost in the double mutant lines, whereas expression of *NtHMA α* and *NtHMA β* was not completely abolished

1
2
3 in the *HMA*-silenced lines (Sup. fig. S3). Partial expression of functional
4 *HMA* gene is therefore required and sufficient for plant development.
5
6

7
8 The number of seeds in capsules of the *NtHMA α /NtHMA β* -silenced lines
9 was greatly reduced compared to the wild type or to lines in which only one
10 of the two *NtHMA* genes was silenced (Fig. 6A). This was caused by a
11 defect of the pollen. Indeed, pollinating the *NtHMA α /NtHMA β* -silenced lines
12 with wild-type pollen led to the production of capsules of a normal size
13 harboring a normal quantity of seeds while pollinating wild-type pistils with
14 pollen from *NtHMA α /NtHMA β* -silenced lines resulted in the production of
15 capsules of very small size harboring a reduced number of seeds (Sup fig
16 S4B). The origin of the pollen defect was neither the amount of pollen nor its
17 viability, which were both similar for the *NtHMA α /NtHMA β* -silenced lines and
18 the wild type (Fig. 7A).
19
20
21
22
23
24
25
26
27

28 Zn concentration was reduced in the anthers of *NtHMA α /NtHMA β* -silenced
29 lines (Fig. 7B) and the pollen exhibited a reduced ability to germinate
30 compared to the wild-type pollen (Fig. 7C). This reduced ability to germinate
31 could at least partially be alleviated when Zn was added to the germination
32 medium (Fig. 7D), thus indicating that Zn is required for pollen germination.
33 These results can be connected to the observation that Zn accumulation
34 was reduced in anthers of the *NtHMA α /NtHMA β* -silenced lines compared to
35 in wild-type anthers (Fig. 7B), which may explain the Zn deficiency of pollen
36 produced by *NtHMA α /NtHMA β* -silenced lines.
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Discussion

In this work, we were interested in identifying and characterizing genes controlling Cd accumulation in leaves of tobacco plants. Attention was drawn to Zn-Cd Heavy Metal ATPases, which have previously been shown to control this character in *A. thaliana* and rice.^{8,34,35} Two tobacco orthologs to *AtHMA2* and *AtHMA4*, *NtHMA α* and *NtHMA β* , were newly identified. As detailed above, our results strongly suggest that *NtHMA α* and *NtHMA β* are the only two *N. tabacum* orthologs to *AtHMA2* and *AtHMA4*. This is not surprising since only one ortholog to the sub-cluster formed by *AtHMA2*, *AtHMA3* and *AtHMA4* has been found in the sequenced genomes of diploid Solanaceae such as *S. lycopersicum*, *N. benthamiana*, *N. sylvestris* and *N. tomentosiformis*, the latter two being the parents of the *N. tabacum* amphiploid. Interestingly, the *NtHMA α* and *NtHMA β* copies could be tracked back in the *N. sylvestris* and *N. tomentosiformis* genomes, clearly supporting the hypothesis that each of them was originating from one of these parents.

NtHMA α* and *NtHMA β* are likely to share similar functions as *AtHMA2* and *AtHMA4

Although the phylogenetic analysis showed that the tobacco *HMA*s were located in a different phylogenetic clade compared to the *A. thaliana* or rice *HMA*s, phenotypic data indicated that their role is similar to the role of *AtHMA2*, *AtHMA4* or *OsHMA2*. *Nthma α* /*Nthma β* double mutant plants displayed a similar dwarf phenotype as *Athma2*/*Athma4* double mutant plants. The expression pattern of the tobacco genes was very similar to that of the *A. thaliana* ones. More importantly, mutations in *NtHMA α* or *NtHMA β* led to reduced accumulations of Cd in shoots in a similar manner to what has been described for *AtHMA2*, *AtHMA4* or *OsHMA2*. Altogether, these results suggest that *NtHMA α* and *NtHMA β* control Cd (and Zn) translocation from the roots to the shoot in tobacco and that this function would be a result of their contribution to the loading/unloading of Cd (and Zn) in the xylem vessels. It has already been proposed that in tobacco, roots control the amount of Cd translocated to the shoot and that xylem loading is the limiting step in this regard.³⁶

***NtHMA α* and *NtHMA β* share partially redundant functions**

N. tabacum is a 0.2-Myr-old allopolyploid ($2n = 4x = 48$) resulting from the fusion of *N. sylvestris* ($2n = 2x = 24$) and *N. tomentosiformis* ($2n = 2x = 24$) and most of its genes are still present in two copies.²⁵ The discovery that two orthologs to *AtHMA2* and *AtHMA4* are present in the genome of tobacco was therefore not surprising. It was however not obvious that those two genes would share the same function. Evolution may have modified or suppressed the function of one or both *NtHMA* genes. This was especially relevant in our case considering for instance that the footprint of a transposable element was present in the promoter region of *NtHMA β* . However, our results showed a high degree of functional redundancy between the two genes. The most striking phenotype illustrating a complete functional redundancy was that single mutant plants grew and produced seeds like normal plants, while plants harboring nonsense mutation in each of the two *NtHMA* genes could neither develop nor survive. When the expression of both *NtHMA* genes was strongly reduced but not completely abolished, plant development and reproduction could occur, but pollen germination was significantly reduced. Functional redundancy could also be observed for other traits, but in these cases, the redundancy was partial. For example, mutation in only one of the *NtHMA* genes already resulted in reduced Cd concentrations in leaves. Partial functional redundancy was also evidenced while studying Cd tolerance of the *NtHMA* mutant plants. Altogether, it is therefore clear that the two *NtHMA* genes share partially redundant functions, showing additive effect for the traits for which they are not fully redundant. In particular, inactivating only one of the two genes is not detrimental for the appropriate development of the plant. This is a very similar situation as for the redundancy existing between *AtHMA2* and *AtHMA4*.

Development of tobacco lines with reduced Cd accumulation in the leaves

One of the objectives of this work was to develop tobacco lines with reduced Cd concentration in shoots by inactivating *HMA* genes. Due to the difficulties

1
2
3
4 of growing and commercializing genetically modified crops in Europe, we
5 chose to work with mutants featuring single nucleotide polymorphisms
6 obtained by ethyl methanesulfonate treatment.³¹ A similar approach targeting
7 *OsNramp5* recently permitted the development of non-transgenic rice plants
8 with reduced Cd accumulation in the shoots.³⁷ Here we have demonstrated
9 the successful creation of non-GMO tobacco lines that accumulate less Cd
10 in the shoots than wild-type plants do.

11
12
13
14
15
16
17 Given the severe phenotypes of the plants harboring nonsense mutations in
18 each of the two *NtHMA* genes or in which both genes are silenced, it is clear
19 that the function of *NtHMA α/β* needs to be at least partially preserved for
20 normal growth and development. This indicates that engineering
21 commercially valuable tobacco plants with no Cd in the shoots cannot be
22 achieved through the sole inactivation of the two *NtHMA* genes. However,
23 inactivating only one of the *NtHMA* genes could lead to commercially
24 valuable plants that display a 50% reduction of Cd concentration in shoots.
25 To further decrease the ability of the plant to accumulate Cd in the leaves,
26 an appropriate strategy could be to combine different mutations in *NtHMA*
27 genes together. For example, one could combine a nonsense mutation in
28 one *NtHMA* gene with the use of a leaky mutation in the other *NtHMA* gene.
29 The aim would be to screen for the appropriate combination that would lead
30 to plants displaying the lowest ability to accumulate Cd in the leaves but
31 which would still develop like wild-type plants. Another and maybe more
32 fruitful solution could be to combine mutations in *NtHMA* genes with
33 mutations in other genes such as *NRAMP5* or *IRT1*, which play a role in Cd
34 transport at different steps than *NtHMAs*.³⁶ Finally, as Cd is mainly
35 transported by proteins, which main function is to transport essential metals
36 such as Zn or Fe, the best solution to develop low-cadmium accumulating
37 crops would be to identify mutations that would selectively reduce the
38 transport of Cd without altering the transport of other elements. To-date, this
39 kind of mutation has still not been reported.

40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58 ***NtHMA* genes play a role in pollen germination and seed development**
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

To date, Zn and Cd ATPases have been mainly described for their role in Zn/Cd loading into the xylem or for their role in Cd/Pb loading into root vacuoles.³⁸ We have shown here that *NtHMA α* and *NtHMA β* are also important for pollen development, seed number and seed viability. *NtHMA α* /*NtHMA β* -silenced plants produce pollen that cannot germinate unless Zn is added to the germination media suggesting that *NtHMA α* and *NtHMA β* play a role in zinc uptake. Interestingly, *NtHMA α* was shown to be expressed in anthers and might thus be responsible for the loading of zinc into the pollen sac. The importance of Zn for the pollen has already been reported in lentil.³⁹ Indeed, when grown with low levels of Zn lentils had a reduced number of seeds and showed reduced pollen and seed viability as well as inhibition of pollen tube growth. However, no gene/mechanism underlying this phenotypic observation had been uncovered. In *A. thaliana*, inactivating *AtHMA2* and *AtHMA4* resulted in reduced fertility¹⁵ but again no precise mechanism was proposed to explain this phenotype. In this context, our results clearly demonstrate the direct impact of Zn on pollen germination and propose *NtHMA* genes as being important contributors to the feeding of pollen with zinc.

NtHMA genes were also shown to be important for seed quality and germination. This is likely due to a modified concentration of metals in seeds but additional experiments are needed to analyse this phenotype in detail.

***NtHMA α* and *NtHMA β* possess special cytoplasmic tails**

Plant Cd²⁺/Zn²⁺-ATPases have long C-terminal sequences rich in cysteines and histidines. Work performed on *AtHMA2* showed that its C-terminal tail plays a role in the turn-over and enzyme velocity of the ATPase and that cysteines and histidines are important for the binding of zinc and cadmium^{8,26}. The C-terminal ends of the newly identified *NtHMA α* and *NtHMA β* are different from those of *AtHMA2* and *AtHMA4*. They have twice more cysteines present in tandem, but lack the stretch of histidines. They are also between 2 and 3 times longer than the cytoplasmic tails of *AtHMA2* and *AtHMA4*. Interestingly, these characteristics are also shared by the cytoplasmic tails of the orthologs of *NtHMAs* in Solanaceae. An exciting

1
2
3 working hypothesis would be that the longer cytoplasmic tail of NtHMA
4 proteins may enhance the capacity of plant to load cadmium into the xylem
5 and explain why tobacco, unlike most other plant species, translocates most
6 of the cadmium to its shoots instead of storing it in roots.⁷
7
8
9

10 11 **Conclusion**

12
13 Our study is a good example of how the fundamental work performed on
14 *Arabidopsis thaliana* can be used for crops with agronomic or industrial
15 interest. Using the sequences of *AtHMA2* and *AtHMA4*, two *HMA* genes
16 could be identified in tobacco. These genes have proven to be good targets
17 to reduce Cd accumulation in shoots. Non-GMO lines with significantly
18 reduced Cd accumulation in shoots were obtained. Like *AtHMA2* and
19 *AtHMA4*, *NtHMA α* and *NtHMA β* are involved in both Cd and Zn distribution
20 in the plant. In this regard, the function of *NtHMA α* and *NtHMA β* seems to be
21 redundant. A particularly novel result is that the *NtHMA α/β* genes seem to be
22 essential for the transport of Zn into the pollen grain and for its germination.
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Material and methods

Plant material and growth conditions

The wild-type tobacco plants BB16NN and BY212 as well as the mutated lines were provided by the tobacco institute of Bergerac (France). BB16NN and BY212 are two modern accessions of Burley tobacco and have both been used to generate the collection of mutagenized lines. To identify lines harboring EMS-induced mutations in the *NtHMA* genes, regions encoding the third cytosolic loop of the HMA protein were amplified by PCR from genomic DNA extracted from pools of the EMS-mutagenized lines using gene specific primers (Supplementary table S1). Mutations present in these amplified products were identified by CE-SSCP (Capillary Electrophoresis-Single Strand Conformation Polymorphism).³¹ To select mutant lines, only missense mutations occurring at highly conserved positions among *HMA* genes were considered and the CODDLE⁴⁰ and SIFT⁴¹ bioinformatics prediction softwares were used to determine which of those mutations were the most likely to alter the function of the protein. For *NtHMA α* the selected missense mutations were P294S and E387K and the non-sense mutation was W265*. For *NtHMA β* the selected missense mutation was G515R and the non-sense mutation was R529* (Supplementary table S1). All the mutant lines were backcrossed twice with their respective wild type (BB16NN or BY212) except the line harboring the non-sense mutation in *NtHMA β* which was only backcrossed once.

In the greenhouse plants were grown in 1.5 liter pots. Light was controlled for a minimum of 16 hours a day; artificial light was used when natural luminosity was below 700 lux.

In hydroponics plants were grown at 20°C, 70% relative humidity and a 16:8 h light:dark cycle with a light intensity of 150 mol.m⁻².s⁻¹ in 2.5 mM KNO₃, 0.5 mM NaH₂PO₄, 2.5 mM Ca(NO₃)₂, 0.5 mM MgSO₄, 0.1 mM Fe_{III}NaEDTA, 0.05 mM H₃BO₃, 0.05 mM MnSO₄, 15 μ M ZnSO₄, 3 μ M Na₂MoO₄, 2.5 μ M KI, 0.05 μ M CuSO₄ and 0.044 μ M CoCl₂. Nutrient solutions were renewed every 3 to 4 days over the course of the entire experiment.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Plants were also grown in sterile conditions on plates containing Murashige and Skoog medium (MS, Duchefa) solidified with 1% (w/v) agar and buffered with 0.5 g.L⁻¹ MES-KOH pH 5.7. The media were supplemented with various Cd (CdCl₂) or Zn (ZnCl₂) concentrations. Growing conditions were 21°C and 16 hours of light. Root lengths were measured with ImageJ using the NeuronJ plugin.

Cloning and sequencing of *NtHMAα* and *NtHMAβ* genes

Tobacco sequences similar to the *Arabidopsis thaliana* *AtHMA2* and *AtHMA4* sequences were identified in a library of tobacco BAC end sequences (compsysbio.achs.virginia.edu/tobfac). These sequences corresponded to short fragments of the *NtHMAα* and *NtHMAβ* genes. From these sequences primers were designed to amplify the *NtHMAα* and *NtHMAβ* genes from the BB16NN genomic DNA. Amplified fragments were used as probes to screen by Southern analysis a tobacco genomic BAC library that was made at CNRGV (reference NTA-B-BB16NN available at http://cnrgv.toulouse.inra.fr/library/genomic_resource/Nta-B-BB16NN). The presence of either *NtHMAα* or *NtHMAβ* was assessed using PCR on the different BAC clones that were identified. For each of the *NtHMA* genes one BAC clone was sequenced (accession numbers HF675180 and HF937053 for the BAC clones harboring *NtHMAα* and *NtHMAβ*, respectively). RT-PCR amplification was performed from BB16NN root cDNA and sequencing of the corresponding fragments led to the identification of the complete coding sequences of the *NtHMA* genes (accession numbers HF675181 and HF937054 for the cDNAs encoding *NtHMAα* and *NtHMAβ*, respectively). The primers used during this work are listed in supplementary table S2.

Phylogenetic analysis

Phylogenetic analyses were performed using Seaview (pbil.univ-lyon1.fr/software/seaview.html).⁴² HMA protein sequences from *Arabidopsis* and their identified orthologs in other species (rice, grape, *A. lyrata*, sorghum, *A. halleri*, tomato and *N. benthamiana*) were retrieved from online databases: AhHMA3 (AJ556182) and AhHMA4 (EU382072), AIHMA1 (XM_002866914), AIHMA2 (XM_002867321), AIHMA3 (XM_002867320),

1
2
3 AtHMA1 (NM_119890), AtHMA2 (AY434728), AtHMA3 (AY055217),
4 AtHMA4 (AF412407), AtHMA5 (Q9SH30), AtHMA6 (Q9SZC9), AtHMA7
5 (Q9S7J8), AtHMA8 (B9DFX7), MtHMA-A (AC130275), MtHMA-B(135313)
6 NbenHMA4 (Nbv3K585727037), OsHMA1 (NM_001064952), OsHMA2
7 (HQ646362), OsHMA3 (AB557931), SbHMA2 (XM_002438908), SbHMA3-1
8 (XM_002459533), SbHMA3-2 (XM_002459534), SIHMA4 (XP_004242843),
9 VvHMA1 (XM_002278513). TBLASTX analysis led to the identification of the
10 sequences of the SIHMA (AEKE02004109), NsyIHMA (ASAF01121754),
11 NtomHMA (ASAG01036389 and ASAG01036390) and NbHMA homologs.
12 The 9 OsHMA protein sequences were obtained from the supplementary
13 data of Baxter *et al.*⁴³ The first parts of the proteins, containing all the
14 transmembrane domains, were aligned using the muscle algorithm and the
15 phylogenetic tree was constructed using the Maximum Likelihood method
16 from G-blocks distributed along the corresponding sequence of the proteins.
17 Bootstrap values were obtained using 1000 replicates.

30 31 **Obtention of *N. tabacum* lines harbouring amiRNA constructs targeting** 32 **the *NtHMA* genes**

33
34
35 AmiRNA constructs were made according to the guidelines detailed in
36 Ossowski *et al.*³², using the online resource associated to this publication
37 (wmd3.weigelworld.org). Two constructs were designed to silence
38 specifically *NtHMA α* , two constructs were designed to silence specifically
39 *NtHMA β* and one construct was designed to silence both genes
40 (Supplementary Table S1 and Supplementary Table S2). The amiRNA
41 sequences were compared with the available tobacco sequences to check
42 that they could not target any other sequence than the *NtHMA* sequences.
43 The amiRNA constructs were introduced into the pGREEN⁴⁴ vector and then
44 in the GV3101 *Agrobacterium* strain, together with the pSOUP plasmid.
45 Transgenic tobacco lines harboring the amiRNA constructs were produced
46 in the BB16NN background using a standard leaf disk transformation
47 procedure.⁴⁵ Quantitative PCR was used on homozygous lines to select the
48 transgenic lines in which *NtHMA* genes were the most specifically and
49 efficiently silenced. At least 3 independent transgenic plants were selected
50 for each construct.
51
52
53
54
55
56
57
58
59
60

RNA Preparation and Quantitative RT-PCR

Total RNA was extracted using the RNeasy kit (Qiagen). The cDNAs were prepared from 1 µg of total RNA using M-MLV Reverse Transcriptase RNase H minus (Promega) and oligo(dT) 15 primers (Promega) following manufacturer's instructions. To analyze transcript accumulation in wild-type and *HMA*-silenced plants gene-specific primers were designed for real time quantitative PCR (Supplementary Table S2). To warrant the gene copy specificity of the amplification, plasmids containing either *NtHMA α* or *NtHMA β* were used as controls in every experiment. Real-time quantitative RT-PCR was performed using a LightCycler480 with Fast-Start DNA MasterPLUS SYBER GREEN I (Roche Applied Science, www.roche-applied-science.com). Analysis of relative transcript accumulation was done using LightCycler480 SW1.5. The purity of the amplicons was assessed for every sample by analyzing the amplicon dissociation during the gradual increase of the temperature from 72°C to 95°C at the rate of 0.11°C/sec, using the T_m calling method proposed by the LightCycler480 software. Expression level of *NtHMA α* and *NtHMA β* was determined relative to the already described housekeeping genes cyclophilin and ribosomal protein L2 (accession numbers X62500 and Z14081, respectively).⁴⁶ Real time quantitative RT-PCR was performed using at least three independent biological replicates and three technical replicates of each biological sample. The student t test used to determine different statistical groups was performed with the open software R (R Development Core Team).

Metal concentration measurements

Preliminary tests showed that metal extraction in HCl and mineralization gave similar results to assess metal concentration (data not shown). We therefor used the former, less hazardous, procedure. Plant samples were dried at 75°C for 48 h before measurement of dry weight. Dried plant samples were incubated in 1N HCl at 80°C for 30 min to extract cations. Concentrations of Cd, Zn or Fe were determined by ICP-MS (ICP-MS 7500cx, agilent technologies, reporting limit of 0.2 ng/ml and standard

1
2
3 recovery of 5 µg/L). Analyses were performed on individual plants using at
4
5 least 3 plants per genotype.
6
7

8 **GUS staining experiments**

9
10 A 3-kb-long PCR-amplified (for primer see Supplementary Table S2)
11 fragment corresponding to the promoter of *NtHMAα* was fused upstream the
12 *gusA* coding sequence into the pCambia1391 vector. This construct was
13 introduced into the *A. tumefaciens* GV3101 strain and then into tobacco as
14 described above. GUS staining was performed as described.⁴⁷ The staining
15 solution contained 50 mM NaPO₄ pH7, 0.5 mM ferricyanide, 0.5 mM
16 ferrocyanide, 0.05 % triton, 1 mM X-Gluc. Coloration appeared after two
17 hours of incubation at 37°C in the staining solution. Chlorophyll was removed
18 from leaf explants by successive 10-min-long incubations in solutions with
19 increasing EtOH concentrations (50%, 70%, 90% and 100%). After complete
20 chlorophyll removal, tissues were rehydrated thanks to successive 10-min-
21 long incubations in solutions with decreasing EtOH concentrations (90%,
22 70% and 50%).
23
24
25
26
27
28
29
30
31
32
33

34 **Analyses of pollen**

35
36 For each plant, pollen grains were collected only from the first three opening
37 flowers. Pollen was collected while flowers were opening for the first time, to
38 avoid pollen loss or dehiscence. Pollen grains were suspended in water and
39 counted using the Nageotte hemocytometer (Grosseron, ref. 0703614).
40 Pollen viability was assessed using the Alexander stain⁴⁸. Pollen germination
41 was performed as described⁴⁷ using a buffer containing 1.6 mM H₃BO₃, 200
42 µM CaCl₂, 6 % sucrose and 1 mM morpholine ethanesulfonic acid, pH 5.7.
43 Pollen was incubated at 20°C over 3 hours under weak agitation in the dark
44 before the reaction was stopped. Pollen germination rate was calculated
45 using 100 pollen grains from each flower.
46
47
48
49
50
51
52
53
54

55 **Seed germination**

56
57 For each genotype, the capsules of at least three plants were collected and
58 at least three capsules per plants were harvested. To determine the number
59 of seeds per capsules, 200 seeds per genotype were weighted. The total
60

1
2
3 number of seeds in a given capsule was then deduced from the weight of all
4 the seeds in this capsule.
5
6

7
8 Plants were sown on plates containing MS media to determine the
9 germination efficiency (number of germinated plants divided by total number
10 of seeds x100).
11
12

13 **SXRF Microspectroscopy of Intact Seed**

14
15
16 Elemental maps of dry, mature seed were collected at the bending magnet
17 beamline BL2-3 at the Stanford Synchrotron Radiation Lightsource (Palo
18 Alto, CA). This beamline uses Kirkpatrick-Baez optics to achieve optimized
19 beam dimensions of 2 μm in both horizontal and vertical planes. Detailed
20 SXRF mapping of tobacco seeds was performed with a 7 μm x 7 μm step
21 size and a 50 ms dwell time using a silicon (111) monochromator for energy
22 selection, and a Vortex (SII) detector. The rapid flyscanning capabilities of
23 this beam line preclude the need for intrusive sample preservation.
24
25 Elemental maps were collected at an energy of 11 keV. SXRF images were
26 quantified and analyzed using Microanalysis Toolkit.⁵⁰ Reference
27 fluorescence maps were collected from a suite of known-abundance thin-film
28 standards for fluorescence count quantification. Arrays of seed were placed
29 on metal-free Kapton™ double-sided tape, maintaining identical orientation
30 for each. Four replicate seeds of each line were included in the analysis,
31 with same-batch wild type seeds included in the array. During quantification
32 the effect of sample thickness on fluorescence yield was accounted for by
33 measuring X-ray absorption with the p-type n-type intrinsic diode
34 downstream of the sample. Region-of-interest analysis (masking) isolated
35 individual seeds for statistical analysis (mean intensity, total intensity and
36 area) and subtraction of background fluorescence.
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Fundings

This work was supported by an INRA fellowship to VH. Part of this work was also supported by grants from l'Association de Recherche sur les Nicotianées and from the National Institute of Environmental Health Sciences and the Environmental Protection Agency: Superfund Research Program (grant No. P42 ES007373-14), Children's Environmental Health and Disease Prevention (EPA RD-83459901 and P20 ES018175) to TP.

Supplementary data

Supplementary figure S1: Comparaison of *NtHMA* genes from different species.

Supplementary figure S2: Conservation of three HMA motifs.

Supplementary figure S3: Accumulation of NtHMA α and NtHMA β transcripts in transgenic lines expressing amiRNA constructs targeting NtHMA α and/or NtHMA β .

Supplementary figure S4: Inactivating or silencing the two NtHMA genes modifies development and fertility.

Supplementary figure S5: SXRF Fe and Zn quantitative analysis of seeds.

Supplementary table S1: List of genetically modified lines used.

Supplementary table S2: List of primer used.

Acknowledgments

We thank A. Adiveze, H. Afonso, C. Ausseil, C. Baracco, H. Baudot, C. Dasen, T. Dessup, X. Dumont, C. Fizames, J. Garcia, S. Gélín, F. Lecocq, C. Muller, V. Papy, V. Rafin, G. Ruiz and C. Zicler, for technical and administrative supports as well as J. Danku and D. Salt for ICP-MS analysis of seed metal content. Portions of this research were carried out at the Stanford Synchrotron Radiation Lightsource, a Directorate of SLAC National Accelerator Laboratory and an Office of Science User Facility operated for the U.S. Department of Energy Office of Science by Stanford University.

References

- 1 **Broadley MR, White PJ, Hammond P, Zelko I, Lux A**, *The New Phytologist*, 2007, **173**, 677-702
- 2 **Clemens S, Palmgren MG, Krämer U**, *Trends in Plant Science*, 2002, **7**, 309-315
- 3 **Dguimi HM, Debouba M, Ghorbel MH, Gouia H**, *Comptes Rendus de Biologies*, 2008, **332**, 58-68
- 4 **Verbruggen N, Hermans C, Schat H**, *Current Opinion in Plant Biology*, 2009, **12**, 364-372
- 5 **Järup L, Akesson A**, Current status of Cd as an environmental health problem, *Toxicology and Applied Pharmacology*, 2009, **238**, 201-208
- 6 **Godt J, Scheidig F, Grosse-Siestrup C, Esche V, Brandenburg P, Reich, Groneberg D**, *Journal of Occupational Medicine and Toxicology*, 2006, **1**, 22-27
- 7 **Maaroufi Dguimi H, Debouba M, Ghorbel MH, Gouia H**, *Comptes rendus de l'Académie des sciences*, 2009, **332**, 58-68
- 8 **Wong CK, Jarvis RS, Sherson SM, Cobbett CS**, *The New Phytologist*, 2009, **181**, 79-88
- 9 **Hanikenne M, Talke IN, Haydon MJ, Lanz C, Nolte A, Motte P, Kroymann J, Weigel D, Kramer U**, *Nature*, 2008, **453**, 391-395
- 10 **Korenkov V, King B, Hirschi K, Wagner GJ**, *Plant Biotechnology Journal*, 2009, **7**, 1-8
- 11 **Elmayan T, Tepfer M**, *The Plant Journal*, 1994, **6**, 433-440
- 12 **Dorlhac de Borne F, Elmayan T, de Roton C, de Hys L, Tepfer M**, *Molecular Breeding*, 1998, **4**, 83-90
- 13 **Gonzalez P, Baudrimont M, Boudou A, Bourdineaud JP**, *Biometals*, 2006, **3**, 225-235

- 1
2
3 **14 Connolly EL, Fett JP, Guerinot ML**, Plant Cell, 2002, **14**, 1347-1357
4
5
6 **15 Hussain D, Haydon MJ, Wang Y, Wong E, Sherson SM, Young J,**
7 **Camakaris J, Harper JF, Cobbett CS**, Plant Cell, 2004, **16**, 1327-1339
8
9
10 **16 Verret F, Gravot A, Auroy P, Leonhardt N, David P, Nussaume L,**
11 **Vavasseur A, Richaud P**, FEBS Letters, 2004, **576**, 306-312
12
13
14 **17 Korenkov V, Hirschi K, Crutchfield JD, Wagner GJ**, Planta, 2007, **226**,
15 1379-1387
16
17
18
19 **18 Morel M, Crouzet J, Gravot A, Auroy P, Leonhardt N, Vavasseur A,**
20 **Richaud P**, Plant physiology, 2009, **149**, 894-904
21
22
23 **19 Thomine S, Wang R, Ward JM, Crawford NM, Schroeder JI,**
24 **Proceedings of the National Academy of Sciences of the USA**, 2000, **97**,
25 4991-4996
26
27
28
29
30 **20 Oomen RJ, Wu J, Lelievre F, Blanchet S, Richaud P, Barbier-Brygoo**
31 **H, Aarts MG, Thomine S**, The New Phytologist, 2009, **181**, 637-650
32
33
34 **21 Huang J, Zhang Y, Peng JS, Zhong C, Yi HY, Ow DW, Gong JM**, Plant
35 **Physiology**, 2012, **158**, 1779-1788
36
37
38
39 **22 Takahashi R, Bashir K, Ishimaru Y, Nishizawa NK, Nakanishi H**, Plant
40 **signaling and behavior**, 2012, **12**, 1605-1607
41
42
43 **23 Yamaji N, Xia J, Mitani-Ueno N, Yokosho K, Feng Ma J**, Plant
44 **physiology**, 2013, **162**, 927-939
45
46
47
48 **24 Yukawa M, Tsudzuki T, Sugiura M**, Molecular Genetics and Genomics,
49 2006, **275**, 367-373
50
51
52 **25 Clarkson JJ, Lim KY, Kovarik A, Chase MK, Knapp S, Leitch AR**, The
53 **New Phytologist**, 2005, **168**, 241-252
54
55
56 **26 Sierro N, Battey JN, Ouadi S, Bovet L, Goepfert S, Bakaher N,**
57 **Peitsch MC, Ivanov NV**, Genome biology, 2013, **14**, 1-17
58
59
60

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- 27 Naim F, Nakasugi K, Crowhurst RN, Hilario E, Zwart AB, Hellens RP, Taylor JM, Waterhouse PM, Wood GC**, PLoS One, 2012, **7**, e52717
- 28 Eren E, Kennedy DC, Maroney MJ, Argüello JM**, The Journal of Biological Chemistry, 2006, **281**, 33881-33891
- 29 Bernard C, Roosens N, Czernic P, Lebrun M, Verbruggen N**, FEBS letters, 2004, **569**, 140-148
- 30 Williams LE and Mills RF**, Trends in plant science, 2005, **10**, 491-502
- 31 Julio E, Laporte F, Reis S, Rothan C, Dorlhac de Borne F**, Molecular Breeding, 2008, **21**, 369-381
- 32 Ossowski S, Schwab R, Weigel D**, The Plant journal, 2008, **53**, 674–690
- 33 Punshon T, Hirschi K, Yang J, Lanzirotti A, Lai B, Guerinot ML**, Plant Physiology, 2012, **158**, 352-362
- 34 Ueno D, Milner MJ, Yamaji N, Yokosho K, Koyama E, Clemencia Zambrano M, Kaskie M, Ebbs S, Kochian LV, Ma JF**, The Plant journal, 2011, **66**, 852-862
- 35 Satoh-Nagasawa N, Mori M, Nakazawa N, Kawamoto T, Nagato Y, Sakurai K, Takahashi H, Watanabe A, Akagi H**, Plant & cell physiology, 2012, **53**, 213-224
- 36 Lugon-Moulin N, Zhang M, Gadani F, Rossi L, Koller D, Krauss M, G.L Wagner**, Advances in agronomy, 2004, **83**: 111-180
- 37 Ishikawa S, Ishimaru Y, Igura M, Kuramata M, Abe T, Senoura T, Hase Y, Arao T, Nishizawa NK, Nakanishi H**, Proceedings of the National Academy of Sciences o the USA, 2012, **109**, 19166-19171
- 38 Palmer CM, Guerinot ML**, Nature chemical biology, 2009, **5**, 333-340
- 39 Pandey N, Pathak GC, Sharma CP**, Journal of Trace Elements in Medicine and Biology, 2006, **20**, 89-96

1
2
3
4 **40 Till BJ, Reynolds SH, Greene EA, Codomo CA, Enns LC, Johnson**
5 **JE, Burtner C, Odden AR, Young K, Taylor NE, Henikoff JG, Comai L,**
6 **Henikoff S**, Genome Research, 2003, **13**, 524–530
7
8

9
10 **41 Ng PC, Henikoff S**, Nucleic Acid Research, 2003, **31**, 3812–3814
11

12 **42 Gouy M, Guindon S, Gascuel O**, Molecular Biology and Evolution,
13 2010, **27**, 221-224
14

15
16 **43 Baxter I, Tchieu J, Sussman MR, Boutry M, Palmgren MG, Gribskov**
17 **M, Harper JF, Axelsen KB**, Plant physiology, 2003, **132**: 618-628
18
19

20
21 **44 Hellens RP, Edwards EA, Leyland NR, Bean S, Mullineaux PM**, Plant
22 Molecular Biology, 2000, **42**, 819-832
23
24

25
26 **45 Horsch RB, Fraley RT, Rogers SG, Sanders PR, Lloyd A, Hoffmann**
27 **N**, Science, 1984, **223**, 496-498
28
29

30
31 **46 Nicot N, Hausman JF, Hoffmann L, Evers D**, Journal of Experimental
32 Botany, 2005, **56**, 2907-2914
33

34
35 **47 Jefferson RA, Burgess SM, and Hirsh D**, Proceedings of the National
36 Academy of Sciences, 1986, **83**, 8447-8451
37
38

39
40 **48 Alexander MP**, Stain Technology, 1969, **44**, 117-122
41

42
43 **49 Thorsness MK, Kandasamy MK, Nasrallah ME, Nasrallah JB**, Plant
44 Cell, 1993, **5**, 253-261
45

46
47 **50 Webb SM**, AIP Conference Proceeding, 2005, **1365**, 196-199
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure legends

Figure 1: Phylogenetic analysis of HMA genes. Protein sequences encoded by the first 8 exons were aligned using the Muscle algorithm and the phylogenetic tree was constructed using the Maximum Likelihood method from G-blocks distributed along the sequence. Bootstrap values were obtained using 1000 replicates and are indicated before each embranchment of the phylogenetic tree. The tree is drawn to scale, with branch length corresponding to the number of amino acid substitutions per site. Ah = *Aribopsis halleri*; Al = *Arabidopsis lyrata*; At = *Arabidopsis thaliana*; Mt = *Medicago truncatula*; Nben = *Nicotiana benthamiana*; Nsyl = *Nicotiana sylvestris*; Ntom = *Nicotiana tomentosiformis*; Os = *Oryza sativa*; Sb = *Solanum tuberosum*; Sl = *Solanum lycopersicum*; Vv = *Vitis vinifera*. HMA protein sequences from *Arabidopsis* and their identified orthologs in other species were retrieved from online databases: AhHMA3 (AJ556182) and AhHMA4 (EU382072), AIHMA1 (XM_002866914), AIHMA2 (XM_002867321), AIHMA3 (XM_002867320), AtHMA1 (NM_119890), AtHMA2 (AY434728), AtHMA3 (AY055217), AtHMA4 (AF412407), AtHMA5 (Q9SH30), AtHMA6 (Q9SZC9), AtHMA7 (Q9S7J8), AtHMA8 (B9DFX7), MtHMA-A (AC130275), MtHMA-B(135313) NbenHMA4 (Nbv3K585727037), OsHMA1 (NM_001064952), OsHMA2 (HQ646362), OsHMA3 (AB557931), SbHMA2 (XM_002438908), SbHMA3-1 (XM_002459533), SbHMA3-2 (XM_002459534), SIHMA4 (XP_004242843), VvHMA1 (XM_002278513). TBLASTX analysis led to the identification of the sequences of the SIHMA (AEKE02004109), NsylHMA (ASAF01121754), NtomHMA (ASAG01036389 and ASAG01036390) and NbenHMA homologs.

Figure 2: Accumulation of *NtHMA α* and *NtHMA β* transcripts. Transcript accumulation in roots and shoots of plants grown for five weeks in hydroponics in control conditions (A) or in roots of plants treated one week before sample collection with different CdCl₂ (B) or ZnCl₂ (C) supplementation. Transcript accumulation was determined for at least three independent plants using quantitative PCR with three technical replicates for each plant. *NtHMA α* and *NtHMA β* transcript levels were normalized by comparison with transcripts of the ribosomic protein L2 and cyclophilin. Error bars represent standard error and the letters represent different statistical groups based on a Student t test.

Figure 3: Spatial localization of the expression of *NtHMA α* . GUS expression was monitored in transgenic plants expressing the GUS reporter gene under the control of a sequence corresponding to a 3kb region upstream of the start codon of *NtHMA α* . Arrows point to where the staining was observed. (A-B-C-D) Blue staining observed in the root stele (A), and more precisely in the xylem parenchyma cells (B-C-D). (E-F) Blue staining

1
2
3
4
5
6
7
8
9
10
11
12
13
14
observed in leaf veins. The staining was observed in the primary vein when the leaves were young (~3 cm in length) (E) but it was only visible in tertiary veins when the leaves were getting older (F). (G-H) blue staining was also observed in anthers (G) but not in the pollen as shown in the longitudinal section (H). The bar present at the bottom left corner of each picture represents the scale. The images correspond to three-week-old plants cultivated on solidified MS media in Petri dishes except for anthers, which were obtained from plants grown in the greenhouse.

15
16
17
18
19
20
21
22
23
24
25
26
Figure 4: Cadmium and zinc accumulation in shoots. Cd (A) and Zn (B) accumulation in the shoot of five-week-old plants was measured by ICP-MS. Mutants (left) and silenced lines (right) with their corresponding control were grown on MS media in the presence of CdCl₂ at the final concentration of 1 μM in the culture medium. At least 3 independent plants were analyzed for each genotype. Error bars represent standard error. * : P < 0.05, compared to the wild type. WT = wild type; EV = transgenic plants harboring an empty vector.

27
28
29
30
31
32
33
34
35
Figure 5: Cadmium tolerance analysis. (A) Root length of four-week-old plants grown on plates on MS media in the presence of a CdCl₂ concentration of 100 μM. At least 5 plants were used for each genotype. Error bars represent standard error. ** = P < 0.005 compared to the wild type. (B) Picture of four-week-old wild-type and *NtHMAα/NtHMAβ*-silenced plants grown on a plate in the presence of 100 μM of CdCl₂.

36
37
38
39
40
41
42
43
44
45
46
47
48
49
Figure 6: Effect of silencing the *NtHMA* genes on seeds. (A) Number of seeds per capsule obtained after selfing. Wild-type plants and plants transformed with an empty vector are shown as controls in comparison to plants expressing an amiRNA construct targeting either *NtHMAα* or *NtHMAβ* or both genes. For plants in which both *NtHMA* genes were silenced, self-pollination was performed manually using the pollen of 3 different flowers from the same plant for no seed could be obtained with natural self-pollination. Five capsules from 5 different plants were used (B) Efficiency of germination of seeds originating from wild-type and *NtHMAα/NtHMAβ*-silenced plants (n=3). Error bars represent standard error.

50
51
52
53
54
55
56
57
58
59
60
Figure 7: Effect of *NtHMA* gene silencing on pollen. (A) Total number of pollen grains and number of dead pollen grains (as determined by Alexander's staining) per flower. (B) Zinc concentration in anthers of plants originating from different genotypes. (C) Pictures of pollen isolated from wild-type and *NtHMAα/NtHMAβ*-silenced plants taken after 5 hours of incubation in germination solution. (D) Pollen germination in media supplemented with 0 to 15 μM of ZnCl₂. (A-B-D) At least 3 different plants were analyzed for each line. For each plant, the anthers from 3 different flowers were collected.

1
2
3 Error bars represent standard errors. *: $P < 0.05$; **: $P < 0.005$, compared
4 to the wild type.
5
6

7 **Supplementary figure S1: Comparaison of NtHMA genes with AtHMA2,**
8 **AtHMA3 and AtHMA4.** (A) Intron-exon organization of NtHMA α , NtHMA β ,
9 AtHMA2, AtHMA3 and AtHMA4. Exons are represented by rectangles and
10 are numbered. Introns are not to scale. Shaded rectangles are well-
11 conserved exons among the 5 HMA genes (identity higher than 50% at the
12 amino acid level). White rectangles represent the remaining exons and
13 correspond to the part of the proteins, which encodes the C-terminal
14 cytoplasmic loop and is less conserved (identity lower than 50% at the
15 amino acid level). (B) Percentage of amino acid identity between the first 8
16 exons of NtHMA α , NtHMA β , AtHMA2, AtHMA3 and AtHMA4.
17
18
19
20

21 **Supplementary figure S2: Conservation of three HMA motifs.** Focus on
22 the protein sequence alignments corresponding to the regions harboring the
23 GICCTSE (metal binding domain, MBD), DKTGT (P-domain) and CPx/SPC
24 motifs, which are the signatures of the clade gathering Zn, Cd and Pb P_{1B}-
25 ATPases. Alignments were performed using Clustalw
26 (www.ebi.ac.uk/Tools/msa/clustalw2/). Black arrows indicate where the
27 motifs are in the alignments and red arrows indicate the two HMAs from
28 tobacco.
29
30
31
32

33 **Supplementary figure S3: Accumulation of NtHMA α and NtHMA β**
34 **transcripts in transgenic lines expressing amiRNA constructs targeting**
35 **NtHMA α and/or NtHMA β .** Transcript accumulation was determined in five-
36 week-old plants grown *in vitro*. Three independent transgenic lines were
37 analyzed for each construct and three plants were used for each transgenic
38 line. Transcript accumulation was determined using quantitative PCR with
39 three technical replicates. NtHMA α and NtHMA β transcript levels were
40 normalized by comparison with transcripts of the ribosomal protein L2 and
41 cyclophilin. Error bars represent standard. * : $P < 0.05$; ** : $P < 0.005$,
42 compared to the transcript level of the same gene in the wild type.
43
44
45
46
47

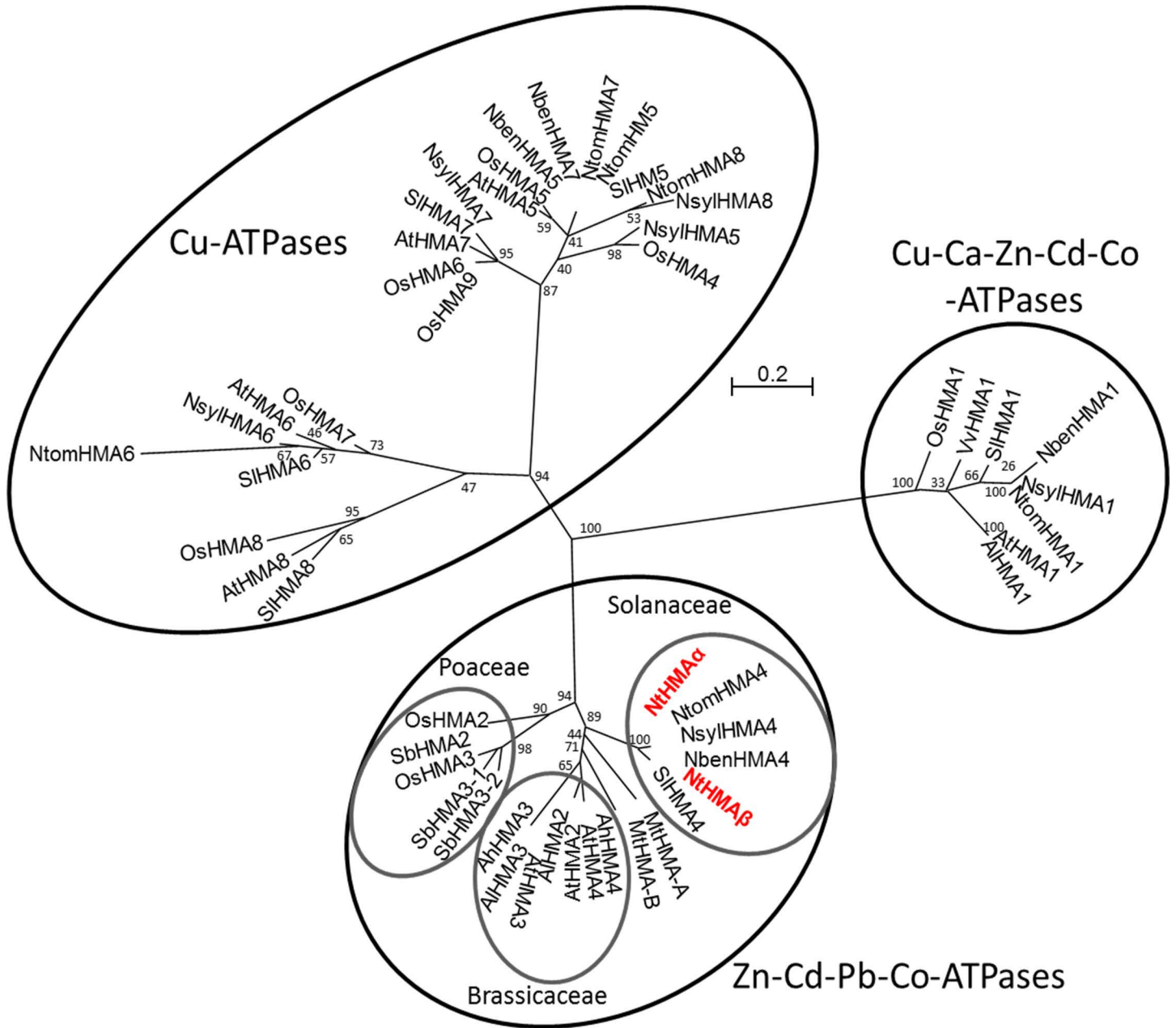
48 **Supplementary figure S4: Inactivating or silencing the two NtHMA**
49 **genes modifies development and fertility.** (A) Picture of a set of F2 plants
50 issued from the cross between two parents harboring each a nonsense
51 mutation in one of the NtHMA genes. Dwarf plants (circled in red) are the
52 only plants that harbor homozygous nonsense mutations in both NtHMA α
53 and NtHMA β . The development of the other plants does not differ from the
54 development of wild-type plants. (B) Leaves of plants in which the two
55 NtHMA genes are silenced (top) are compared to leaves of plants
56 transformed with the empty vector (bottom). The main differences between
57 the two lines concern the secondary veins organization and the shape of the
58 leaves. (C) Representative mature capsules collected from different
59
60

1
2
3 combinations of reciprocal crosses between wild-type plants and
4 *NtHMA α /NtHMA β* -silenced plants.
5
6

7 **Supplementary figure S5: SXRF Fe and Zn quantitative analysis of**
8 **seeds.** (A) Iron and zinc distribution in the seed as visualized by SXRF.
9 Scales on the right side correspond to Fe and Zn signal intensities
10 (normalized fluorescence counts) analyzed using Microanalysis Toolkit⁴⁹. (B)
11 Quantification of the signal visualized in panel A using Microanalysis Toolkit.
12 (C) Fe and (D) Zn content quantified by ICP-MS analysis. Error bars
13 represent standard errors (n = 3). * : P < 0.05 ; ** : P < 0.005.
14
15
16

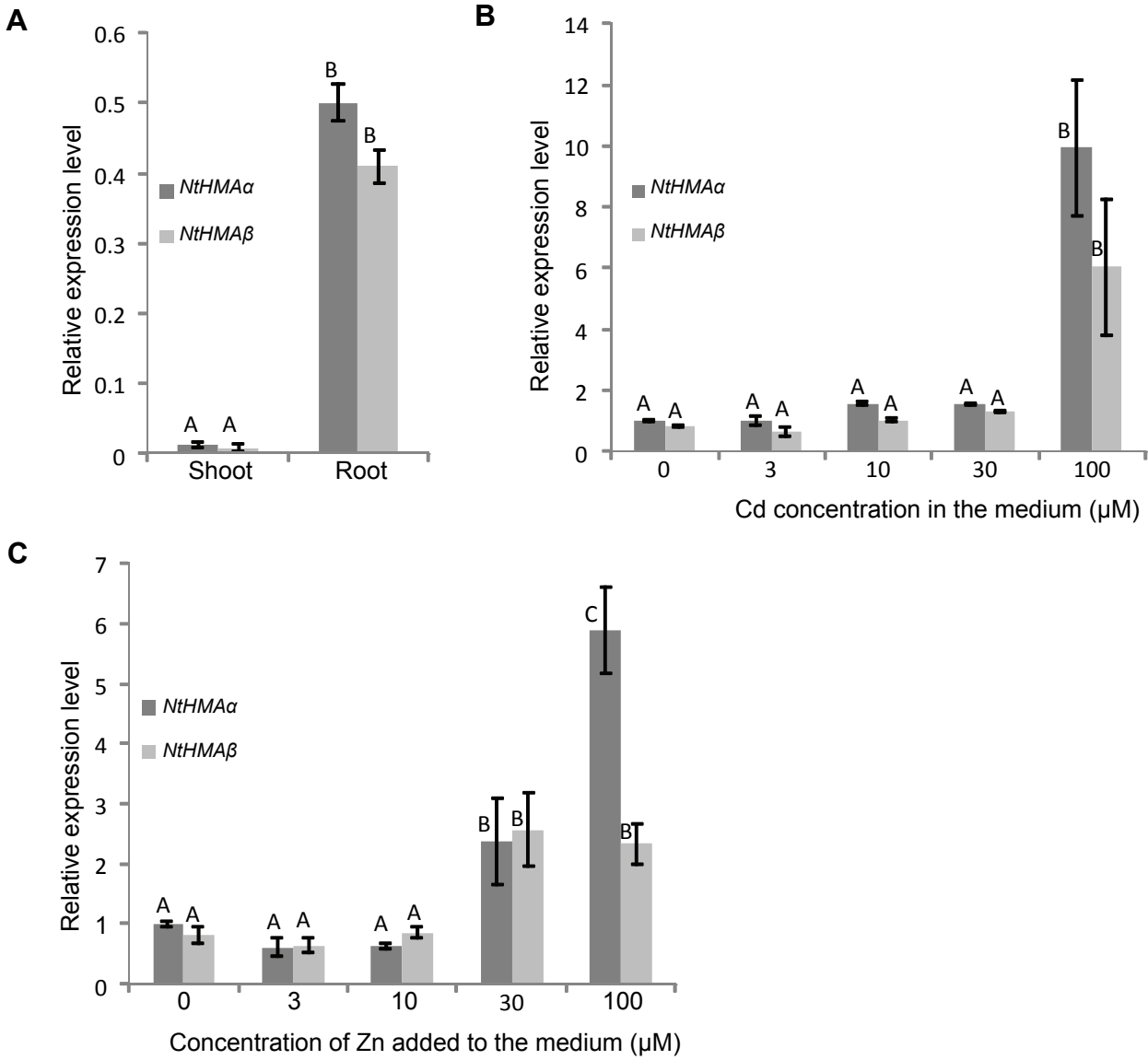
17 **Supplementary table S1: list of genetically modified lines.** The “**”
18 indicate the apparition of a new stop codon.
19
20

21 **Supplementary table S2: list of the primers used.**
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

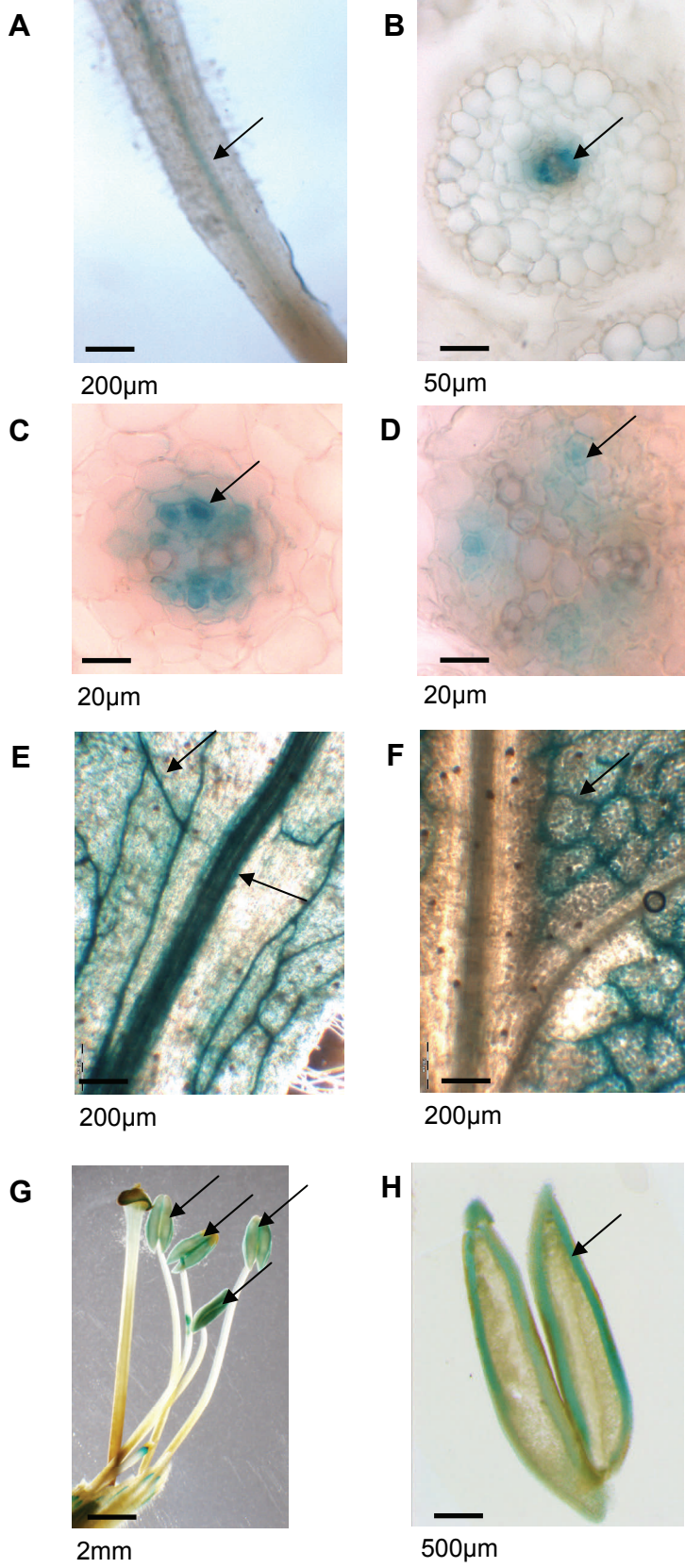


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 2

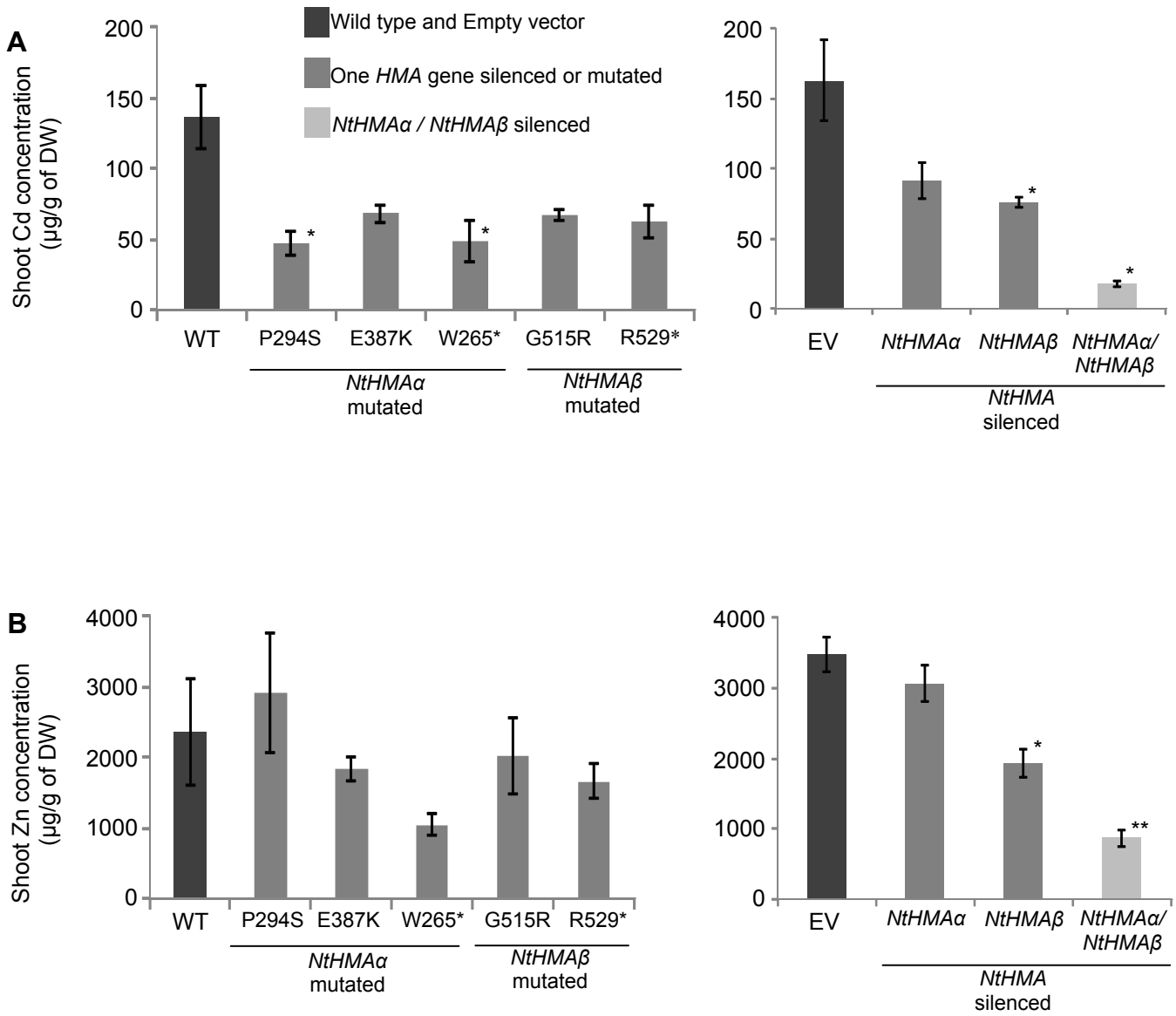


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 4



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 5

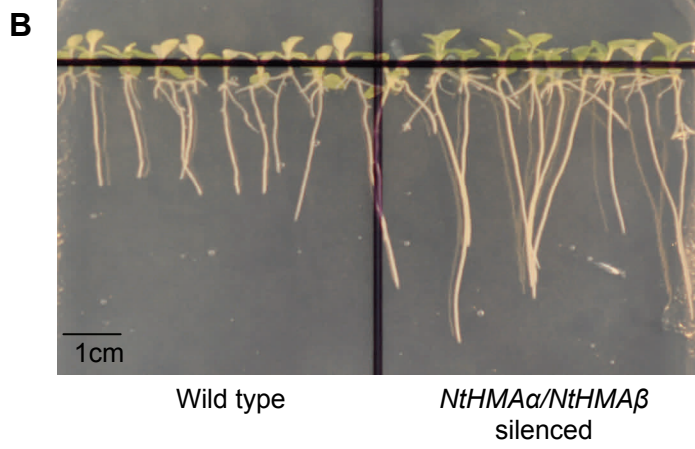
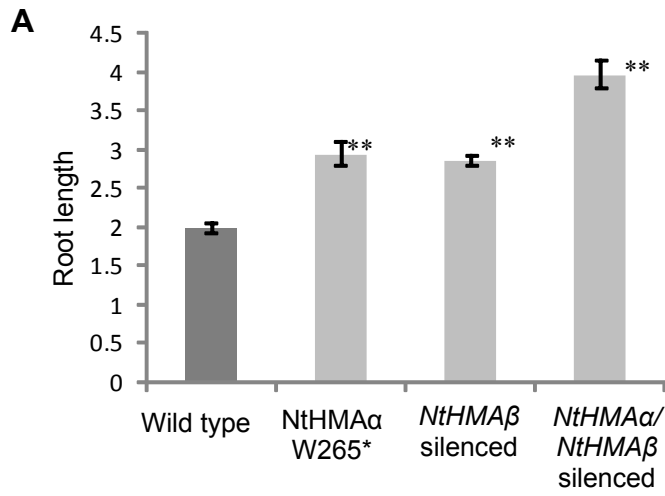
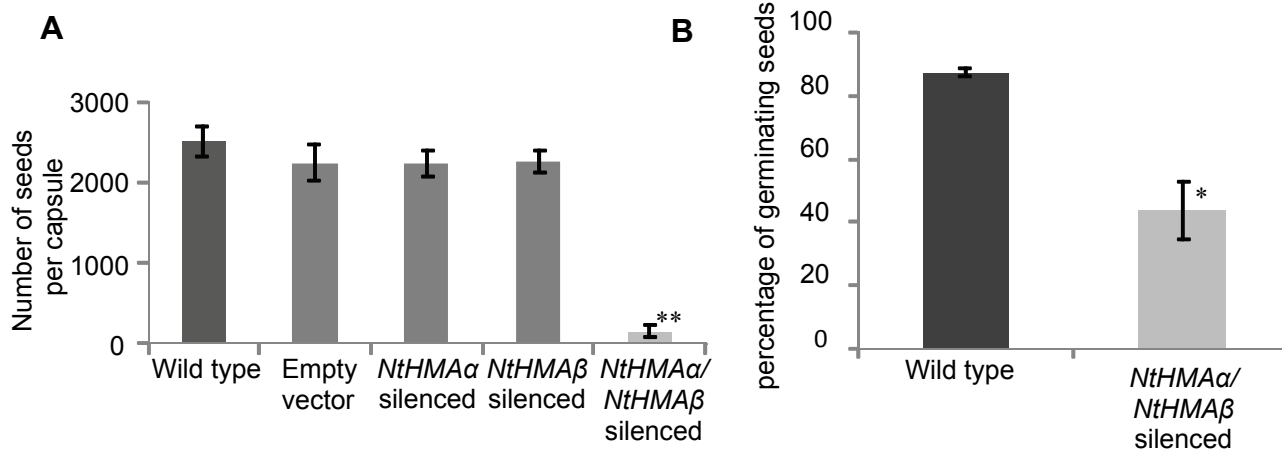


Figure 6



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 7

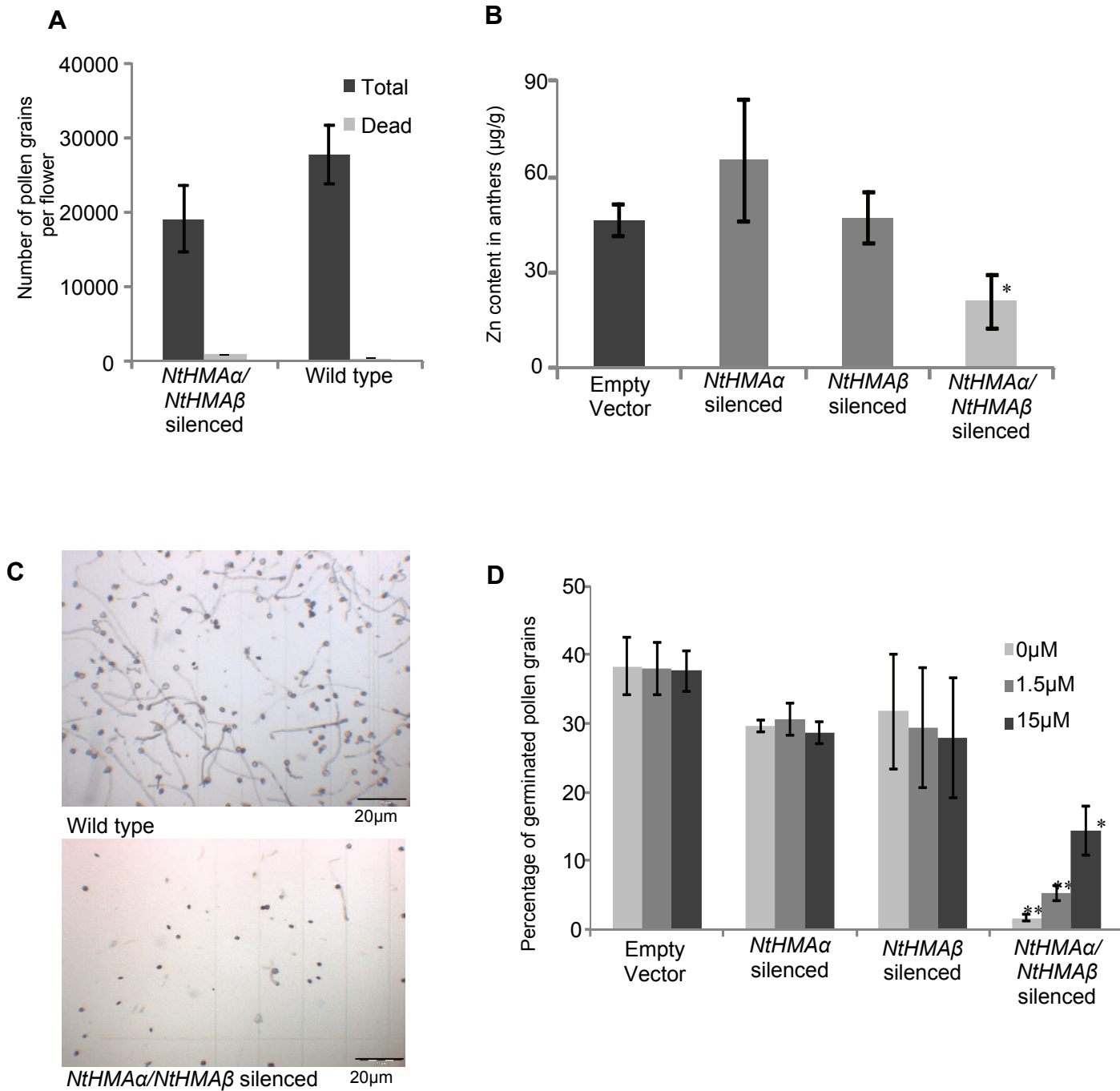
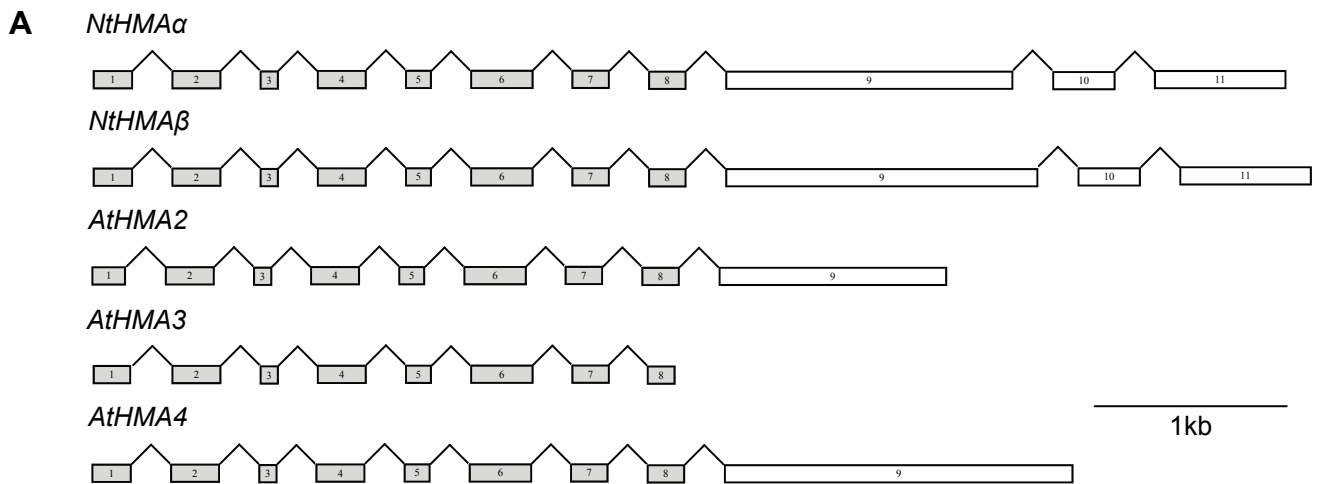


Fig sup 1



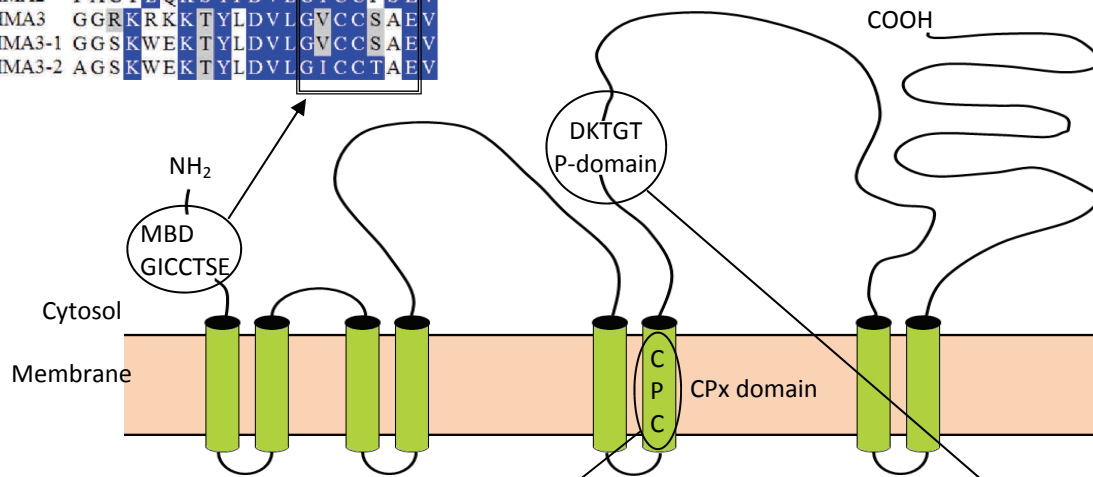
B

	<i>NtHMAα</i>	<i>NtHMAβ</i>	<i>AtHMA2</i>	<i>AtHMA3</i>	<i>AtHMA4</i>
<i>NtHMAα</i>	100				
<i>NtHMAβ</i>	98	100			
<i>AtHMA2</i>	67	67	100		
<i>AtHMA3</i>	48	47	55	100	
<i>AtHMA4</i>	66	65	81	55	100

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Fig sup 2

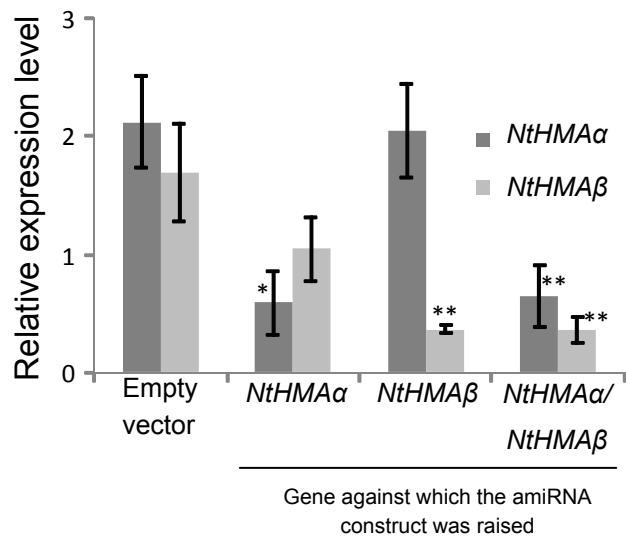
1
 2
 3
 4
 5
 6 → NiHMAa E T **K** **K** **L** **S** **K** **S** **Y** **F** **D** **V** **L** **G** **I** **C** **C** **T** **S** **E** **V**
 7 NSylHMA4 E T **K** **K** **L** **S** **K** **S** **Y** **F** **D** **V** **L** **G** **I** **C** **C** **T** **S** **E** **V**
 8 NbHMA4 E T **K** **K** **L** **S** **K** **S** **Y** **F** **D** **V** **L** **G** **I** **C** **C** **T** **S** **E** **V**
 9 → NiHMAb D T **K** **N** **L** **S** **K** **S** **Y** **F** **D** **V** **L** **G** **I** **C** **C** **T** **S** **E** **V**
 10 NtomHMA4 D T **K** **N** **L** **S** **K** **S** **Y** **F** **D** **V** **L** **G** **I** **C** **C** **T** **S** **E** **V**
 11 SIHMA4 D I **N** **K** **L** **S** **K** **S** **Y** **F** **D** **V** **L** **G** **I** **C** **C** **T** **S** **E** **V**
 12 MtHMA-A - - - I **K** **R** **S** **N** **F** **E** **V** **L** **G** **M** **C** **C** **A** **T** **E** **A**
 13 AhHMA3 **K** **K** **K** **N** **L** **Q** **T** **S** **Y** **F** **D** **V** **V** **G** **I** **C** **C** **T** **S** **E** **V**
 14 AIHMA3 **K** **K** **K** **N** **L** **Q** **T** **S** **Y** **F** **D** **V** **V** **G** **I** **C** **C** **S** **S** **E** **V**
 15 AtHMA3 **K** **K** **M** **N** **L** **Q** **T** **S** **Y** **F** **D** **V** **V** **G** **I** **C** **C** **S** **S** **E** **V**
 16 AhHMA4 **K** **V** **K** **K** **L** **Q** **K** **S** **Y** **F** **D** **V** **L** **G** **I** **C** **C** **T** **S** **E** **V**
 17 AtHMA4 **K** **V** **K** **K** **L** **Q** **K** **S** **Y** **F** **D** **V** **L** **G** **I** **C** **C** **T** **S** **E** **V**
 18 AIHMA2 - - **K** **K** **M** **T** **K** **S** **Y** **F** **D** **V** **L** **G** **I** **C** **C** **T** **S** **E** **V**
 19 AtHMA2 - - **K** **K** **M** **T** **K** **S** **Y** **F** **D** **V** **L** **G** **I** **C** **C** **T** **S** **E** **V**
 20 MtHMA-B **G** **G** **K** **K** **L** **Q** **K** **S** **Y** **F** **D** **V** **V** **G** **L** **C** **C** **S** **S** **E** **V**
 21 OsHMA2 **E** **G** **G** **R** **C** **Q** **K** **S** **Y** **F** **D** **V** **L** **G** **I** **C** **C** **P** **S** **E** **V**
 22 SbHMA2 **P** **A** **G** **T** **L** **Q** **K** **S** **Y** **F** **D** **V** **L** **G** **I** **C** **C** **P** **S** **E** **V**
 23 OsHMA3 **G** **G** **R** **K** **R** **K** **T** **Y** **L** **D** **V** **L** **G** **V** **C** **C** **S** **A** **E** **V**
 24 SbHMA3-1 **G** **G** **S** **K** **W** **E** **K** **T** **Y** **L** **D** **V** **L** **G** **V** **C** **C** **S** **A** **E** **V**
 25 SbHMA3-2 **A** **G** **S** **K** **W** **E** **K** **T** **Y** **L** **D** **V** **L** **G** **I** **C** **C** **T** **A** **E** **V**



26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37 → NiHMAa L A L V T L V S A C P C A L V L S T P V A M
 38 NSylHMA4 L A L V T L V S A C P C A L V L S T P V A M
 39 NbHMA4 L A L V T L V S A C P C A L V L S T P V A M
 40 → NiHMAb L A L V T L V S A C P C A L V L S T P V A M
 41 NtomHMA4 L A L V T L V S A C P C A L V L S T P V A M
 42 SIHMA4 L A L V A L V S A C P C A L V L S T P V A M
 43 MtHMA-A L A L V V L L S G C P C A L I L S T P V A I
 44 AhHMA3 L A L V V L V S G C P C G L I L S T P I A T
 45 AIHMA3 L A L V V L V S G C P C G L I L S T P V A T
 46 AtHMA3 L A L V V L V S G C P C G L I L S T P V A T
 47 AhHMA4 L A L V V L V S G C P C G L I L S T P V A T
 48 AtHMA4 L A L V V L V S G C P C G L I L S T P V A T
 49 AIHMA2 L A L V V L V S A C P C G L I L S T P V A T
 50 AtHMA2 L A L V V L V S A C P C G L I L S T P V A T
 51 MtHMA-B F A L V V L V S A C P C A L I L S T P V A T
 52 OsHMA2 L A L V L L V S A C P C A L V L S T P I A T
 53 SbHMA2 L A L V L L V S A C P C A L V L S T P V A T
 54 OsHMA3 L A L V M L V S A C P C A L V L S T P V A S
 55 SbHMA3-1 L A L V L L V S A C P C A L V L S T P V A T
 56 SbHMA3-2 L S L V L L V S A C P C A L V L S T P V A T

37 → NiHMAa Y L E T L A K I K I M A F D K T G T I T K G E F
 38 NSylHMA4 Y L E T L A K I K I M A F D K T G T I T K G E F
 39 NbHMA4 Y L E T L A K I K I M A F D K T G T I T K G E F
 40 → NiHMAb Y L E T L A K I K I M A F D K T G T I T R G E F
 41 NtomHMA4 Y L E T L A K I K I M A F D K T G T I T R G E F
 42 SIHMA4 Y L E T L A K I K I M A F D K T G T I T R G E F
 43 MtHMA-A Y L E T L S R I K I T V A F D K T G T I T R G E F
 44 AhHMA3 C L E T L A K I K I V A F D K T G T I T K A E F
 45 AIHMA3 C L E T L A K I K I V A F D K T G T I T K A E F
 46 AtHMA3 C L E T L A K I K I V A F D K T G T I T K A E F
 47 AhHMA4 Y L D T L S K I K I A A F D K T G T I T R G E F
 48 AtHMA4 Y L D T L S K I K I V A F D K T G T I T R G E F
 49 AIHMA2 Y L E T L A K I K I V A F D K T G T I T R G E F
 50 AtHMA2 Y L E T L A K I K I V A F D K T G T I T R G E F
 51 MtHMA-B A L E T L A K I K V M A F D K T G T I T K G E F
 52 OsHMA2 V L E S L A S I K V A A F D K T G T I T R G E F
 53 SbHMA2 V L E T L A R I K I A A F D K T G T I T R G E F
 54 OsHMA3 V L E S L G E I R A V A F D K T G T I T R G E F
 55 SbHMA3-1 V L E S L G E I R V A A F D K T G T I T K G Q F
 56 SbHMA3-2 V L E S L G E V S V A A F D K T G T I T R G E F

Fig sup 3




1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

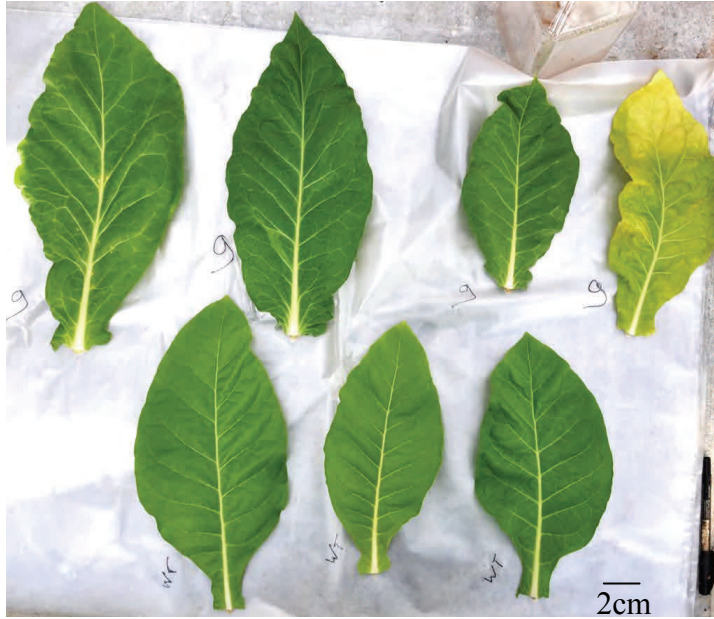
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

A



 *Nthmaα*⁻/*Nthmaα*⁻
Nthmaβ⁻/*Nthmaβ*⁻
Homozygous mutant
plants

B



NtHMAα/*NtHMAβ*
silenced

Empty vector

C

Wild Type ♂

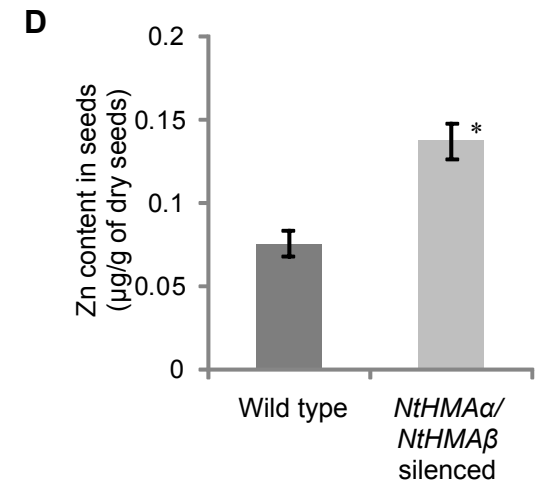
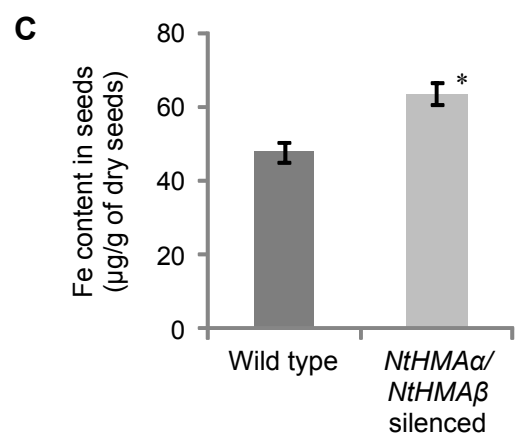
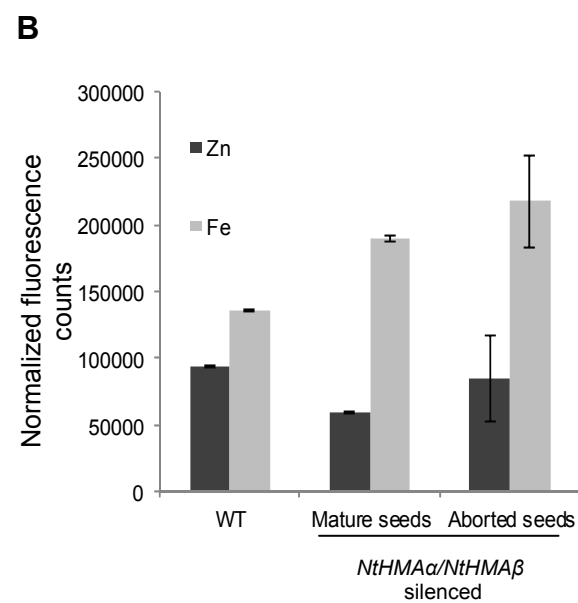
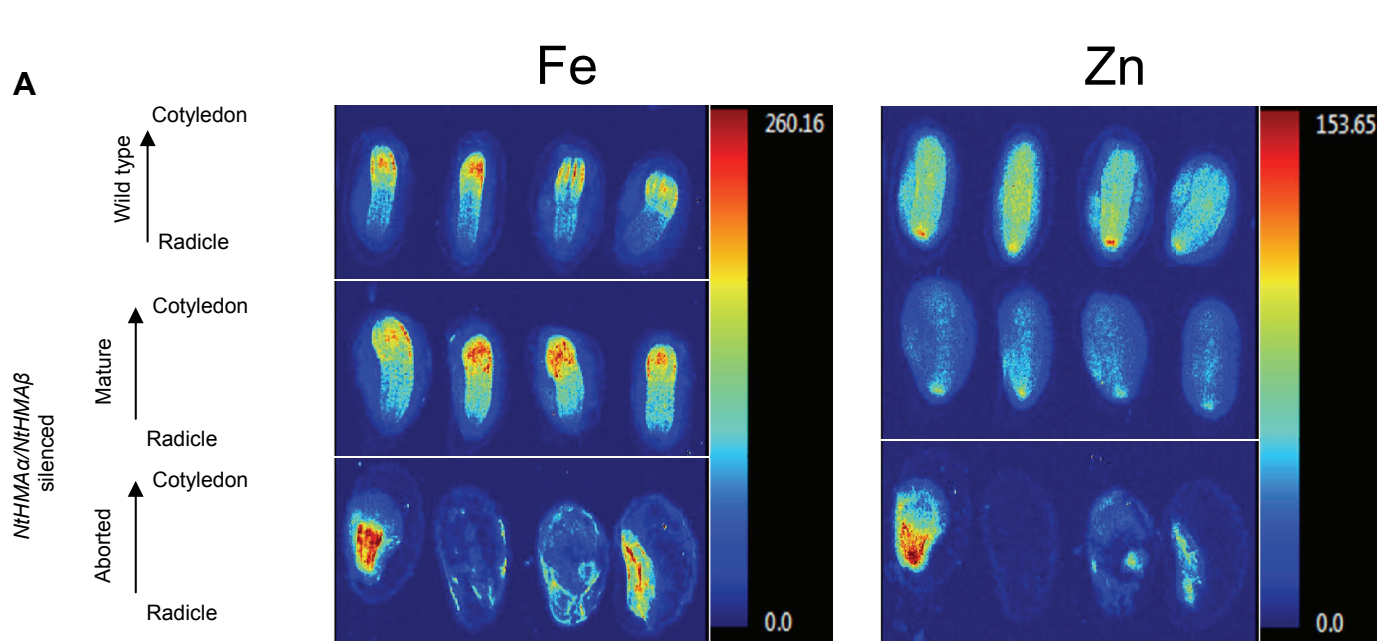
NtHMAα/*NtHMAβ*
Silenced ♂



Wild
Type ♀



NtHMAα/
NtHMAβ
Silenced ♀



Metallomics Accepted Manuscript

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table sup 1

Supplementary table 1: list of genetically modified lines.

Type of line	Gene mutated	Type of mutation	Mutation
	<i>NtHMAα</i>	Missense	P294S
	<i>NtHMAα</i>	Missense	E387K
Mutant	<i>NtHMAα</i>	Nonsense	W265*
	<i>NtHMAβ</i>	Missense	G515R
	<i>NtHMAβ</i>	Nonsense	R529*
	Gene silenced	Exon targeted	
	<i>NtHMAα</i>	11	
Silenced	<i>NtHMAβ</i>	11	
	<i>NtHMAα/NtHMAβ</i>	9	

The "***" indicate the apparition of a new stop codon.

Designed for	Name	Sequence	Comment
	I miR-s	GATGACTCTTATTATTTCCGCATTCTCTCTTTTGTATTCC	
	II miR-a	GAATGCGGAAATAATAAGAGTCATCAAAGAGAATCAATGA	Targeting <i>NtHMAβ</i>
	III miR*s	GAATACGGAAATAATTAGAGTCTTCACAGGTCGTGATATG	
	IV miR*a	GAAGACTCTAATTATTTCCGTATTCTACATATATATTCCT	
	I miR-s	GATGGATCTTATCATTTCGCATTCTCTCTTTTGTATTCC	
	II miR-a	GAATGCGGAAATGATAAGATCCATCAAAGAGAATCAATGA	Targeting <i>NtHMAα</i>
	III miR*s	GAATACGGAAATGATTAGATCCTTCACAGGTCGTGATATG	
	IV miR*a	GAAGGATCTAATCATTTCGTATTCTACATATATATTCCT	
amiRNA construct	I miR-s	GATTCATTAGTATAACATGGCCTTCTCTCTTTTGTATTCC	
	II miR-a	GAAGGCCATGTTATACTAATGAATCAAAGAGAATCAATGA	Targeting <i>NtHMAα/NtHMAβ</i>
	III miR*s	GAAGACCATGTTATAGTAATGATTACACAGGTCGTGATATG	
	IV miR*a	GAATCATTACTATAACATGGTCTTCTACATATATATTCCT	
	I miR-s	GATGATTTTAGCTAGAGTCTCAATCTCTCTTTTGTATTCC	
	II miR-a	GATTGAGACTCTAGCTAAAATCATCAAAGAGAATCAATGA	Targeting <i>NtHMAα</i>
	III miR*s	GATTAAGACTCTAGCAAAAATCTTCACAGGTCGTGATATG	
	IV miR*a	GAAGATTTTGTAGCTAGAGTCTTAATCTACATATATATTCCT	
	I miR-s	GATGATTTTAGCAAGAGTCTCAATCTCTCTTTTGTATTCC	
	II miR-a	GATTGAGACTCTTGTCTAAAATCATCAAAGAGAATCAATGA	Targeting <i>NtHMAβ</i>
	III miR*s	GATTAAGACTCTTGCAAAAATCTTCACAGGTCGTGATATG	
	IV miR*a	GAAGATTTTGTCAAGAGTCTTAATCTACATATATATTCCT	
Sequence of	BamH1-ATGNtHMAa	CGCGGATCCGAGAAGGAGAAAAATGGTGGAAAGT	5' segment
	RevHMAa5'	CGCGAATCCCGACTTGCAACATGAAGCTTTGTCTTTGTGGT	
	FwHindEco	TTCGCATGCTCCACCACAAAGACA	Middle segment
	RevEcorR1	ATGGACCAGACATCTTTGAAGAGC	
	FwEcorR1	GGAAATACAAAGTCTCGGACACC	3' segment
	Stop-XhoNtHMAa	CCGCTCGAGCTTCGGATTGCATCTACTCTATGACA	
Sequence of	BamH1-ATGNtHMAb	CGCGGATCCGAAAGAAGAGAATGGTGGAAAGT	5' segment
	RevHMAb5'	TGCGGAGCATTTCGACTTGCAACAA	
	Béta Fw Hind	CACACAGACATGGGAAAAATGTTGTAGATCTG	Middle segment
	Béta Rev Eco	CATTGATCCTGAAATTGCAGTTTGGCATGTTTGATTGA	
	Béta Fw Eco	AACTCAAGCTCCAGGACATGTGGAAATATG	3' segment
	Stop-XhoNtHMAb	CCGCTCGAGTCAGATTGCATCTACTCTATGACA	
quantitative PCR	Fw134.5	ACAAAGTGCTCGACACCAA	
	Rev134.5m.alpha	CTTCTCGTTGCAGAGTCCT	<i>NtHMAα</i> and <i>NtHMAβ</i>
	Rev134.5m.béta	CTTCTCGTTGCAGAGTCTA	
	L2-jc-f	GGCGAAATGGGTGTTTGTATC	Housekeeping Ribosomal protein L2
	L2-jc-r	CGTTCCGTTCCCGAAGTCG	
	cyclophilin-jc-f	CTCTATGCCGACACCGTTCC	Housekeeping Cyclophilin
	cyclophilin-jc-r	TCACACGGTGGAGGTTGAG	
Promoter cloning	FwPROMHMAa	TTCCACCATGGTTCTCTCTCTAAGGA	<i>NtHMAα</i>
	RevPROMHMAa	TCT ACC TAA ATC CCT AAT TTC TAC AGT AA	
SSCP	HMA-A1-seq3	GGTGAATAGCATTCTTGCTGT GCACAACATAAGATTCACTAAC	Detection of P294S and W265*
	HMA-B-seq3	GTCTGATTTGACTGGTGATG AAGAATATGTATGAGTGGAACC	Detection of E387K
	HMA-DT-seq47	GACCTGTTTAGCACTAATGCC TTATAATCATTTCAGCGTAATGCAG	Detection of G515R and R529*

SSCP stand for single strand conformation polymorphism.