An evolutionary perspective on differential regulation of zinc and cadmium homeostasis genes in *Arabidopsis thaliana* and *Noccaea caerulescens*

Ya-Fen Lin
An evolutionary perspective on differential regulation of zinc and cadmium homeostasis genes in *Arabidopsis thaliana* and *Noccaea caerulescens*

Ya-Fen Lin
Thesis committee

Promotor
Prof. Dr M. Koornneef
Personal chair at the Laboratory of Genetics
Wageningen University

Co-promotor
Dr M.G.M. Aarts
Associate Professor, Laboratory of Genetics
Wageningen University

Other members
Prof. Dr H.J. Bouwmeester, Wageningen University
Prof. Dr M.E. Schranz, Wageningen University
Prof. Dr S. Kärenlampi, University of Eastern Finland, Kuopio
Prof. Dr N. Verbruggen, Université Libre de Bruxelles, Belgium

This research was conducted under the auspices of the Graduate School of Experimental Plant Sciences.
An evolutionary perspective on differential regulation of zinc and cadmium homeostasis genes in *Arabidopsis thaliana* and *Noccaea caerulescens*

Ya-Fen Lin

Thesis

submitted in fulfilment of the requirements for the degree of doctor at Wageningen University
by the authority of the Rector Magnificus Prof. Dr M.J. Kropff, in the presence of the Thesis Committee appointed by the Academic Board to be defended in public on Monday 24 March 2014 at 4 p.m. in the Aula.
Ya-Fen Lin
An evolutionary perspective on differential regulation of zinc and cadmium homeostasis genes in Arabidopsis thaliana and Noccaea caerulescens
220 pages.

With references, with summaries in English and Dutch

Dedicated to my beloved parents and family

Jimmy and lovely Teck-yew
Contents

**Chapter 1** General introduction 9

**Chapter 2** The molecular mechanism of zinc and cadmium stress response in plants 19

**Chapter 3** The expression of the ZNT1 zinc transporter from the metal hyperaccumulator *Noccaea caerulescens* confers enhanced zinc and cadmium tolerance and accumulation to *Arabidopsis thaliana* 49

**Chapter 4** The selective advantage of a mutant *nicotianamine synthase* allele in a natural *Noccaea caerulescens* population 95

**Chapter 5** A comprehensive set of transcript sequences of the heavy metal hyperaccumulator *Noccaea caerulescens* 133

**Chapter 6** General discussion 167

Reference 180
Summary 207
Samenvatting 209
Acknowledgement 212
Curriculum Vitae 215
Publication 216
Education Statement 218
Chapter 1

General Introduction
Certain heavy metals, such as Zn, Fe, Cu and Mn, are micronutrients and thus are essential for plant growth and development (Appenroth, 2010). However, when the concentrations of these micronutrients exceed a certain level, the excess metals become toxic and this has a negative impact on plant health. Other metals, such as Cd, Hg and Pb, are non-essential and toxic to plants even at low concentrations. For example, exposure to Cd damages the root system, causes leaf chlorosis, disrupts water balance, inhibits stomatal opening, and also affects the photosynthetic apparatus (Claire-Lise & Nathalie, 2012, Clemens, 2006).

Human activities such as metal mining and smelting, industrial processes, and the use of fertilizers, have increased the concentrations of metals in the soil (McLaughlin et al., 2000, Singh et al., 2011, Zhang & Wong, 2007). When such pollutants exceed the levels that plants can tolerate, there is a negative impact on crop yield, the balance of the local ecosystems, and eventually human health (Lürling & Scheffer, 2007, Singh & Aggarwal, 2011). Although all plants are sensitive to heavy metals, the tolerance level differs between species. Dose–response curves can be used to explain how the toxicity of metals depends on the external bioavailable metal concentration, the exposure time, the genotype of the plant, and its general health status (Lin & Aarts, 2012) (Chapter 2). When plants are exposed to high concentrations of metal, various species-dependent strategies can be used to cope with the resulting metal stress. Based on these strategies, plants can be divided into four major groups: metal-sensitive species, metal-resistant excluder species, metal-tolerant non-hyperaccumulator species, and metal-hypertolerant hyperaccumulator species (Lin & Aarts, 2012) (Chapter 2). It is therefore important to determine how particular plant species deal with high metal concentrations by adapting their physiology and metabolism to evolve into heavy metal hypertolerant and hyperaccumulator species. The resulting knowledge can be used to develop plants suitable for environmental clean-up purposes, a phenomenon known as phytoremediation.

**Metal hyperaccumulator plant species**

Among the various strategies plants have evolved to deal with metal stress (Lin & Aarts, 2012) (Chapter 2, Fig. 2), metal-hypertolerant hyperaccumulator plant species have drawn the most attention because they can potentially be used for phyto-extraction, i.e. the use of plants to accumulate metals that are subsequently isolated from the biomass and even re-used (Pilon-Smits, 2005). Quantitative criteria to define hyperaccumulator plant species were first proposed by (Brooks et
al., 1977) for a Ni hyperaccumulator, in which the Ni concentration was more than 1000 µg/g leaf dry weight, i.e. 100–1000 times higher than that normal. Recently, van der Ent et al. (2013) proposed that plants should be defined as hyperaccumulators if the metal concentration in the leaves was 2–3 orders of magnitude higher than normal when growing in standard soils and one order of magnitude higher than normal when growing in metalliferous soils. Based on these criteria, more than 500 plant taxa have been defined as hyperaccumulators of one or more metal elements, including 450 Ni hyperaccumulators (the largest group), 12 Zn hyperaccumulators, and 2 Cd hyperaccumulators (van der Ent et al., 2013).

The best model Zn/Cd hyperaccumulator species are Arabidopsis halleri and Noccaea caerulescens (Krämer, 2010, Milner&Kochian, 2008). N. caerulescens (2n = 14) is a self-compatible plant that can produce seeds either by selfing or inter-crossing. It is more amenable than A. halleri, which is self-incompatible and therefore more difficult to study at the genetic level. N. caerulescens has evolved to hyperaccumulate several metals (Zn, Ni, Cd, Pb) and although less closely related to A. thaliana than A. halleri, it still shares 88.5% sequence identity at the transcript level (Rigola et al., 2006). This makes N. caerulescens the most suitable plant species in which to study the mechanisms of heavy metal hypertolerance and hyperaccumulation (Assunção et al., 2003a, Milner&Kochian, 2008). Furthermore, many natural N. caerulescens ecotypes have been identified in serpentine, metalliferous and non-metalliferous soils (Fig. 1). Although they are grown in the same condition for the same duration, their sizes and appearances differ, e.g. the accessions from Plombières and the Moravsko slezské Beskydy mountains are larger than the accessions from Ganges, Mezica, and Col du Mas de l’Ayre and they differ in the shape of the leaves. Also, the metal tolerance and accumulation abilities are different, e.g. La Calamine and Ganges display higher Zn/ Cd tolerance than Lellingen and Monte Prinzera; Ganges accumulates Zn and Cd to higher concentrations than La Calamine (Assunção et al., 2003b). The natural genetic differences between different N. caerulescens accessions are ideal to study the genetics and genomics of heavy metal homeostasis, and the resulting knowledge can be applied in the field of phytoremediation.
Fig 1. Appearance of different *Noccaea* accessions

*Noccaea caerulescens* (J. & C. Presl) F. K. Mey accessions were collected from the (1) Krušné hory Mountains, Bohemia, Czech Republic (KH); (2) Durfort, France (DF); (3) Plombières, Belgium (PB); (4) Lellingen/Wilwerwilz, Luxembourg (LE); (5) La Calamine, Belgium (LC) (first row, left to right); (6) Moravsko slezské Beskydy Mountains, Moravia, Czech Republic (MB); (7) St. Laurent Le Minier/Ganges, southern France (GA); (8) La Calamine, Belgium (LC42); (9) Monte Prinzena, Italy (MP), (10) St Félix de Paillières, France (SF) (second row, left to right); and (15) Col du Mas de l’Ayre, southern France (CMA) (fourth row). Four *Noccaea praecox* (Wuljen) F. K. Mey accessions from (11) Zaplana, (12) Mezica, (13) Zerjav, and (14) unpolluted soil in Slovenia (third row, left to right) are also shown. The image also shows two accessions of the close relative *Microthlaspi perfoliatum* (L.) F. K. Mey, from Bologna, Italy (16), and from the pustas of Hungary (17) (fourth row).
Evolution of metal hyperaccumulators

The discovery of heavy metal hyperaccumulators in natural habitats with high heavy metal loads has drawn the interest of evolutionary biologists seeking to learn how such adaptations arose. Several hypotheses have been proposed to explain the selective benefits of metal hyperaccumulation in plants (Boyd, 2004). (1) The tolerance/disposal hypothesis suggests that metal hyperaccumulation allows the metal concentration in plant tissues to be reduced by accumulating excess metal in the leaves (tolerance) which are eventually shed to the soil (disposal). (2) The interference hypothesis, also known as the elemental allelopathy hypothesis (Boyd&Martens, 1998), proposes that perennial hyperaccumulator plants use their canopies to enrich soil metal concentration and therefore interfere with the establishment of less tolerant plant species. (3) The drought resistance hypothesis proposes that metal hyperaccumulation in plants increases drought tolerance. However, these three hypotheses are supported by limited experimental evidence and their veracity remains to be established. In contrast, (4) the elemental defense hypothesis, also known as the inorganic defense hypothesis (Boyd, 2010), suggests that high metal concentrations can protect plants from certain pathogens and herbivores, and this hypothesis is well supported by scientific data (Boyd, 2007).

The inorganic defense hypothesis actually incorporates two variations that may explain how metal hyperaccumulation evolved to fulfil a role in plant defense (Boyd, 2012). The defensive enhancement hypothesis suggests that heavy metal hyperaccumulation evolved in some plant species because the enhanced metal concentration can protect plants from pathogens and herbivores (protective benefits). Recently, it has been shown that the threshold for the defensive effect can even fall below the defined hyperaccumulation threshold (Cheruiyot et al., 2013). The joint effects hypothesis, an extension of the defensive enhancement concept, suggests that metal hyperaccumulation can be achieved at a lower metal concentration if another inorganic compound or other defenses are also present. This takes the form of a co-evolutionary arms race, in which the herbivores/pathogens evolve metal tolerance such that the plant must accumulate higher and higher metal concentrations to maintain its advantage (Hörger et al., 2013).

Based on current knowledge, a potential elemental defense scenario to explain the evolution of hyperaccumulation *N. caerulescens* was proposed by (Hörger et al., 2013). This combines selective pressure caused by pathogens and
herbivores together with the gene flow from non-metallicolous populations to increase the potential to establish hyperaccumulation traits in metallicolous populations. Therefore, the evolution of hyperaccumulation can be investigated by comparing the genetic differences among metallicolous/non-metallicolous populations of *N. caerulescens* growing in metalliferous/non-metalliferous habitats, and examining the different forms of defense against herbivores and pathogens.

**Molecular mechanisms of metal hyperaccumulation**

When heavy metals enter plants, whether they are essential micronutrients or non-essential elements, the concentration and distribution of metals must be tightly regulated by controlling (1) mobilization at the root site, (2) uptake by and sequestration in the roots, (3) transport through the xylem, (4) unloading the xylem and distribution to tissues such as shoots and leaves, and (5) final intracellular distribution and sequestration (Clemens *et al.*, 2002). The molecular basis of the Zn and Cd stress responses in metal-sensitive and metal-adapted species is compared and reviewed in Chapter 2 (Lin&Aarts, 2012). The regulatory mechanisms listed above must be tightly regulated to ensure survival when plants are exposed to heavy metals, and two of the most important regulatory mechanisms involve metal transporters and chelators. Metal transporters such as AtZIP4/NcZNT1 (discussed in Chapter 3) can control the uptake of metals into the cell and therefore play a decisive role in metal tolerance/accumulation. Following uptake, metal-binding chelators such as nicotianamine (discussed in Chapter 4) can facilitate the transport, loading and unloading of metals in plant tissues, thus facilitating metal detoxification and hyperaccumulation and the ability of plants to adapt to metalliferous environments.

Metal tolerance strategies and regulatory mechanisms in plants have been discussed in several reviews (Hassan&Aarts, 2011, Krämer *et al.*, 2007, Palmgren *et al.*, 2008, Sinclair&Krämer, 2012, Tamas *et al.*, 2006). When exposed to metals, non-accumulator species attempt to prevent the uptake of excess or toxic metals, whereas metal hyperaccumulators enhance the accumulation of essential micronutrients such as Zn as well as toxic metals such as Cd by (1) boosting the uptake of metals by root cells; (2) reducing the metal sequestration capacity of root cells adjacent to the vasculature; (3) enhancing metal loading into the xylem; (4) facilitating the unloading of metals from the xylem into the shoots; (5) enhancing the transport of metals through neighboring cells to metal storage tissues (e.g. the
epidermis and mesophyll); and (6) by increasing the metal-sequestration capacity of these sink cells (Hassan & Aarts, 2011).

The study of metal homeostasis in hyperaccumulator species such as *N. caerulescens* and *A. halleri* relies on the plentiful resources available for the non-accumulator *A. thaliana*, particularly the genome sequence, microarrays, and knockout mutant libraries representing all genes. For example, microarray-based transcript-profiling studies have been used to investigate differential gene expression between *N. caerulescens* and *A. thaliana* (van de Mortel et al., 2006, van de Mortel et al., 2008). Direct genetic studies in *N. caerulescens* are possible (e.g. to determine differences between accessions) reflecting the availability of efficient growth systems (Assunção et al., 2003b) and genetic maps (Assunção et al., 2006), but there is no complete physical map and no genome sequence. The available cDNA sequences (Rigola et al., 2006) and BAC genomic libraries (Iqbal et al., 2013) have already contributed to the understanding of hyperaccumulation in *N. caerulescens*, but the available sequence information is still limited, which makes it difficult to isolate the coding and regulatory sequences of candidate genes.

The development of the second (next) and the third (next-next) generation sequencing techniques has made it possible to achieve efficient whole-transcriptome and whole-genome sequencing in metal hyperaccumulators such as *N. caerulescens*. The resulting well-annotated genome and transcriptome sequences will not only accelerate the functional analysis of metal homeostasis genes, but will also provide insight into the adaptation of *N. caerulescens* by identifying genetic factors that are unique to hyperaccumulators. This will be facilitated by comparing *N. caerulescens* sequences to those of non-accumulator Brassicaceae species such as *A. thaliana* and *A. lyrata* (Arabidopsis, 2000, Hu et al., 2011, Krämer, 2010, Slotte et al., 2013, Wang et al., 2011, Wu et al., 2012). Second-generation sequencing techniques will also improve our understanding on metal hyperaccumulation in *N. caerulescens* by using RNA sequencing to generate high-density markers for genetic studies including quantitative trait locus (QTL) analysis or association mapping of metal tolerance and accumulation traits (Verbruggen et al., 2013). The re-sequencing of accessions differing in metal accumulation and/or tolerance will help to identify genes that can be selected to improve metal tolerance, as described for *A. lyrata* (Turner et al., 2010).
The main questions

The aim of this thesis is to gain insight into the regulatory basis of Zn and Cd hypertolerance/hyperaccumulation in *N. caerulescens* by comparative analysis with its non-accumulator relative *A. thaliana*. The ecological impact of metal hyperaccumulation and the evolution of hyperaccumulation as a defensive trait will also be discussed. The main questions addressed in this thesis are listed below.

1. Transporters such as AtZIP4 are known to play an important role in the regulation of Zn transport in *A. thaliana* (Assunção et al., 2010). Therefore, we will investigate whether this transporter functions differently in *N. caerulescens*, and if it is subject to differential regulation in *A. thaliana* and *N. caerulescens* through the analysis of cis-regulatory elements (Chapter 3).

2. Enzymes such as nicotianamine synthase (NAS), which synthesize metal chelators, have been shown to play an important role in Zn translocation and hyperaccumulation in *A. halleri*. We will investigate how this enzyme confers a potential selective advantage in hyperaccumulator species.

3. Previously an *N. caerulescens* expressed sequence tag (EST) database was created to identify metal-related genes, although the number remains limited (Rigola et al., 2006). We will sequence *N. caerulescens* RNA using next-generation sequencing techniques to generate sequences representing further metal-related genes with a role in metal homeostasis.

Outline of the thesis

This thesis describes the basis of metal homeostasis in the metal hyperaccumulator species *N. caerulescens*, focusing on the essential nutrient Zn and the non-essential element Cd. Our current understanding of the molecular basis of heavy metal homeostasis in plant cells is reviewed in Chapter 2. The ability of the *N. caerulescens* zinc transporter NcZNT1 to confer enhanced Zn and Cd tolerance and accumulation in the non-accumulator species *A. thaliana* is described in Chapter 3. Following uptake into plants, Zn and Cd form a complex with chelators such as nicotianamine, which is synthesized by nicotianamine synthase (NAS). A transposon was shown to be inserted within the *NAS1* gene of a natural *N. caerulescens* population, and the resulting *nas1* mutant is compared to wild-type plants in Chapter 4. The potential benefit of this mutant *nas1* allele in nature is also explored by analyzing growth and pest tolerance. In Chapter 5, we describe a comprehensive set of *N. caerulescens* transcript sequences that will support future studies investigating the molecular mechanisms of metal adaptation, thus
contributing to advanced sequencing studies such as whole-genome sequencing and the re-sequencing of different *N. caerulescens* accessions. The main findings of each chapter and future perspectives are summarized in Chapter 6.
The molecular mechanism of zinc and cadmium stress response in plants

Ya-Fen Lin, Mark G. M. Aarts

Laboratory of Genetics, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

This chapter has been published on Cellular and Molecular Life Sciences (2012) 69:3187–3206

Fig 2a of this chapter has been used as the cover of Cellular and Molecular Life Sciences, Volume 69, Issue 19, October 2012
ABSTRACT

When plants are subject to high metal exposure, different plant species take different strategies in response to metal-induced stress. Largely, plants can be distinguished in four groups: metal-sensitive species, metal-resistant excluder species, metal-tolerant non-hyperaccumulator species, and metal-hypertolerant hyperaccumulator species, each having different molecular mechanisms to accomplish their resistance/tolerance to metal or reduce the negative consequences of metal toxicity. Plant responses to heavy metals are molecularly regulated in a process called metal homeostasis, which also includes regulation of the metal-induced Reactive Oxygen Species (ROS) signaling pathway. ROS generation and signaling plays an important dual role in heavy metal detoxification and tolerance. In this review, we will compare the different molecular mechanisms of nutritional (Zn) and non-nutritional (Cd) metal homeostasis between metal-sensitive and metal-adapted species. We will also include the role of metal-induced ROS signal transduction in this comparison, with the aim to provide a comprehensive overview of how plants cope with Zn/Cd stress at the molecular level.

KEYWORDS

Plant stress adaptation, molecular regulation, Zn, Cd, ROS
INTRODUCTION

Heavy metal contamination, in soil caused by human activities, such as mining and industrial activities, is a serious problem all over the world (Ikenaka et al., 2010, McLaughlin et al., 2000, Zhang&Wong, 2007). Heavy metals are a poorly defined set of chemical elements, mostly belonging to the so-called transition metals, but often also including elements like Pb and Tl. Quite a few of the lower molecular weight heavy metals are essential minerals, like Zn, Co, Cu, Ni, Mn, Mo and Fe, but most are non-essential for biology. All of them are potentially toxic, depending on their bioavailable concentrations and sensitivity of the organism that is exposed (Jarup, 2003, Kien et al., 2010, Rascio&Navari-Izzo, 2011). Most toxic to humans are the elements that resemble essential minerals, such as Hg, Pb, and Cd, which are more likely to enter the cell through the existing mineral uptake machinery. Cd exposure, for example, can cause emphysema and osteoporosis, leading to irreversible damages to lungs, kidneys, and bones in humans (Straif et al., 2009). Toxicity of these metals is not limited to humans or animals, but affects many organisms, including plants. Under excess heavy metal exposure, plants will display reduced biomass, leaf chlorosis, inhibited root growth, and morphological alterations, often leading to plant death at excessive exposures (S.K, 2010). While humans and animals can move and thus avoid heavy metal-contaminated areas, plants can not and need to evolve ways to deal with the heavy metals they encounter in their direct environment.

Among plant species, there is a wide variation in sensitivity to heavy metal exposure. In general, plants have developed two major strategies to resist high heavy metal exposure (Marschner, 1991, Verbruggen et al., 2009). The first one is the **excluder strategy**, in which plants try to avoid heavy metals entering the roots, for instance by restricting soil metal bioavailability or by reducing expression of metal uptake transport genes. The solubility of metals in the rhizosphere is affected by pH, cation exchange capacity, concentrations of organic compounds or metal chelating compounds, properties of the minerals, and the activity of microorganisms (Ghosh&Singh, 2005). Secretion of protons and exudation of carboxylates from roots can acidify the rhizosphere and increase metal solubility, bioavailability, and toxicity (Martinez&Motto, 2000). Organic acids that are excreted by roots can form a complex with heavy metals in the rhizosphere and thus inhibit the uptake of metals (Murphy&Taiz, 1995). Microorganisms can decrease metal solubility in soil by ways of biosorption, extracellular sequestration, transportation, bioprecipitation, and chelation of metals by siderophores
(Haferburg&Kothe, 2007), but bacterial siderophore excretion can also enhance bioavailability. With the excluder strategy, plants are trying to keep the metal concentrations in roots low, despite the elevated metal concentration in the soil. However, in areas highly contaminated with heavy metal, such as close to a metal smelter, it is often too demanding to exclude toxic metals from plant roots due to the high metal concentrations in the soil. To cope with such situation, plants have developed a different mode of action, which involves taking up metals and reducing the damage.

This second strategy is the tolerance strategy, which relies on confinement and detoxification of metals in a controlled way. This not only permits plants to withstand high metal exposure but also to accumulate metals to sometimes extremely high concentrations. Again a division can be made, this time between plants that tolerate high uptake of metals, but restrict their accumulation to roots, and plants that accumulate metals and preferentially transport metals to the above-ground parts (Pollard et al., 2002). The latter type of plants is often referred to as heavy metal hyperaccumulators, a term coined by Jaffré et al. (Jaffré et al., 1976). Both types combine a high tissue tolerance to toxic metals with a high ability to accumulate them. In either type, metals are detoxified by chelation in the cytosol, sequestration in vacuoles, or confinement in the apoplast. If metals are translocated from roots to above-ground tissues, via xylem, and distributed over aerial tissues, they are often compartmentalized or sequestrated in photosynthetically inactive tissues, like epidermis, or in storage tissues, such as trichomes and old leaves (Kupper et al., 2000). Plant heavy meal hypertolerance and/or hyperaccumulation are clear ecophysiological adaptations to metalliferous soils (Evangelou et al., 2004).

Whichever strategy plants use to limit the negative effects of metal toxicity, the metal response needs to be tightly regulated to reduce damage by toxic metals but ensure proper homeostasis of essential minerals. Plant responses to heavy metal stress are the combined results of cellular transport mechanisms and activation of signal transduction pathways. These processes are metal-dependent and plant species-dependent. In this review, we will discuss how plants respond at the cellular and molecular level to high exposures of zinc (Zn) and cadmium (Cd), two metals commonly found in the environment with toxic effects on plants and humans. In addition, we will discuss how plants regulate the response to metal exposure through signal transduction. Furthermore, we will compare the different molecular mechanisms and signal transductions between the majority of plant
species, which are heavy metal sensitive, and the few species that have evolved extreme adaptation to heavy metal exposure.

**Plant response to heavy metals**

When discussing the toxic effects of heavy metals, one should be aware that the toxicity of metals to plants depends on the external bioavailable metal concentration, the exposure time, the plant genotype, and the general condition of the plant. In addition, the dose-response curves to essential elements and non-essential elements are different (Fig. 1) (Alloway, 1995). Therefore, it is convenient to categorize heavy metals into two groups, the essential micronutrients and the non-essential elements. Heavy metal micronutrients, such as Co, Cu, Fe, Mn, Mo, Ni, and Zn, play an essential role in plant cell growth and development (Hansch&Mendel, 2009). For example, Zn is a cofactor of many enzymes through which it is involved in protein binding, regulation of enzyme activity, transcriptional regulation, translational regulation, and signal transduction (Broadley et al., 2007a). Because of the importance of essential micronutrients in plant physiology, it is imperative to maintain homeostasis of these heavy metals in plant cells. This means that metal concentrations should be maintained within a relatively narrow range to avoid deficiency and toxicity effects. Plants fortunately possess tight regulatory mechanisms to maintain heavy metal micronutrient homeostasis. Only when this mechanism has reached its capacity, and heavy metal micronutrient concentrations rise above a certain threshold level, their toxicity will be imminent (Appenroth, 2010). Non-essential heavy metals, such as Cd, Pb, Tl, and Hg, which are not known to have a biological function, are generally toxic to plants (Jarup, 2003). Cd for instance is a powerful enzyme inhibitor. Cd exposure also results in the degeneration of mitochondria, and in aberrations of mitosis, leading to inhibition of cell proliferation and cell division in meristems (Das et al., 1997). Cd also damages the photosynthetic apparatus, causing production of reactive oxygen species (ROS) in photosynthetically active tissues (Siedlecka&Krupa, 1996). These non-essential heavy metals can evoke a strong response of plant cells even when only applied at low concentrations (Appenroth, 2010). The molecular mechanism to exclude, detoxify, or compartmentalize non-essential heavy metals plays a crucial role in plant survival under heavy metal exposure.
The consequences of prolonged metal exposure for plant cells are membrane disintegration, ion leakage, lipid peroxidation, DNA/RNA degradation, and eventually cell death. At the cellular level, higher plants potentially use one or more of the following six ways to avoid or endure heavy metal exposure (Fig. 2) (Hall, 2002b). (1) Reduce metal bioavailability; (2) control metal influx; (3) chelate metals; (4) promote metal efflux; (5) compartmentalize and sequester metals; (6) detoxify metal-induced Reactive Oxygen Species (ROS). In the subsequent sessions, we will introduce the differential molecular mechanisms of plant responses to Zn and Cd. The differences in the respective regulatory mechanisms will be compared, including comparison between metal-sensitive and metal-tolerant plant species.
plant heavy metal stress response
Fig 2. Molecular mechanism in response to heavy metal stress in plants
When plants are exposed to high metal concentrations (in this case Zn, blue dots, and Cd, red dots), they can be discerned in four types depending on their response (A-D). (A) heavy metal sensitive plants, which cannot keep metals out of their roots nor prevent transport to the shoot, and which will succumb due to the toxic effects of metals on root and shoot cells; (B) heavy metal resistant excluder plants, which are able to keep metals outside the roots or take care of rapid efflux in case toxic metals have entered root cells; (C) heavy metal tolerant non-hyperaccumulator plants, in which metals can enter root cells where they are sequestrated into root vacuoles preventing translocation to shoots; (D) heavy metal hypertolerant hyperaccumulator plants, in which metals are actively taken up through the root, and largely loaded into xylem for root to shoot transport. In the shoot, the metals are safely sequestered in vacuoles. The molecular mechanism to keep metal homeostasis in plant cells is shown in (E). It involves metal influx transporters (dark purple), such as ZIP and NRAMP1 proteins, which are responsible for the uptake of metals into cytosol; metal tonoplast located transporters (green), such as MTP1, 3, ZIF1, HMA3, and ABCC1 and 2, which are responsible for metal (Zn or Cd), chelator (NA), or metal-chelator complex (Cd-PC) sequestration into the vacuole, or remobilization from the vacuole (NRAMP3/4); and metal efflux transporters (blue), such as HMA4, which acts to exclude excess metals out of the cytosol and is involved in the metal translocation towards the shoot. The cell wall is shown in brown, the vacuole is shown in yellow, and the cytosol is show in grey. ZIP, ZRT-IRT like Protein family; NRAMP, Natural Resistance Associated Macrophage Protein; MTP, Metal Tolerance Protein; ZIF1, Zinc-Induced Facilitator 1; HMA, Heavy Metal ATPase; YSL, Yellow Stripe 1-Like; PCR, Plant Cadmium Resistance; PDR8, Pleiotropic Drug Resistance 8; NA, Nicotianamine; GSH, Glutathione; PC, Phytochelatin; S, Sulphide; LMW, Low Molecular Weight; HMW, High Molecular Weight.

Reducing metal bioavailability
The bioavailability of metals in soil determines if metals are accessible for plants to be taken up or not. Bioavailability is largely determined by the rhizosphere. Rhizosphere microbes can change metal solubility, mobility, availability, specificity, precipitation by alteration of the soil pH, and secretion of metal chelators (Gadd, 2010). The presence of mycorrhiza, a symbiotic interaction between specific soil fungi and roots of most vascular plant species, can enhance heavy metal detoxification and tolerance of plants by selective metal exclusion and metal chelator secretion of the fungi (Jentschke&Godbold, 2000). For instance, the ectomycorrhizal fungus Paxillus involutus possesses specific Zn retaining capacity in its mycelium, which can immobilize Zn and prevent its transport to host plants, thus increasing the apparent Zn tolerance of its host, Scots pine (Pinus sylvestris).
There is also genetic variation among different mycorrhizal fungi genotypes, since a Zn-sensitive strain of the ectomycorrhizal fungus, *Suillus bovinus*, inoculated on Scots pine seedlings, provided a stronger Zn stress response (reduced chlorophyll concentration and inhibited N assimilation) compared to pine seedlings inoculated with a Zn-tolerant fungal strain (Adriaensen et al., 2006). The same holds for Cd tolerance. *Pinus pinaster* seedlings inoculated with the same fungus showed higher shoot biomass, less Cd-sensitivity symptoms, and lower shoot Cd concentrations than seedlings inoculated with *Rhizopogon roseolus* (Sousa et al., 2012). Next to mycorrhizal fungi, also bacterial microbes can affect plant metal exposure. Application of *Pseudomonas aeruginosa* to black gram (*Vigna mungo*) seeds or of *Methylobacterium* or *Burkholderia* sp. to tomato, can reduce Cd accumulation in plant roots and enhance plant growth (Ganesan, 2008). *P. aeruginosa* can also reduce the uptake of Cd in pumpkin and mustard by decreasing Cd bioavailability in the soil (Madhaiyan et al., 2007, Sinha&Mukherjee, 2008). To prevent the entrance of Zn into their plant host, ectomycorrhizal fungi tend to sequester Zn into their vacuoles, while Cd is bound to thiol-containing compounds such as glutathione and γ-glutamylcysteine (Courbot et al., 2004). It is clear that the effects of microbes and mycorrhiza to heavy metal stress response usually depends on the plant species, the microbial varieties or mycorrhiza types, and the metal. Several studies have recently been published on the selection of metal-hypertolerant microbes and mycorrhiza for bioremediation purposes, either in using microorganisms to remove metal pollutants (Gadd, 2010, Silar&Dairou, 2011), or indirectly by microorganisms that promote soil metal bioavailability thus enhance metal uptake by metal hyperaccumulator plants (Göhre&Paszkowski, 2006, Meier et al., 2012, Miransari, 2011).

Another way of preventing entry of metals into the plant is by binding metals to specific root exudates or to the plant cell walls. Root exudates, including low molecular weight compounds (amino acids, organic acids, sugars, phenolics) and high molecular weight compounds (polysaccharides and proteins), can be actively secreted or passively leaked from root cells to the soil (Bais et al., 2006). The presence of root exudates affects the metal availability either by direct chelation of metals or indirectly by acidification of the rhizosphere and reduction of elements (Bertrand&Poirier, 2005). Thus, root exudates can enhance or inhibit metal uptake by plant roots and further regulate metal tolerance and accumulation of plants. For instance, barley cultivar “Sahara” accumulates more Zn than
“Clipper” because of the higher root exudation of organic acids (malate, maleate, fumarate, and cis-aconitate) and amino acids (alanine, valine, proline, aspartic acid, and glutamic acid) (Rasouli-Sadaghiani et al., 2011). Exudation can also account for differences between species, such as in the comparison between Oenothera picensis, with higher organic acid root exudation than Imperata condensate, accounting for its higher tolerance to elevated soil Cu concentrations (Meier et al., 2011).

The cell wall, the first plant structure getting in touch with metals, contains suberin and low-methylesterified pectin. Suberin is often found in root endodermis and exodermis cell walls, where it acts as a barrier to control the uptake of water and mineral ions to provide control over nutritional mineral accumulation and transportation (Baxter et al., 2009). Low-methylesterified pectin comprises polysaccharides for the binding of divalent and trivalent metal ions (Krzesłowska, 2011). Most essential and non-essential metals are commonly entering plants as divalent cations (Zn$^{2+}$, Cd$^{2+}$, Fe$^{2+}$, Mn$^{2+}$ or Cu$^{2+}$). Thus, altering the composition of the cell wall to enhance binding of metal cations to the cell wall can contribute to reducing the inadvertent uptake of toxic metals by plants. Zn-treated tomato suspension cells for instance produced a higher cell wall biomass, with higher Zn-retaining ability by cell wall polymers and higher Zn-binding capacity compared to non-treated suspension cells (Muschitz et al., 2009). When comparing hyperaccumulating (HE) and non-hyperaccumulating ecotypes (NHE) of Sedum alfredii, both showed similar Zn absorption abilities, however, the Zn affinity to root cell walls was different (Li et al., 2007): Zn bound to root cell walls of HE plants was more available for xylem loading than NHE. This implies that metal binding capacity and affinity of the cell wall can be modified by the plant, and thus contribute to preventing metals from entering the plant. Modification of cell wall lignins can be studied at the transcriptomic level using microarray analysis, focusing on expression of lignin biosynthesis related genes, such as 4-coumarate-CoA ligases 4CCL1, 4CCL2, 4CCL3, 4CCL8, 4CCL9 and 4CCL14; cinnamoyl-CoA reductase CCR2; and hydroxycinnamyl alcohol dehydrogenase CAD1 (Boerjan et al., 2003, Neutelings, 2011, Wagner et al., 2009). When comparing root gene expression under deficient, sufficient, and excess Zn conditions between the Zn-hypertolerant and hyperaccumulating species Noccaea (formerly Thlaspi) caerulescens, and the non-tolerant and non-hyperaccumulating -related species Arabidopsis thaliana, a large group of lignin biosynthesis-related genes are expressed at higher levels in N. caerulescens. This may relate to the formation of
the extra endodermis reinforcements that are found in roots of *N. caerulescens* (Broadley *et al.*, 2007b, Van De Mortel *et al.*, 2006). A similar phenomenon was found in response to Cd stress. When comparing differential gene expression of roots under Cd exposure from *N. caerulescens* (Cd tolerant), and *A. thaliana* (Cd sensitive), lignin biosynthesis genes and cell expansion-related genes were found to be up-regulated in *A. thaliana* under excess Cd conditions, while *N. caerulescens* maintained its constitutively high expression of these genes (Van De Mortel *et al.*, 2008). This means that also Cd-sensitive species like *A. thaliana* make use of lignin depositions as a physical barrier to prevent Cd to enter the plant. Metal precipitation is another way to restrict metal entry, as was found in the heavy metal hyperaccumulator *Arabidopsis halleri*, in which precipitates of Zn and Cd phosphates accumulate in the cell walls of the root epidermis (Kupper *et al.*, 2000). These examples demonstrate the potential role of structural modifications of root cell walls to reduce the level of bioavailable metals and thus induce heavy metal tolerance.

**Control of metal influx**

Metals can enter the plant symplast in several ways: by simple diffusion, by passive transport through channel proteins, or by active transport through carrier proteins, the latter of which is most important, at least offers most control. These carrier proteins are generally referred to as metal transporters. They often have different affinities for different heavy metals, which is why they can confer some level of discrimination at the plasma membrane between desired essential elements and unwanted non-essential elements. When plants face an excess of Zn or Fe, they will rapidly try to inhibit excessive metal influx through down-regulation of the relevant transporters at the transcriptional and often also post-transcriptional or post-translational level. Fe status, for instance, alters the function of the *IRT1* transporter through transcriptional and post-translational regulation (Barberon *et al.*, 2011, Kerkeb *et al.*, 2008, Vert *et al.*, 2002). Zn and Cd are most likely crossing the plasma membrane via members of the ZIP transporter family (ZRT-IRT like protein; Zinc-regulated transporter, Iron-regulated transporter Protein) (Guerinot, 2000). In *A. thaliana* this family comprises 15 genes, *ZIP1* to *ZIP12* and *IRT1* to *IRT3*). The *A. thaliana ZIP4* gene appears to be an important factor to control Zn import, as it encodes a plasma membrane-localized Zn specific transporter, which expression is upregulated under Zn deficiency, and repressed by excess Zn in *A. thaliana* (Assunção *et al.*, 2010, Grotz *et al.*, 1998, Van De Mortel *et al.*, 2006). Interestingly, the *ZNT1* gene, a *ZIP4* orthologue from *N. caerulescens*, is much
higher expressed, almost irrespective of the Zn exposure levels, when compared to its orthologues in *A. thaliana* or *Thlaspi arvense*, another related non-hyperaccumulator (Assunção *et al.*, 2001, Van De Mortel *et al.*, 2006), which is well in accordance with the constitutive Zn hyperaccumulation in *N. caerulescens*. In addition to ZIP4, also the ZIP1, ZIP3, ZIP5, ZIP9, ZIP10, ZIP12 and IRT3 genes of *A. thaliana* appear to be involved in Zn-uptake, as their expression in roots is induced in response to Zn deficiency (Assunção *et al.*, 2010, Talke *et al.*, 2006, Van De Mortel *et al.*, 2006). Expression of all of these ZIP genes is under control of two basic-region leucine zipper (bZIP) transcription factors, bZIP19 and bZIP23, which act redundantly to control the initial Zn deficiency response of *A. thaliana* (Assunção *et al.*, 2010). Whether the control of ZIP Zn transporters in metal hyperaccumulators is also controlled by the two bZIP transcription factors is not yet known. Although these transcription factors in *A. thaliana* hardly respond at the transcriptional level to alterations in Zn exposure levels, the expression of their target ZIP genes is tightly controlled, with strongly reduced expression at high Zn exposure.

Still, also at elevated levels, Zn is able to enter the root cells, through other metal uptake transporters or in a more passive way. Zn excess often evokes a Fe deficiency response, with induced expression of Fe uptake transporters like *IRT1*. Since these also have affinity for Zn, inadvertent uptake of Zn cannot be completely avoided. This means that the essential mineral homeostasis mechanisms are better at controlling metal uptake under deficiency conditions than at high exposure levels.

Cd chemically resembles Zn and Fe. To date, there is no Cd-specific influx transporter found for plant cells and the uptake of Cd is likely to occur through available metal uptake ZIP transporters (or alike) which have high specific transport affinity for Zn or Fe, but also low affinity for Cd (Korshunova *et al.*, 1999, Pence *et al.*, 2000). The uptake of Cd in root cells thus appears to be an opportunistic event. Still, the Zn/Cd specialized hyperaccumulating Ganges accession of *N. caerulescens*, shows a much higher maximum Cd influx than the Cd-tolerant, and Cd-excluding, accession Prayon, which results in a fivefold higher Cd concentration in xylem sap of Ganges (Lombi *et al.*, 2001). This accession also shows strong induction of *IRT1* under Fe-deficiency conditions, which may account for the high Cd uptake, although this has not been proven conclusively (Lombi *et al.*, 2002, Plaza *et al.*, 2007).
Metal chelation

Metals that get past the plasma membrane need to be chelated by various ligands in order to reduce their undesired interaction with cellular compounds. Such ligands can be oligopeptides, organic acids, amino acids, or proteins. Metal-ligand complexation is an important part of the molecular mechanism of metal homeostasis. Therefore, the production of metal chelators is of importance in the plant metal exposure response. Heavy metal hyperaccumulating species usually show higher chelator accumulation than non-hyperaccumulating species, which supports the importance of chelators in metal detoxification, metal sequestration, and metal efflux. The next section will particularly address the role of Zn and Cd chelators.

Nicotianamine (NA) is a nonproteinogenic amino acid that shows high binding affinity in vitro to a range of transition metals, such as Cu, Ni, Co, Zn, Fe, and Mn. The formation of Zn-NA complexes has also been detected in vivo, in Schizosaccharomyces pombe (Beneš et al., 1983, Trampczynska et al., 2010). Metal-NA complexes can be transported over cellular membranes by YSL proteins, which are thus important components in the regulation of metal homeostasis in plants (Curie et al., 2009). NA is synthesized from three molecules of S-adenosylmethionine by the enzyme nicotianamine synthase (NAS) (Higuchi et al., 1999). The expression of NAS genes determines the NA concentrations and thus contributes to NA-mediated metal homeostasis. NAS genes are found in varying copy numbers in different plant species. Solanaceous species appear to have only one NAS gene, Graminae species mostly have three NAS genes, and in A. thaliana (and other Brassicaceae), there are four NAS genes: NAS1, 2, 3, 4, which act functionally redundant in metal homeostasis (Klatte et al., 2009). Differences in NAS1, NAS2, and NAS3 gene expressions are seen when comparing the Zn/Cd hyperaccumulator species A. halleri and N. caerulescens to related non-accumulators (Hammond et al., 2006, Van De Mortel et al., 2006, Weber et al., 2004). In correspondence with the NAS transcript levels, A. halleri has a higher NA concentration in roots than A. thaliana (Deinlein et al., 2012). Knocking down NAS2 gene expression in A. halleri by RNA interference (RNAi) results in reduced NA contents in roots, which inhibits the root-to-shoot translocation of Zn and renders these plants virtually non-hyperaccumulators (Deinlein et al., 2012). In addition to YSL transporters, which transport NA-metal complexes over the plasma membrane (Curie et al., 2009, Gendre et al., 2007), recently a tonoplast localized transporter, zinc-induced facilitator 1 (ZIF1), was shown to be involved.
in transporting NA from the cytosol into the vacuole, thus facilitating NA–Zn complex formation in the vacuoles (Haydon et al., 2012). In A. thaliana, the ZIF1 gene is up-regulated when exposed to excess Zn (Haydon & Cobbett, 2007) and plants tend to retain the Zn in the vacuoles of root cells. As a result, less Zn is available for root-to-shoot translocation of Zn, preventing Zn to accumulate in the shoots where it is potentially more harmful than in roots due to interference with the photosynthetic machinery. The formation of NA–Zn complexes contributes to Zn loading into xylem in Zn hyperaccumulators, and turns out to be a useful storage form of chelated Zn in roots of non-accumulators under excess Zn exposure. NAS, YSL, and ZIF1 genes, controlling NA synthesis and (metal-chelated) membrane transport, form an important regulatory mechanism of controlling metal distribution of the plant.

Glutathione (GSH), a γ-Glu-Cys-Gly tripeptide, is important because of its tripartite role in metal detoxification as metal chelator, a cellular antioxidant, and as ROS signaling molecule (Jozefczak et al., 2012, Seth et al., 2012). GSH can act as metal chelator through its thiol groups, which have high metal binding affinity. The expression of GSH synthesis genes, GSH1 and GSH2, is induced by Cd treatments in A. thaliana, contributing to Cd tolerance, while a decrease in GSH levels reduces Cd tolerance (Semane et al., 2007, Wójcik & Tukiendorf, 2011). Also in rice, Cd-tolerant plants have higher levels of GSH than Cd-sensitive plants (Cai et al., 2011). In addition, GSH plays an essential role in Fe-mediated Zn tolerance in A. thaliana (Shanmugam et al., 2012). GSH acts as a moderator of cellular oxidation status, as it can exist in a reduced (GSH) and an oxidized form (GSSG). The reduced form can donate a reducing equivalent (H⁺ + e⁻) to unstable molecules like ROS. The resulting reactive GSH will soon react with another reactive GSH to become GSSG. It thus also acts as a ROS signaling molecule, mainly determined by the GSH:GSSG ratio, which reflects the oxidative state of the plant cell (Jozefczak et al., 2012). The transition between GSH and GSSG also affects activity of several antioxidant enzymes, such as glutathione reductase (GR), ascorbate peroxidase (APX), and catalases (Seth et al., 2012). Under Cd treatment, a decreased GSH/GSSG ratio was observed in A. thaliana, which was accompanied by enhanced GR and APX activities (Semane et al., 2007). The hyperaccumulating ecotype of Sedum alfredii shows a higher GSH/GSSG ratio and less ROS production than the non-hyperaccumulating ecotype under excess Cd (Tian et al.). In summary, GSH plays an important role in the detoxification of heavy metals and metal-induced oxidative stress response.
**Phytochelatins** (PCs) are oligomers of GSH (Glu-Cys)\textsubscript{n}-Gly (n=2–10), synthesized by the enzyme pythochelatin synthase (PCS). The synthesis of PCs was initially found to be essential for tolerance to Cd, as concluded from analysis of \textit{pcs} (\textit{cad}) mutants defective in PCS (Howden \textit{et al.}, 1995a, Howden \textit{et al.}, 1995b, Larsson \textit{et al.}, 2002) but later it was also shown to be relevant for tolerance to excess Zn (Tennstedt \textit{et al.}, 2009). When Cd enters plant cells, low molecular weight (LMW) PCs will first form a PC–Cd complex. This LMW PC–Cd complex is subsequently sequestered into vacuoles by ATP-binding cassette (ABC) transporters; for example, two ABCC-type transporter from \textit{A. thaliana}, \textit{ABCC1} and \textit{ABCC2} can enhance Cd tolerance and accumulation through vacuolar sequestration of PC–Cd (Park \textit{et al.}, 2012, Song \textit{et al.}, 2010b). LMW PC–Cd complexes bind sulfides to form a stable, high-molecular weight (HMW) PC–Cd complex, which is stored in the vacuole (Clemens, 2006, Cobbett&Goldsbrough, 2002, Hall, 2002a). While Cd detoxification by PC–Cd complexation and vacuolar storage is important for most plant species (Cobbett, 2000), it does not appear to play a role in Cd detoxification in Cd hyperaccumulator species (de Knecht \textit{et al.}, 1992, Schat \textit{et al.}, 2002, Shah, 2011). When characterizing and comparing PCS between the Cd-hyperaccumulators \textit{A. halleri} and \textit{N. caerulescens}, and the non-hyperaccumulator \textit{A. thaliana}, it turned out that the \textit{PCS1} gene from \textit{A. halleri} and \textit{N. caerulescens} showed lower expression than its orthologue from \textit{A. thaliana}, in accordance with PC accumulation data (Meyer \textit{et al.}, 2011). This demonstrates that PCs do not constitute the major Cd detoxification pathway in Cd hyperaccumulators. Instead of PCs, the induction of antioxidative mechanisms appears to play a more important role in conferring Cd tolerance. For example, the root superoxide dismutase (SOD) activity, leaf peroxidase (POD) activity, catalase (CAT) activity, and free proline concentrations are higher in the Cd-tolerant species \textit{Solanum nigrum}, compared to the non-hyperaccumulator \textit{S. melongena}; while there is no indication of elevated PC levels in the tolerant species (Sun \textit{et al.}, 2007). Similarly, investigations of PC synthesis and Cd accumulation in the Cd-hyperaccumulating ecotype of \textit{Sedum alfredii} showed that Cd is hardly retained in roots, but transported and accumulated in leaf cell walls, in which PCs take no part (Zhang \textit{et al.}, 2010).

For metal non-hyperaccumulating species, the balance between GSH and PCs is important for Cd tolerance. For example, overexpressing the \textit{AtPCS1} gene can result in both Cd-sensitive and Cd-tolerant transgenic plants (Lee \textit{et al.}, 2003a, Lee \textit{et al.}, 2003b). The explanation is that elevated synthesis of PCs results in the
depletion of GSH, which contributes to enhanced oxidative stress and the Cd-
sensitive phenotype. However, if synthesis of PCs is not tremendously increased,
though sufficient for increased PC–Cd complexation, the transgenic plants can
maintain their GSH levels for sufficient contribution to Cd-induced antioxidative
stress response and consequently they become Cd-tolerant (Seth et al., 2012). In
conclusion, PC accumulation plays an important role in Cd detoxification in non-
tolerant, non-hyperaccumulating plant species, but is of little importance in Cd
hyperaccumulators, which appear to rely more on efficient Cd sequestration and
enhanced ROS detoxification.

**Metallothioneins (MTs)** are small cysteine-rich proteins that are found in
most eukaryotes and contain metal-binding motifs that provide sulfhydryl for
interacting with bivalent metal ions (Cobbett & Goldsbrugh, 2002). Based on
the type of cysteine residues, plant MTs are classified into four types. In *A. thaliana*
there are six MTs, belonging to four types: *MT1a, MT2a, MT2b, MT3, MT4a*, and
*MT4b*. Yeast complementation experiments with these six *MTs* showed that most
can enhance tolerance to and accumulation of Cu. Only *MT4* types confer Zn
tolerance and accumulation and *MT1, 2*, and *3* types enhance tolerance to Cd, but
often not Cd accumulation (Guo et al., 2008). In *A. thaliana, MT1a* is responsible
for Cu homeostasis under elevated Cu, and is required for Cd tolerance, Cd
accumulation, and Zn accumulation (Guo et al., 2008, Zimeri et al., 2005).

Phytohormones abscisic acid (ABA) and gibberellic acid (GA) regulate the
contribution of *MT4a* and *MT4b* to Zn accumulation in seeds and Zn nutrient
supplementation of young seedlings (Ren et al., 2012). MTs and PCs can work
cooporatively to protect *A. thaliana* from Cu and Cd toxicity (Guo et al., 2008).
*M1* and *MT2* are expressed at much higher levels in *N. caerulescens* compared to
*A. thaliana*, implying that they are important for metal tolerance (Roosens et al.,
2005). Indeed, *MT1, 2*, and *3* from *N. caerulescens* can confer Cd tolerance to
yeast or increase the intracellular Cd concentrations (Hassinen et al., 2007,
Roosens et al., 2005). The higher *NcMT2* expression in metallicolous *N.
caerulescens* accession La Calamine, compared to the non-metallicolous accession
Lellingen, is thought to contribute to the metal-adapted phenotype through
improved Cu homeostasis at high Zn and Cd exposure (Hassinen et al., 2007,
Hassinen et al., 2009). MTs not only play a role as metal chelators but also as ROS
scavengers to reduce oxidative stress. The redox sensing residue on MT, cysteine,
acts as a ROS scavenger (Green & Paget, 2004, Hassinen et al., 2011). Zn-induced
*MT1a* can improve rice stress tolerance through the regulation of zinc-finger
transcription factors via alternating Zn homeostasis, and also participate in the ROS scavenging pathway by altering the levels of antioxidant enzymes catalase, peroxidase and ascorbate peroxidase (Yang et al., 2009). Also rice MT2b is an reactive oxygen scavenger involved in the \( \text{H}_2\text{O}_2 \) signaling pathway (Steffens&Sauter, 2009). Similar functions are found for \( A. \text{thaliana} \) MT2a and cotton MT3a (Xue et al., 2009, Zhu et al., 2009). MT3 from \( \text{Tamarix hispida} \), is a ROS scavenger that contributes to increased Cd, Zn, and Cu tolerance; also, the enhanced activity of superoxide dismutase, catalase, and glutathione peroxidase under metal treatments participates in scavenging of ROS (Yang et al., 2011). The exact role of MTs as ROS scavenger in metal hyperaccumulator species is still not clear.

**Promotion of metal efflux**

Another solution to overcome excessive entering of toxic metals into plant cells is to release these metals again from the cells. They can either be returned back to the soil solution or remain in the apoplast. The direction of metal efflux corresponds to the metal accumulation phenotype of the plant. In non-tolerant, non-hyperaccumulating plants, root efflux transporters direct metals to the soil solution, while in heavy metal hyperaccumulators, the efflux system is directed towards loading of metals into the xylem, on their way to the shoot.

Zn efflux is an important factor in plant Zn homeostasis, and thus Zn efflux transporters have been found. However, so far, no Cd-specific efflux transporter has been found in plants, which means that Cd efflux always accompanies the transport of other metals by action of efflux transporters with higher affinities for other metals than Cd. An important cellular metal efflux transporter family is the \( \text{P}_{1\text{B}} \)-type ATPase family (Williams&Mills, 2005). One prominent Zn efflux transporter is HMA4 (Heavy Metal ATPase 4). This plasma membrane-localized transporter is normally responsible for loading Zn into the xylem. It is involved in Zn and Cd uptake, as found in \( A. \text{thaliana} \) where it is upregulated in roots under elevated Zn exposure but repressed by Cd exposure (Mills et al., 2003). A \( hma4 \) null mutant results in low Zn and Cd translocation ability, while overexpressing \( HMA4 \) enhances root tolerance to Zn, Cd, and Co and increases Zn and Cd accumulation (Verret et al., 2004). Another \( \text{P}_{1\text{B}} \)-type ATPase, HMA2, is equally important for Zn homeostasis in plants (Eren&Argüello, 2004). In \( \text{Arabidopsis} \), \( HMA2 \) and \( HMA4 \) act redundantly in Zn and Cd root-to-shoot translocation (Hussain et al., 2004, Wong&Cobbett, 2009). HMA4 also appears to be the major
determinant in explaining root-to-shoot transport of Zn and Cd in the heavy metal hyperaccumulators *A. halleri* and *N. caerulescens*. The gene is much higher expressed in these species compared to related non-accumulators (Courbot *et al.*, 2007, Hammond *et al.*, 2006, Hanikenne *et al.*, 2008, Van De Mortel *et al.*, 2006). This appears to be caused by local gene multiplication events in both species (Ó Lochlainn *et al.*, 2011). RNAi-mediated knock-down of *HMA4* gene expression in *A. halleri* resulted in loss of the Zn/Cd hypertolerance and hyperaccumulation ability in *A. halleri*, including increased sensitivity to high Zn and Cd exposure and reduced root to shoot translocation efficiency (Hanikenne *et al.*, 2008).

In *A. thaliana* there are other efflux transporters involved in Zn or Cd redistribution, translocation, and detoxification, such as the ones encoded by the *Plant Cadmium Resistance 1* and 2 (*PCR1* and *PCR2*) genes. Activity of the plasma membrane-localized Cd efflux transporter, PCR1, enhances Cd tolerance by exporting Cd out of the cell and thus reducing Cd contents (Song *et al.*, 2004). Also, PCR2 functions as a Zn efflux transporter, which contributes to Zn distribution and detoxification in *A. thaliana* (Song *et al.*, 2010a). Another plasma membrane-localized transporter, pleiotropic drug resistance 8 (*PDR8*), confers Cd tolerance in *A. thaliana* by pumping Cd$^{2+}$ across the plasma membrane, out of root epidermal cells (Kim *et al.*, 2007). It not clear if these transporters are also relevant in heavy metal hyperaccumulating species.

**Metal sequestration and remobilization**

If a plant is not able to prevent entry or enhance efflux, it will have to face the symplastic entry of metals and deal with it. This means plants need to sequester the metals at “safe” sites. This can either be in the cell, in specific organelles, in storage tissues, or even outside the cells. To give an example of the latter, in *N. caerulescens*, the Cd-hyperaccumulating ecotype Ganges can store much more Cd in the cell walls of epidermal cells than the poor Cd-accumulating ecotype Prayon (Cosio *et al.*, 2005). Metal sequestration is an important strategy to reduce the cytoplasmic metal concentrations. Preferred organelle is the vacuole, which provides a well-controlled internal storage reservoir, normally already acting as storage buffer for nutrient minerals to account for temporary deficiencies (Vögeli-Lange&Wagner, 1990). Tonoplast transporters, which not only transport metals but also chelators (e.g., ZIF1 for NA) (Haydon *et al.*, 2012), are needed for vacuolar import. Non-tolerant species mostly promote enhanced vacuolar sequestration in the roots, thus preventing transport to photosynthetic leaves where metals can be
potentially more harmful than in roots. In contrast, the heavy metal adapted species have strong metal sequestration abilities in shoot cells and reduce root vacuolar sequestration (Schneider et al., 2013). This difference in compartmentalization promotes metal loading into the xylem and creates a safe metal storage sink in leaves. Several proteins are involved, of which we will discuss the most prominent ones.

The Metal Tolerance Protein 1 (MTP1), belongs to the cation diffusion facilitator (CDF) protein family and is probably the most important Zn vacuolar sequestration transporter in plants. It acts as a tonoplast located Zn$^{2+}$/H$^+$ antiporter (Desbrosses-Fonrouge et al., 2005, Kawachi et al., 2008, Kobae et al., 2004). In heavy metal hyperaccumulators like Thlaspi (currently Noccaea) goesingense, N. caerulescens, A. halleri and also the hyperaccumulating ecotype of Sedum alfredii, this gene is higher expressed in shoots than in comparable non-hyperaccumulators (Assunção et al., 2001, Gustin et al., 2009, Peer et al., 2003, Persans et al., 2001, Zhang et al., 2011). The higher MTP1 expression in A. halleri compared to its close non-hyperaccumulator relatives A. lyrata or A. thaliana, is provided by several additional copies of the gene, distributed over four loci, compared to only one locus in the other two species (Shahzad et al., 2010). Two of these loci are co-segregating with zinc tolerance QTLs in a back cross population between both species (Dräger et al., 2004). Next to MTP1, tonoplast localized MTP3 also contributes to Zn tolerance and Zn sequestration in response to excess Zn in A. thaliana (Arrivault et al., 2006), but the role of MTP3 in hyperaccumulators is still not clear. Other MTPs, such as MTP8 and MTP11, show increased gene expression in N. caerulescens and A. halleri, compared to A. thaliana, which suggests they may be important for heavy metal homeostasis and tolerance in these species (Chiang et al., 2006, Hammond et al., 2006, Krämer et al., 2007, Talke et al., 2006, Van De Mortel et al., 2006).

Heavy Metal ATPase 3 (HMA3) is another member of the P$_{1B}$-type ATPase superfamily to which HMA2 and HMA4 belong (Leonhardt et al., 2012). It was found to be a prominent candidate for Cd sequestration. The HMA3 transporter from rice, isolated from the low Cd-accumulating rice cultivar Nipponbare, is located at the tonoplast of root cells to limit Cd root-to-shoot translocation through selective sequestration of Cd into the root vacuoles (Ueno et al., 2009, Ueno et al., 2010). HMA3 also confers Cd vacuolar storage in A. thaliana and overexpression in this species enhances Cd, Co, Pb and Zn tolerance, probably through regulation of the vacuolar sequestration of these metals, even
though actual Zn transport of this protein has not been confirmed in yeast (Gravot et al., 2004, Morel et al., 2009). The gene is much higher expressed in shoots of *A. halleri* than in *A. thaliana*, independent of Zn exposure levels. Expression of the *A. halleri* HMA3 in yeast increased Zn tolerance, suggesting it to be also involved in regulation of cellular Zn status in *A. halleri* (Becher et al., 2004). Similarly, the HMA3 protein of *N. caerulescens* is much higher expressed in the Cd-hyperaccumulating accession Ganges, than in the poor Cd accumulator Prayon. Like for HMA4 and MTP1, copy number variation appears to largely account for the difference in expression level between both accessions. Unlike HMA3 from *A. thaliana*, HMA3 from *N. caerulescens* is not found to transport Zn, only Cd (Ueno et al., 2011). The HMA3 expression pattern in *N. caerulescens* also deviates from that in *A. thaliana*, with higher expression in leaf epidermis and mesophyll cells, which is in line with the shoot Cd hyperaccumulation ability of *N. caerulescens*.

Next to HMA3, the chloroplast envelope-located **HMA1 transporter** is found to be involved as metal sequestration transporter. It acts as a Cu-ATPase to import Cu into chloroplasts and also contributes to Zn detoxification by reducing Zn content in the plastid (Kim et al., 2009, Seigneurin-Berny et al., 2006). In yeast, the ATPase activity of AtHMA1 can be activated by Zn, Cu, Cd, and Co and the increased activity enhances metal tolerance of yeast (Moreno et al., 2008). Knocking down HMA1 gene expression in *A. thaliana* results in reduced Cu/ZnSOD enzyme activity, which means it is involved in the antioxidant defense (Seigneurin-Berny et al., 2006). Although the function of HMA1 in hyperaccumulator species has not been studied in detail, enhanced HMA1 expression is observed in *A. halleri* where it may contribute to Zn hypertolerance (Becher et al., 2004).

When examining the gene expression of transporters involved in metal efflux (HMA4) and metal sequestration (HMA3 and MTP1), these genes show higher expression in hypertolerant and/or hyperaccumulating species than in non-tolerant and/or non-accumulating species. The enhanced expression is often due to multiple copies of the genes, which is likely to be caused by gene duplication during evolution of metal exposure adaptation. A study of speciation between hyperaccumulator *A. halleri*, and non-hyperaccumulator *A. lyrata*, indicates that the historical split between both species coincides with the initial duplication of the HMA4 gene, which suggests that this may have contributed to, if not caused, the separation of both species (Roux et al., 2011).
The Natural Resistance-Associated Macrophage protein (NRAMP) family, contains six members in A. thaliana. Although these proteins are involved in the regulation of heavy metal homeostasis, they have quite different roles to play. The metal influx AtNRAMP1 transporter confers Fe, Mn, and Cd uptake in yeast (Thomine et al., 2000). In A. thaliana, this plasma membrane-localized transporter works as a major high-affinity Mn transporter, needed for acquisition of Mn from the soil (Cailliatte et al., 2010). AtNRAMP 3 and AtNRAMP 4 act as tonoplast localized metal efflux transporters, needed for vacuolar remobilization of Fe and Mn, but also contributing to Cd tolerance. They are essential for remobilization Fe during seed germination at low Fe supply and to maintain optimal photosynthesis and plant growth at low Mn supply (Lanquar et al., 2005, Lanquar et al., 2010). AtNRAMP3 can transport Cd and AtNRAMP4 can transport both Zn and Cd. Overexpressing AtNRAMP4 results in a Zn- and Cd-hypersensitive phenotype under Fe deficiency in A. thaliana (Lanquar et al., 2004). The N. caerulescens NRAMP3 and NRAMP4 show the same metal transport abilities as their A. thaliana orthologues; however, NcNRAMP3 and NcNRAMP4 are expressed at higher levels than the A. thaliana NRAMPs, and thus contribute to enhanced root to shoot metal transport (Oomen et al., 2009, Wei et al., 2009). AtNRAMP6 is localized to a vesicular-shaped endomembrane compartment and functions as an intracellular metal transporter that regulates the distribution and availability of Cd within A. thaliana (Cailliatte et al., 2009). Its role in heavy metal hyperaccumulation is still not clear.

Detoxification of metal-induced Reactive Oxygen Species

Heavy metal toxicity induces the production of Reactive Oxygen Species (ROS), such as superoxide (\(O_2^-\)), hydroxyl radicals (OH\(^*\)), hydrogen peroxide (H\(_2\)O\(_2\)) and singlet oxygen (\(^1\)O\(_2\)), which are highly reactive molecules that interact with various cellular components leading to oxidative damage to macromolecules like nucleic acids, proteins, sugars, and lipids, causing oxidative stress in plant cells and (intra)cellular membrane damage, that could result in cell death in severe cases (Gadjev et al., 2008). To resist such oxidative stress, plant cells possess a comprehensive antioxidant system to scavenge ROS and detoxify them. This is mainly conferred by the action of several antioxidant enzymes, such as superoxide dismutase (SOD), catalase (CAT), ascorbate peroxidase (APX), and glutathione reductase (GR) (Matilla-Vázquez & Matilla, 2012). SOD is the only enzyme able to scavenge O\(_2^-\). CAT, APX, and GR play a crucial role in the scavenging and detoxification of H\(_2\)O\(_2\). Thus, the presence and concentration of these antioxidant
enzymes indicates the status of the antioxidative defence in plant cells. In heavy metal-sensitive plant species, ROS levels will rise substantially if there are not sufficient antioxidant enzymes available. Consequently, ROS-induced cellular damage will induce local programmed cell death and will generally affect plant growth and development. Heavy metal tolerant species normally produce high levels of ROS scavenging antioxidant enzymes, providing a sufficiently efficient antioxidant defense mechanism against heavy metal-induced oxidative stress (Sharma & Dietz, 2009).

Zn plays a dual role in the ROS-induced oxidative stress in plant cells. Zn excess leads to ROS production, while Zn is also an important cofactor of SOD, which catalyses the removal of O_2\(^\cdot\) and thus protects cells from ROS-induced damage (Cakmak, 2000). Zn is not a redox-active metal, which means it is not directly involved in the Haber–Weiss and Fenton reactions, as are known for Fe. Nevertheless, excess Zn leads to the production of OH\(^\cdot\) and CH_3\(^\cdot\) radicals, as observed in cell walls of Verbascum thapsus, and the accumulation of H_2O_2 in the leaf apoplast (Morina et al., 2010). Increased activity of antioxidant enzymes SOD and APX and also monodehydroascorbate reductase (MDHAR) accounts for the accumulation of ROS-scavenging ascorbate and phenolics, thus enhancing the survival of V. thapsus under excess Zn exposure-induced oxidative stress (Morina et al., 2010). Excess Zn-induced toxicity in Brassica napus was found to be caused by intervention in its nutrient balance and induction of oxidative stress due to decreased SOD and APX activities (Wang et al., 2009). Higher activities of the antioxidant enzymes SOD, CAT, APX, GR, and DHAR also accounted for the higher Zn tolerance of Zn hyperaccumulating ecotypes of S. alfredii, compared to non-hyperaccumulating ecotypes, under high Zn exposure (Jin et al., 2008).

The Zn- and Cd-induced antioxidative mechanism in plant cells is mainly acting through enhancing activity or presence of antioxidant enzymes or increasing accumulation of antioxidants. For example, the enhanced presence of antioxidative enzymes (such as CAT and SOD) under Cd stress in Cd hyperaccumulating species/ecotypes like N. caerulescens, Brassica juncea, S. alfredii and S. nigrum results in less ROS accumulation and increased Cd detoxification ability than in the related non-hyperaccumulating species/ecotypes, such as N. tabacum, S. alfredii and S. melongena (Semane et al., 2010, Tian et al., Wang et al., 2008).

The ROS response mechanism induced by Zn or by Cd exposure can act through different defense pathways. For example, the Zn/Cd hyperaccumulator
Arabis paniculata adapts to Zn excess stress by enhancing expression of proteins involved in energy metabolism and protein metabolism to accelerate plant growth and correct misfolded proteins; but it resists Cd exposure stress by promoting the antioxidative defense and cellular metabolism to maintain cellular redox homeostasis (Zeng et al., 2011). Similarly, N. caerulescens under high Cd exposure shows higher APX activity but lower SOD activity than when under high Zn exposure (Wójcik et al., 2006). Although Cd and Zn may induce the antioxidant mechanisms by different pathways, Zn plays a synergistic role in Cd-induced antioxidant defense because of its role as an enzyme cofactor. Zn supplementation enhances the activities of antioxidant enzymes SOD, CAT, APX and GR, and increases the accumulation of antioxidants like ascorbic acid and GSH, as found in Cd-treated Solanum lycopersicum, S. alfredii and Triticum aestivum (Cherif et al., 2011, Jin et al., 2009, Sanaeiostovar et al., 2012). Next to Zn, also low Mg can have a synergistic effect on Cd toxicity in plants, probably by maintaining Fe status, increasing the antioxidative capacity and protecting the photosynthetic apparatus (Hermans et al., 2011).

**Signal transduction of heavy metal stress**

Next to controlling cellular metal contents, by the activity of various transporters, and reducing the toxic effects due to ROS generation, plants need to sense the presence of metals. More specifically, they need to sense alterations in the cytoplasmic and organellar concentrations of heavy metals to anticipate overaccumulation and prepare for detoxification. In this section, we report on heavy metal-induced signal transduction in response to excess metal exposure. When talking about the signaling transduction, a basic scheme should be kept in mind (Fig. 3), meaning that plant stressors (such as Cd) can induce signaling perception through an external signal (such as calcium or miRNA), upon which the signals are rapidly transmitted (e.g., through MAP kinases) to the responsible transcription factors. These will interact with gene promoter elements to induce the required gene expression response (Lichtenthaler, 1998).

**Dual role of ROS**

Not only Zn but also ROS play a dual role in metal stress response. ROS act both as oxidative molecules, aggressively reacting with cellular macromolecules, and as signal transduction molecules (Sandalio et al., 2012). For example, overproduction of H$_2$O$_2$ leads to serious oxidative damage and is thus a danger to cellular function. However, H$_2$O$_2$ is also an important signaling molecule, which regulates plant
development, hormone signaling, programmed cell death, and biotic and abiotic stress response and tolerance (Matilla-Vázquez & Matilla, 2012). Thus, keeping control over ROS generation in plant cells during metal exposure is important to keep control of developmental processes and general stress response. In this section, we will further outline the role of ROS in heavy metal stress signal transduction and response.

Different metals induce different pathways to regulate the induction of ROS signals (Sharma & Dietz, 2009). Although the mode of Zn-regulating ROS signal transduction is still not clear, the action of another heavy metal micronutrient, Cu, has been well studied. Excess Cu induces the accumulation of ROS either directly, by acting as a redox-active molecule, or indirectly, by inhibiting microRNA398 (miR398) expression, which was found to be regulated by SQUAMOSA promoter-binding (like) protein binding to the GTAC(T) motif of the miR398 promoter (Nagae et al., 2008, Yamasaki et al., 2009). Cd-induced ROS accumulation is mediated and indirectly regulated by NADPH oxidase and Ca, but also includes involvement of miR398. NADPH oxidase is localized on the plasma membrane and produces superoxide (O$_{2}^{−}$) from oxygen (O$_{2}$) when challenged with excess metals (Garnier et al., 2006, Groppa et al., 2012, Remans et al., 2010). Excess Cd also inhibits the expression of miR398, which normally targets copper/zinc superoxide dismutase genes $CSD1$ and $CSD2$ for mRNA degradation.
NADPH oxidase

ROS sensor: OXI1, EX1, EX2

Cytosol

MKKs
- (ANP1; MKK1; unknown)

MKS
- (MKK4/MKK5; MKK1/MKK2; MKK3)

MPKs
- (MPK3/MPK6; MPK4; MPK7)

Nucleus

Cis- and Trans-regulation

TF1
- Ch 1: metal sensitive related genes

TF2
- Ch 2: metal resistant related genes

Tolerant
- chitinases and HSPs
- reduce metals in cells
- induce SOD, CAT, APX, GR
- accumulate GSH, PCs, NA, MTs

Sensitive
- inhibit growth
- suppress photosynthesis
- induce senescence
- DNA and protein damage
- program cell death

plant heavy metal stress response
Fig 3. Signal transduction in response to heavy metal stress in plants

Heavy metals, such as Cd (red), but also Cu (blue) can affect ROS accumulation through three ways, (1) excess Cd\(^{2+}\) or Cu\(^{2+}\) induces miR398 expression which inhibits the function of Cu/ZnSOD (CSD) and further induces ROS accumulation; (2) excess Cd\(^{2+}\) inhibits the regulatory role of Ca\(^{2+}\), which also stimulates ROS accumulation through inhibited CSD activity; (3) excess Cd\(^{2+}\) enhances NADPH oxidases, which leads to additional H\(_2\)O\(_2\) production. The accumulated ROS are detected by ROS sensors, the OXI1, EX1, and EX2 proteins, and used as a signal to induce MAPK (Mitogen-Activated Protein Kinase; MPK) cascades (green rectangle), which are a series of phosphorylations from MKKKs (MAPK kinase kinases) to MKKs (MAPK kinases), and then to MPKs (MAP kinases). The MAPK cascades can activate functions of transcription factors (trans-regulatory elements) in the nucleus, which regulates gene expression through the binding to cis-regulatory elements, and further controls the plant response to heavy metals, to be sensitive or tolerant. MAPK cascade-mediated gene expressions in ROS signalling pathways are well studied by Pitzschke et. al., (Pitzschke et al., 2009). ROS, Reactive Oxygen Species; OXI1, Oxidative Signal-Inducible 1; EX1 or 2, Executer 1 or EX2;SOD, Superoxide Dismutase; CAT, Catalase; APX, Ascorbate Peroxidase; GR, Glutathione Reductase; GSH, Glutathione; PCs, Phytochelatins; NA, Nicotianamine; MTs, Metallothioneins; ANP1, Arabidopsis Nicotiana protein kinase 1-related kinase 1.

These encode the most important SODs in plant cells (Sunkar et al., 2006, Zhou et al., 2008). Reduced miR398 results in increased CSD 1 and CSD2 expression, which enhances CSD accumulation, and subsequent detoxification of superoxide into less toxic H\(_2\)O\(_2\) in mitochondria (Fridovich, 1995, Kliebenstein et al., 1998, Sunkar et al., 2006, Zhu et al., 2011). Thus, the reduction of miRNA398 induces ROS detoxification and antioxidant accumulation, which further reduces ROS accumulation in plant cells (Ding&Zhu, 2009). Cd-induced production of ROS can also be mediated by a different process, involving protein phosphatases, Ca\(^{2+}\) channels, and cGMP, as found in pea (Romero-Pertas et al., 2004). Excess Cd can lead to Ca\(^{2+}\) deficiency, which results in the downregulation of antioxidant enzymes CAT and CSD (Costa et al., 2010, Rodríguez-Serrano et al., 2009); as a result, less ROS are scavenged.

ROS and stress response

Heavy metal stress-induced ROS production and accumulation in plant cells is detected by the OXI1 (Oxidative Signal-Inducible 1) protein for H\(_2\)O\(_2\), and the EX1 (Executer 1) or EX2 proteins for superoxide (Lee et al., 2007, Rentel et al., 2004). These proteins can stimulate the rapid signal transduction of a downstream Mitogen-Activated Protein Kinase (MAPK) cascade. A MAPK cascade is a series
of phosphorylation steps from MKKKs (MAPK kinase kinase), via MKKs (MAPK kinase), to MPKs (MAPK) (Jonak et al., 2002). MAPKs are serine/threonine kinases, which can phosphorylate a large number of transcription factors (trans-regulatory elements), such as DREB (Dehydration-Responding Element Binding proteins), bZIP (basic region leucine ZIPper), NAC (NAM, ATAF1,2, CUC2), AP2 (Activator Protein 2), RAV (Related to ABA-insensitive3/Viviparous1), WRKY (containing a conserved WRKYGQK domain and a zinc-finger-like motif), MYB (MYeloBlastosis), bHLH (basic Helix-Loop-Helix), Zat (Zinc finger, C2H2-EAR-motif-containing repressor), and GRAS (GAI, RGA, SCR), and further change the expression of cis-regulatory elements, such as IDRS (Iron-Dependent Regulatory Sequences), and ABA (abscisic acid), SA (salicylic acid) or IAA (auxin) responsive elements (Hong-Bo et al., 2010, Mittler et al., 2004). For example, activation of OXI1, a serine/theonine kinase, is required for full activation of MAP kinases MPK3 and MPK6 (Rentel et al., 2004). Colcombet and Hirt (Colcombet&Hirt, 2008) proposed that ROS act upstream of several MAPK cascades, one of which is the H$_2$O$_2$-activated ANP1 (Arabidopsis Nicotiana protein kinase 1-related kinase 1) phosphorylation of M KK4/MKK5 and subsequently MPK3/MPK6; a second one is the H$_2$O$_2$-activated MEKK1 phosphorylation, via MKK1/MKK2, to MPK4 (Rodriguez et al., 2010); and finally, the H$_2$O$_2$-activated phosphorylation of an as yet unknown component via MKK3, to MPK7. The activation MAPK cascade is plant-species and metal variety-dependent. In alfalfa (Medicago sativa) seedlings, excess Cu can activate four MAPKs (SIMK, SAMK, MMK2, MMK3) responses, and does this faster than Cd. It can also activate SIMKK, which Cd cannot (Jonak et al., 2004). In A. thaliana, Cd can trigger activation of MPK3/MPK6 in a ROS dosage-dependent way (Liu et al., 2010). In rice, increasing myelin basic protein kinase (MBP) activities by excess Zn exposure can be inhibited by ROS treatments (Lin et al., 2005).

The signal coming from a MAPK cascade will be further transduced to transcription factors, which can bind to specific cis-regulatory elements and induce metal responsive gene expression. In metal-sensitive plants, the MAPK cascades activate trans- and cis-regulatory mechanisms that are insufficient to provide metal-tolerance, while in metal tolerant plants, MAPK cascades successfully activate an adequate metal tolerance mechanism. In metal-sensitive plants, the Cd-ROS-MAPK signal causes damage through different processes: (1) interruption of the hormonal signaling pathway; (2) suppression of photosynthesis; (3) damaging macromolecules; (4) induction of senescence; and (5) induction of programmed
cell death. Interfering with hormone signaling will inhibit plant growth and development. For example, Cd-induced ROS can activate auxin oxidase, which degrades auxin, and changes the auxin-regulated morphogenetic response in A. thaliana rosette leaves (Blomster et al., 2011, Elobeid&Polle, 2012). The effect on photosynthesis includes restraining photosystem II (PSII) activity, inhibiting the PSII photoreaction, lowering photophosphorylation, reducing the activity of chloroplast enzymes RuBPC and phosphoribulokinase, decreasing photosynthetic pigments (such as total chlorophyll content and chlorophyll a/b ratio), diminishing net photosynthesis in leaves, and reducing chloroplast metabolism (Clijsters&Assche, 1985). Macromolecules like proteins and nucleic acids are damaged by oxidation, impaired DNA repair, and poor protein folding. Enhanced lipid peroxidation causes membrane damage that may lead to cell death (Ercal et al., 2001, Sharma et al., 2008). Senescence is induced through H$_2$O$_2$-induced MEKK1 expression, which enhances the expression of transcription factor WRKY53, a positive regulator of leaf senescence in A. thaliana (Miao et al., 2004, Miao et al., 2007). Finally, the metal-induced production of H$_2$O$_2$ acts as signaling molecule to trigger the expression of the WRKY75, Zat11 and NAM transcription factors that stimulate programmed cell death in plants (Gechev&Hille, 2005).

In heavy metal-adapted plants, the Cd-induced ROS-MAPK response is different and does not lead to the same damage as observed in metal-sensitive plants. This appears to be due to the heavy metal-induced accumulation of repair proteins, such as chitinases, and molecular chaperones, such as heat shock proteins (HSPs). Different isoforms of chitinases are expressed in response to metal treatments, implying a specific role in the metal response mechanism (Békésiová et al., 2008). Overexpressing fungal chitinases in tobacco or A. thaliana can confer Cd tolerance in transgenic plants (Brotman et al., 2012, de las Mercedes Dana et al., 2006). Induction of HSPs in response to oxidative heavy metal stress is thought to involve the molecular chaperone activity of HSPs, to prevent misfolding of proteins, protein aggregation, and the degradation of (denatured) proteins under stress (Gupta et al., 2010). Cd stress induces the expression of several HSP proteins, such as HSP70s chaperonin, and mitochondrial HSP60s (Sarry et al., 2006). The interaction between OsHSP70 and MAPKs protects rice roots from Cu toxicity (Chen et al., 2008).
Conclusion

When facing heavy metal stress in nature, plants have evolved one of four different responses to deal with the stress (Fig. 2a). These responses are guided by the molecular mechanism in plant cells (Fig. 2b) and by a series of signal transductions (Fig. 3). ROS play a dual role in these mechanisms and affect the plant response to heavy metal stress. Most plant species are not sufficiently armed against heavy metal stress, which means their response leads to metal sensitivity, exhibited by inhibited root growth and leaf chlorosis. Large amounts of toxic metals will enter plant cells, which have no sufficient mechanisms to detoxify them. Thus, excess metals induce accumulation of ROS, which results in a further negative impact on plant growth and development, suppressed photosynthesis, damage to nucleic acids and proteins, enhanced programmed cell death, and induction of senescence. The other response to metal stress is metal tolerance, in which plant growth and development are not much affected by metal stress, at least plants do not succumb. To resist heavy metal exposure, the entrance of metals into cells needs to be avoided or otherwise regulated; toxic metals ending up in the cytosol should be efficiently detoxified; metal-induced accumulation of ROS should be immediately scavenged; the cell membrane system should be protected; and the injured cells should be quickly repaired. Plant cells can detoxify metal-induced damage through the accumulation of repair proteins (chitinases and HSPs), antioxidants (GSH), antioxidant enzymes (SOD, CAT, APX, and GR) and accumulation of metal chelators (PCs, NA, and MTs).

Metal-tolerant plants usually take one of three strategies to maintain metal homeostasis. One involves heavy metal exclusion, in which metals are excluded outside the cells or bound to cell walls, to reduce the metal concentration in the cell. Another involves sequestration of heavy metals, as performed by heavy metal-tolerant non-hyperaccumulators, in which metals are taken up by root influx transporters and subsequently sequestered into vacuoles or rapidly exported outside the cell. The third is practiced by heavy metal-hypertolerant hyperaccumulator species, in which metals are very efficiently detoxified in vacuoles by several tonoplast transporters, while large amounts of metals are transported from roots to shoots via xylem. Interestingly, the required enhanced expression of involved transporters in heavy metal hyperaccumulators is often caused by gene copy number multiplication, which appears to act prior to subsequent differentiations in gene expression regulation.
Chapter 3

The expression of the ZNT1 zinc transporter from the metal hyperaccumulator *Noccaea caerulescens* confers enhanced zinc and cadmium tolerance and accumulation to *Arabidopsis thaliana*

Zeshan Hassan¹*, Ya-Fen Lin¹*, Sangita Talukdar¹*, Liu Hong¹, Alfred Arulandhu¹, Henk Schat² and Mark G.M. Aarts¹a

¹Laboratory of Genetics, Wageningen University and Research Centre, Droevendaalsesteeg 1, 6708 PB, Wageningen, The Netherlands
²Ecology and Physiology of Plants, Faculty of Biology, Free University, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands

* authors have contributed equally
ABSTRACT

Plants maintain their Zn homeostasis by regulated expression of transition metal transporters. The objective of this research was to perform a functional analysis of the NcZNT1 gene Noccaea caerulescens and a comparative promoter study of NcZNT1 from N. caerulescens and AtZIP4 from Arabidopsis thaliana. In this study, NcZNT1 was found to be higher expressed under Zn deficient conditions in N. caerulescens and localized to plasma membrane. The p35S::NcZNT1 A. thaliana lines showed enhanced tolerance to Zn and cadmium (Cd) excess, increased accumulation of Zn and Cd and up-regulation of the Fe deficiency response compared to wild type line. Besides, pAtZIP4::GUS and pNcZNT1::GUS transformed A. thaliana showed similar GUS staining but only under Zn deficiency while pNcZNT1::GUS transformed N. caerulescens roots had a constitutive expression compared to Zn deficiency induced expression of pAtZIP4::GUS. Putative cis elements in both promoters were identified by 5’ deletion analysis. Furthermore, both pAtZIP4::GFP and pNcZNT1::GFP expressing roots showed GFP expression in cortex, endodermis, pericycle cells and vascular tissues when expressed in A. thaliana roots under Zn deficient conditions. In N. caerulescens, expression of pNcZNT1::GFP was higher than pAtZIP4::GFP, especially under Zn sufficient conditions. In conclusion, NcZNT1 is clearly an important factor in Zn and Cd tolerance and accumulation in N. caerulescens. Differences in cis- and trans-regulators are likely to account for the differences in expression between A. thaliana and N. caerulescens. NcZNT1 expression in the stele of N. caerulescens roots proposes its involvement in long distance metals transport by maintaining the metal influx into cells responsible for xylem loading and ultimately shoot translocation of the metals.

KEYWORDS

Zinc, cadmium, ZNT1, ZIP4, hyperaccumulation, Noccaea caerulescens
INTRODUCTION

Zinc (Zn) is an essential component of several enzymes in plants like RNA polymerase, alcohol dehydrogenase, Cu/Zn superoxide dismutase and carbonic anhydrase (Marschner, 1995; Guerinot & Eide, 1999). Also, a large number of proteins contain Zn-binding structural domains such as the Zn finger domain. Poor growth and less biomass are among the major Zn deficiency symptoms which lead to reduced crop yields (Marschner, 1995). Although Zn is essential for plants, supra optimal concentrations of Zn can be toxic. The toxic effects are due to uncontrolled binding of Zn to proteins and cofactors and rendering them non-functional (Eide, 2003). It is known that cellular damaging Reactive Oxygen Species (ROS) are highly induced under excess Zn conditions in plants (Cuypers et al., 1999). Leaf chlorosis and growth reduction was reported due to Zn toxicity (Marschner, 1995). Cadmium (Cd) is toxic for plants having no known biological function but can be taken up by Zn transporters due to similarity with Zn (Pence et al., 2000). Cd is not able to initiate ROS production directly but enhanced ROS levels were found under Cd exposure (Cuypers et al., 2009). Cd toxicity has harmful effects in plants as it disturbs DNA repair mechanism, reduced water and nutrient uptake, lowered photosynthesis and ultimately leaf chlorosis and reduction in plant growth (Banerjee & Flores-Rozas, 2005; Sanita di Toppi & Gabbrielli, 1999).

In order to deal with fluctuations in metal concentrations, plants have evolved “metal homeostasis” which is the ability to regulate their cellular and organellar metal concentration to maintain a stable and constant condition (Hassan & Aarts, 2011).

There are few species that hyperaccumulate Zn from the soil and store it in their leaves. Zn hyperaccumulator species are defined to accumulate more than 10,000 µg Zn g⁻¹ of dry weight (dw) (1%, w/w) (Baker & Brooks, 1989), whereas most plants contain between 30 and 100 µg Zn g⁻¹ dw and concentrations above 300 µg Zn g⁻¹ dw are generally toxic (Marschner, 1995). Another characteristic of metal hyperaccumulators is that most of the hyperaccumulated metals are found in the shoots rather than in the roots, whereas generally plants try to reduce the shoot heavy metal concentration to avoid toxicity and negative interference with photosynthesis. Two of these Zn hyperaccumulators, Noccaea caerulescens and Arabidopsis halleri, were examined at the transcriptional level, which showed that both species generally express genes constitutively that are normally induced by Zn.
deficiency and at higher levels than related non-hyperaccumulators (Becher et al., 2004; Hammond et al., 2006; van de Mortel et al., 2006).

*Noccaea caerulescens* is an exceptional metal hyperaccumulating species, as it is the only one to hyperaccumulate Zn (30,000 mg kg$^{-1}$), Ni (4000 mg kg$^{-1}$ DW) and Cd (2700 mg kg$^{-1}$ DW) in shoots (Brown et al., 1995; Assunção et al., 2003a). In addition, there is substantial natural variation regarding metal specificity and metal tolerance among different accessions. Accessions originating from serpentine soils are generally good at accumulating Ni. Accessions from calamine or non-metallicolous soils are good Zn accumulators. Cd hyperaccumulation so far has only been observed among populations found in the south of France, in the region around Ganges (Lombi et al., 2000; Assuncão et al., 2003a; Roosens et al., 2003). *N. caerulescens* is one of the few known Cd hyperaccumulator species together with *A. halleri* and *Sedum alfredii* (Yang et al., 2004). It belongs to the Brassicaceae family and shares about 88.5% coding region sequence similarity with the dicot plant reference species *A. thaliana* (Rigola et al., 2006).

Understanding the mode of action of plant metal hyperaccumulation is interesting for evolutionary and applied biology reasons. Most of the metal hyperaccumulation traits evolved relatively recently. For instance within the Brassicaceae family, metal hyperaccumulation is found in at least three genera: *Arabidopsis, Noccaea* (previously known as *Thlaspi*) and *Alyssum*. Although these genera are related (Koch et al., 2009), it is unlikely that their common ancestor was hyperaccumulating, considering there are only few hyperaccumulator species within the lineages derived from such common ancestor. It means that this trait has evolved several times independently of each other and thus this facilitates the study of the molecular origin of such a drastic adaptive evolution. This will provide interesting insight in the selective mechanisms that are prone to such evolutionary changes. The applied interest in metal hyperaccumulation resides in the use of metal hyperaccumulator plants for the remediation of metal polluted soils known as “phytoremediation” (Reeves&Baker, 2000). A disadvantage of the plant species that are currently used for Zn/Cd phytoremediation is that either their biomass is insufficient to support economically viable phytoremediation projects, or their metal extraction capacity is too low (Hassan&Aarts, 2011). Furthermore, “biofortification” is the fascinating concept of improving the human micronutrients deficiencies (e.g. Zn and Fe) by developing crops having improved bioavailable micronutrient content (Palmgren et al., 2008). With increased knowledge on the mode of action of Zn and Cd tolerance, uptake, translocation and accumulation in
Zn/Cd hyperaccumulating species, it may be possible to engineer Zn/Cd hyperaccumulation and tolerance in a high-biomass species for Zn/Cd phytoremediation or in a crop species useful for Zn biofortification purposes. In order for this to be efficient, both the genes involved and their (post-)transcriptional regulation should be known and optimized.

So far, a number of metal transporters are identified and shown to be involved in the metal hyperaccumulation process (Hassan & Aarts, 2011). Previously, we have cloned the ZNT1 gene from *N. caerulescens*, encoding a ZIP-like transporter (Assunção et al., 2001). NcZNT1 resembles most the *AtZIP4* gene from *A. thaliana* with 90% cDNA and 87% amino acid identity. Heterologous expression in yeast showed it to mediate high-affinity Zn uptake and low-affinity Cd uptake (Pence et al., 2000). The *AtZIP4* gene was not studied in great detail, but expression is known to be strongly induced in roots and shoots under Zn deficient conditions (Grotz et al., 1998; van de Mortel et al., 2006). Previously, it was reported that NcZNT1 is predominantly expressed in roots and less in shoots in *N. caerulescens*, but expression of this gene is barely responsive to changes in Zn supply (Assunção et al., 2001). Only at very high Zn concentrations, the expression is somewhat reduced (Pence et al., 2000; van de Mortel et al., 2006). This deregulation of the *N. caerulescens* gene compared to its *A. thaliana* orthologue may be part of the metal adaptation phenomenon of the former species.

Recently Milner and colleagues (2012) have reported that NcZNT1 is able to transport Zn but not Cd, contradicting its previously known Zn and Cd transport ability (Pence et al., 2000). *A. thaliana* lines expressing NcZNT1 were found to be sensitive to excess Zn but not to Cd. However, these authors used what appears to be a 5' truncated NcZNT1 cDNA. In the current study we have performed a detailed analysis of transgenic *A. thaliana* lines expressing NcZNT1 and of the NcZNT1 promoter, to investigate the function of NcZNT1 and its role in metal hyperaccumulation or tolerance. We have determined the response of NcZNT1 gene expression to changes in Zn supply. The phenotype of *A. thaliana* lines expressing full length NcZNT1 cDNA under control of the strong CaMV 35S promoter was examined and the gene expression of known metal transporters in these lines was analysed. To further determine the expression pattern of NcZNT1 and AtZIP4 genes, we have used full promoter-GUS fusion constructs of these genes and have studied GUS expression in *A. thaliana* and *N. caerulescens* under different metal exposure conditions. The possible function of NcZNT1 gene and the relevance for Zn/Cd hyperaccumulation of *N. caerulescens* is discussed. We
conclude that *NcZNT1* plays an important role in Zn and Cd tolerance and accumulation and is involved in establishing a high metal influx into the root vasculature, important for xylem-mediated translocation of metals to the shoot.

**MATERIALS AND METHODS**

*Isolation of AtZIP4 and NcZNT1 promoters*

Genomic DNA was extracted from flowers of *A. thaliana* (accession Columbia) and *N. caerulescens* (accession La Calamine), as described by Aarts *et al.* (2000). To amplify the sequence containing the *AtZIP4* promoter, a PCR reaction was performed on genomic DNA of *A. thaliana* using primers P5 and P15 (Table 1). To amplify the sequence containing the *NcZNT1* promoter, two forward primers (P1 and P2; Table 1) were designed on the gene upstream of *AtZIP4* (*At1g10970*) in *A. thaliana*. This gene, At1g10980, is predicted to encode a “membrane protein PTM1 precursor isolog” by TAIR (www.arabidopsis.org). The reverse primer (P3; Table 1) was designed on the 5’ end of the *NcZNT1* cDNA (GenBank acc. No. AF275751)(Assunção *et al.* 2001). PCR fragments were cloned into pGEM-T-easy (Promega, Leiden, The Netherlands) and plasmids from several colonies for each fragment were sequenced to confirm their identity.

*Isolation of the promoters of AtZIP4-like genes from Cochleria pyrenaica, Arabidopsis halleri and Arabidopsis lyrata*

In order to isolate the *ZIP4* promoters from three related species, *C. pyrenaica* (accession La Calamine, Belgium), *A. halleri* (accession Auby, France) and *A. lyrata* (accession Unhost, Central Bohemia, Czech Republic; Macnair *et al.*, 1999), the latter two kindly provided by Dr. Pierre Saumitou-Laprade (CNRS, Lille, France), genomic DNA was isolated from them and PCR amplified using the same primers as used for amplification of the *NcZNT1* promoter. PCR fragments were cloned into pGEM-T-easy (Promega, Leiden, The Netherlands) and plasmids from several colonies for each fragment were sequenced to confirm their identity.

*Construction of binary plasmids*

To generate a construct encoding a chimeric fusion protein of N-terminal GFP and *NcZNT1*, a *NcZNT1* cDNA fragment of 1289 bp was amplified using primers P27 and P28 (Table 1) and Pfu DNA polymerase (MBI Fermentas, St. Leon-Rot, Germany). The PCR fragment was digested with HindIII and SalI and ligated into
HindIII/SalI digested binary vector pEZR(H)-LN (Narvaez-Vasquez, Pearce and Ryan, 2005) containing the HPT gene for hygromycin resistance.

The p35S::NcZNT1 construct was made by cloning a 1.48-kb NcZNT1 (accession La Calamine, Belgium) cDNA fragment (Assuncao et al., 2001) upon restriction digestion with XbaI and HindIII into the pGD121 (de Folter et al., 2006) vector harbouring the nptII gene for selection on kanamycin resistance upon transformation.

Promoter GUS for stable transformation in Arabidopsis

For making the pAtZIP4::GUS construct, a 1048-bp A. thaliana genomic DNA fragment obtained by PCR using P5 and P15 (Table 1) was digested with HindIII and NcoI and cloned at the ATG start codon of the uidA (GUS) gene in the HindIII-NcoI digested pCAMBIA1301 vector (http://www.cambia.org/daisy/bios/585.html) (Ge et al., 2004) replacing the CaMV 35S promoter. This vector contains the HPT gene for selection on hygromycin resistance in plants. This construct was named F05.

For construction of the pNcZNT1::GUS construct, the NcZNT1 promoter was cloned from N. caerulescens genomic DNA using primers P1 and P3, as described above and the cloned fragment was reamplified with primers P16 and P26 (Table 1) to create suitable restriction sites for cloning. Cloning was performed in the same way as described for the AtZIP4 promoter. This construct was named F16.

In order to generate 5’ deletions of the AtZIP4 and NcZNT1 promoters, 10 forward primers containing appropriate HindIII and NcoI sites for cloning, were designed based on the promoter sequences of AtZIP4 and NcZNT1 (P5 through P14 for AtZIP4 and P16 through P25 for NcZNT1) (Table 1). Each construct was named after the forward primers. In total, 22 constructs, 11 for AtZIP4 and 11 for NcZNT1, were obtained following the same method as described for the F05 and F16 constructs. Two more constructs were made by digestion of the F05 and F16 vectors with SalI, and subsequent self-ligation. These deletion constructs were named F15 in case of NcZNT1 and F26 in case of AtZIP4. All constructs were verified by DNA sequencing.
Table 1. Primers used for the PCR amplification

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence of oligonucleotides (5’-3’)</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>5’-ATCGGGCATGATCATGAGGAA-3’</td>
<td>Forward primer for NcZNT1 promoter isolation; designed on At1g10980 gene</td>
</tr>
<tr>
<td>P2</td>
<td>5’-CCTCTTTGGCCCTACGGAAGA-3’</td>
<td>Forward primer for NcZNT1 promoter isolation; designed on At1g10980 gene</td>
</tr>
<tr>
<td>P3</td>
<td>5’-TTATAAGATCAATCTAATAACA-3’</td>
<td>Reverse primer for NcZNT1 promoter isolation; designed on cDNA of NcZNT1</td>
</tr>
<tr>
<td>P4</td>
<td>5’-TAAAGTCAGGCACCAAATGCGAGTG-3’</td>
<td>Reverse primer on NcZNT1 cDNA</td>
</tr>
<tr>
<td>P5</td>
<td>5’-GTAAGCTTTGGAAGATGTAAGTGATG-3’</td>
<td>Forward primer for AtZIP4 promoter isolation</td>
</tr>
<tr>
<td>P6</td>
<td>5’-CCAAGCTTTAGCTCTTGGTCGATGTAAC-3’</td>
<td>Forward primer on AtZIP4</td>
</tr>
<tr>
<td>P7</td>
<td>5’-ATAAGCTTCCACTGAGAACCCCTGCA-3’</td>
<td>Forward primer on AtZIP4</td>
</tr>
<tr>
<td>P8</td>
<td>5’-TGAAGCTTCTCCATCTTACAAGATTACCCTGCTC-3’</td>
<td>Forward primer on AtZIP4 promoter</td>
</tr>
<tr>
<td>P9</td>
<td>5’-TGAAGCTTCTTGTAGCTCTTGAATCTCCTT-3’</td>
<td>Forward primer on AtZIP4 promoter</td>
</tr>
<tr>
<td>P10</td>
<td>5’-TGAAGCTTCC GTGAGTAAGTTAACCTCCTT-3’</td>
<td>Forward primer on AtZIP4 promoter</td>
</tr>
<tr>
<td>P11</td>
<td>5’-TGAAGCTTCTTGGGAACCAATGTTGCCTTGGT-3’</td>
<td>Forward primer on AtZIP4 promoter</td>
</tr>
<tr>
<td>P12</td>
<td>5’-GAAAGGCTTAAACCCCATGACCCACCGATC-3’</td>
<td>Forward primer on AtZIP4 promoter</td>
</tr>
<tr>
<td>P13</td>
<td>5’-GAAAGGCTTACCTCTTCTTACAAGCTGCTT-3’</td>
<td>Forward primer on AtZIP4 promoter</td>
</tr>
<tr>
<td>P14</td>
<td>5’-GAAAGGCTTTACCTTCTCTACAGTCGCTT-3’</td>
<td>Forward primer on AtZIP4 promoter</td>
</tr>
<tr>
<td>P15</td>
<td>5’-ATCGAGGAAGAACTGATGAAAGGAAAGAT-3’</td>
<td>Reverse primer for AtZIP4 promoter isolation</td>
</tr>
<tr>
<td>P16</td>
<td>5’-ATATCGAGTTCTGACTTGGGCTTTTGA-3’</td>
<td>Forward primer on NcZNT1 promoter</td>
</tr>
<tr>
<td>P17</td>
<td>5’-TTGAGCTTACATTCGATCTCGTCC-3’</td>
<td>Forward primer on NcZNT1 promoter</td>
</tr>
<tr>
<td>P18</td>
<td>5’-GGAAGCTTTGACTGCTAAATGCGAGAAGA-3’</td>
<td>Forward primer on NcZNT1 promoter</td>
</tr>
<tr>
<td>P19</td>
<td>5’-GGAAGCTGCATTGAAAACATCCCTGAAAAAC-3’</td>
<td>Forward primer on NcZNT1 promoter</td>
</tr>
<tr>
<td>P20</td>
<td>5’-TGAAGCTTCTGTTGAGTGGATAGTGTTAA-3’</td>
<td>Forward primer on NcZNT1 promoter</td>
</tr>
<tr>
<td>P21</td>
<td>5’-TGAAGCTTTCTTTTGACTGTTAACC-3’</td>
<td>Forward primer on NcZNT1 promoter</td>
</tr>
<tr>
<td>P22</td>
<td>5’-AGGAAGCTTACCTGACCAATTTCTGTCAAC-3’</td>
<td>Forward primer on NcZNT1 promoter</td>
</tr>
<tr>
<td>P23</td>
<td>5’-TGAAGCTTACATACGAAAAAGAGAGAATACAG-3’</td>
<td>Forward primer on NcZNT1 promoter</td>
</tr>
<tr>
<td>P24</td>
<td>5’-AACAGGCTTTCTGCTGATGTTACCTCCTTGTT-3’</td>
<td>Forward primer on NcZNT1 promoter</td>
</tr>
<tr>
<td>P25</td>
<td>5’-ACCGAGTTTGACAAAGTTGCTGAAAGAAGA-3’</td>
<td>Reverse primer on NcZNT1 promoter</td>
</tr>
<tr>
<td>P26</td>
<td>5’-ACCGAGTGACACCAGGAGAAACAA-3’</td>
<td>Forward primer on NcZNT1 promoter</td>
</tr>
<tr>
<td>P27</td>
<td>5’-CCGAAGCTTACCCAAACAAAGAGATCGTCC-3’</td>
<td>Forward primer on cDNA clone pad-GAL4-2.1 vector with HindIII</td>
</tr>
<tr>
<td>P28</td>
<td>5’-TAAAGTCCAGGCCCCAAATTGCGAGTGTCG-3’</td>
<td>Reverse primer for the NcZNT1 cDNA</td>
</tr>
<tr>
<td>P29</td>
<td>5’-TTCCATGATACATCCGCAT-3’</td>
<td>Forward primer designed on NcZNT1 cDNA</td>
</tr>
<tr>
<td>P30</td>
<td>5’-AACAGTCTGATTACCTTACTGCTGTT-3’</td>
<td>Forward primer for PCR amplification of At-tubulin</td>
</tr>
<tr>
<td>P31</td>
<td>5’-GGTTGGAGTTGACTGATCGATC-3’</td>
<td>Reverse primer for PCR amplification of At-tubulin</td>
</tr>
</tbody>
</table>

Restriction sites incorporated in the primers are underlined.

**Promoter GUS for root transformation in Noccaea**

To generate the pAtZIP4::GUS construct, the AtZIP4 promoter (1046 bp) was amplified from *A. thaliana* (accession Colombia) by using forward primer, 5’-CACCCTTTGGAAGATGGAAGTG-3’ and reverse primer 5’-GGGAACAAAGAGATTATTC-3’. To develop the pNcZNT1::GUS construct, the
NcZNT1 promoter (896 bp) was amplified from *N. caerulescens* (accession La Calamine, Belgium) by using forward primer, 5’-CACCTCTGACTCTTTATCTGGCCT-3’ and reverse primer 5’-GGGAACAAAGAGTGTCTTCTTC-3’. The amplified fragments were cloned separately into pENTR™/D-TOPO® vectors (Invitrogen™, cat. K2400-20). These entry vectors were recombined into the binary destination vector, pKGWFS7-RR, by using the Gateway® LR Clonase™ Enzyme Mix (Invitrogen™, cat. 11791-019). pKGWFS7-RR contains GUS as reporter protein and pAtUBQ10::DsRed as a selection marker, which was used to identify transformed roots based on red fluorescence under a stereo microscope using a DsRed filter (Op den Camp *et al.*, 2011; Karimi *et al.*, 2002). The destination constructs were sequenced to confirm correct cloning of the NcZNT1 and AtZIP4 promoters.

The *pAtZIP4::eGFP* and *pNcZNT1::eGFP* constructs were developed based on the pEZR(H)-LN vector (Narvaez-Vasquez, Pearce and Ryan, 2005), which harbours the eGFP gene for GFP expression and the hygromycin resistance HPT gene as a selection marker for transformation. The *pNcZNT1::eGFP* construct was made by replacing the CaMV35S promoter region of pEZR(H)-LN with the NcZNT1 promoter upon restriction digestion with HindIII and NcoI. Similarly, the *pAtZIP4::eGFP* construct was developed by cloning *pAtZIP4* into pEZR(H)-LN upon restriction digestion with SacI and NcoI.

**Transient NcZNT1-GFP expression in Cowpea protoplasts**

Cowpea protoplasts were prepared and transfected by the *p35S::NcZNT1-GFP* construct as described previously by Shah *et al.* (2002).

**Plant transformation and growth conditions**

**Stable transformation in A. thaliana**

*A. thaliana*, accession Columbia (Col), was transformed with *p35S::NcZNT1, pAtZIP4::GUS, pNcZNT1::GUS, pAtZIP4::eGFP, pNcZNT1::eGFP* and all the deletion constructs of *AtZIP4 and NcZNT1* promoters by the *Agrobacterium tumefaciens*-mediated flower dipping transformation method as described by Clough and Bent (1998). T₁ transformed seedlings were selected on ½ MS agar plates (Murashige and Skoog, 1962) (no sugar, pH 5.8) supplemented with 50 mg L⁻¹ kanamycin or 20 mg L⁻¹ hygromycin (Duchefa Biochemie B.V., Haarlem, The Netherlands) at 24 °C (16/8 hr, light/darkness). 50 independently transformed plants were tested for *NcZNT1* expression by semi-quantitative RT-PCR (data not
shown) and 10 high expressing lines were propagated until homozygous T₃ lines and these lines for each construct were used for experimentation. Plates were incubated in a climate-controlled growth cabinet (25°C 16/8 hr, light/darkness with illumination at a light intensity of 120 µmol m⁻² s⁻¹).

**Root transformation in N. caerulescens**

*N. caerulescens* roots were transformed with *pNcZNT1::GUS, pAtZIP4::GUS, pNcZNT1::eGFP* and *pAtZIP4::eGFP* constructs using modified *Agrobacterium rhizogenes* mediated transformation method as described by (Limpens *et al.*, 2004). Seeds of *N. caerulescens* were sterilized and gminated on ½ MS agar plates (no sugar, pH 5.8) at 24 °C (16/8 hr, light/darkness). Seven-day-old seedlings were cut above the hypocotyl-root boundary and roots were removed. A dot of *A. rhizogenes* (MSU440) containing *pNcZNT1::GUS* or *pAtZIP4::GUS* constructs was applied to the cut surface of each seedling and incubated for 5 day at 20°C/15 °C (day/night, 12 hours light). The *A. rhizogenes* inoculated seedlings were then transferred to ½ MS agar plates (no sugar, pH 5.8) containing 200 mg L⁻¹ tricarcillin (Ticarcillin Disodium Mixture 15:1 & potassium Clarulanate; Duchefa, Netherlands) at 24 °C (16/8 hr, light/darkness). The non-transformed roots, which did not express DsRed or GFP, were cut off once every week under Leica MZ FLIII Fluorescence Stereo Microscope until only transgenic roots were growing.

**Metal exposure**

To determine the *NcZNT1* expression in response to various Zn treatments, seeds of *N. caerulescens* (La Calamine) were grown in modified half strength Hoagland’s nutrient solution (Schat *et al.*, 1996) containing 10 µM ZnSO₄. After three weeks, the seedlings were supplied with different Zn concentrations, Zn deficiency (0.05 µM ZnSO₄), Zn supply (2 µM ZnSO₄), Zn sufficient (10 µM ZnSO₄), or excess Zn (1000 µM ZnSO₄). After another four weeks, shoots and roots were collected separately for gene expression analysis.

To determine the metal tolerance and accumulation of transgenic *p35S::NcZNT1* *A. thaliana* lines, nine plants for each of three independent transgenic lines and one control *A. thaliana* wild type (WT) line were grown hydroponically in modified half strength Hoagland’s nutrient solution containing sufficient Zn (2 µM ZnSO₄) and excess Zn (60 µM ZnSO₄). For each treatment, the transgenic and control lines were grown in the same tray to avoid any effect of variation among the trays. Each tray was containing about nine litres of hydroponic
medium. The plants were grown in a climate chamber (20/15°C day/night temperatures; 250 μmoles light m⁻² s⁻¹ at plant level during 12 h/day; 75% RH). Plants were grown for five weeks for the flowering time and Zn deficiency analysis experiments, while for the excess Zn experiment, plants were grown for four weeks. For the first two weeks, plants were grown in sufficient Zn and for the rest of the period in respective treatments. The nutrient medium was refreshed twice every week. Root and shoot tissues were harvested for metal concentration analysis. Each hydroponics experiment was repeated twice at different time points, while keeping all growth conditions the same.

To determine the response of p35S::NcZNT1 transformed A. thaliana plants to Cd, the same transgenic lines were grown hydroponically on modified half strength Hoagland’s solution with sufficient Zn (2 μM ZnSO₄) for two weeks and then transferred to the same media but containing sufficient Zn (0 μM CdSO₄-2 μM ZnSO₄) and/or excess Cd (2 μM CdSO₄-2 μM ZnSO₄) while keeping the rest of the minerals constant in the media. The nutrient solution was refreshed every week. Plants were grown for four weeks. For mineral concentration analysis the root and shoot tissues were harvested individually.

**Promoter GUS Arabidopsis**

For qualitative GUS analysis in pAtZIP4::GUS and pNcZNT1::GUS expressing A. thaliana, all homozygous lines were grown hydroponically on modified half strength Hoagland’s nutrient solution, which supplied sufficient Zn to plant (2 μM ZnSO₄) and on Zn-deficiency medium (no Zn added). Three of these lines for each construct were selected for quantitative GUS assay, based on the GUS expression results obtained with the qualitative GUS assay. These three lines (six seedlings per line) were grown on vertical half MS and 1% agar plates for two weeks and then grown hydroponically on modified half strength Hoagland’s solution with sufficient Zn (2 μM ZnSO₄) and Zn deficient (no Zn added). Roots were collected for quantitative GUS assay every week. The same was done for all other transformed A. thaliana plants with F05 through F26 constructs, except that roots were collected for quantitative GUS assay after three weeks only.

**Promoter GUS Noccaea**

To compare NcZNT1 and ArZIP4 promoter expression in response to Zn treatment in N. caerulescens roots, pNcZNT1::GUS and pAtZIP4::GUS expressing N. caerulescens were grown hydroponically on modified half strength Hoagland’s
solution with deficient Zn (0.05 µM ZnSO$_4$) and sufficient Zn (10 µM ZnSO$_4$). The nutrient solution was refreshed every week. Plants were grown for four weeks.

**Promoter GFP**

To observe the root tissues localization of *NcZNT1* and *pAtZIP4* promoters in *A. thaliana*, the *pNcZNT1::eGFP* and *pAtZIP4::eGFP* expressing seedlings were grown in modified half Hoagland’s solution with Zn deficiency (0 µM ZnSO$_4$) or with Zn supply (25 µM ZnSO$_4$) for three weeks. Similarly, *N. caerulescens* plants expressing *pNcZNT1::eGFP* and *pAtZIP4::eGFP* in roots were transferred to half strength Hoagland’s solution with Zn deficiency (0.05 µM ZnSO$_4$) or Zn supply (100 µM ZnSO$_4$) for one week.

**RNA isolation and quantitative Reverse Transcriptase-PCR (qRT-PCR)**

To determine *NcZNT1* expression in *N. caerulescens* under different Zn exposure conditions and the expression of known metal transporters in *p35S::NcZNT1* expressing *A. thaliana* lines exposed to sufficient Zn (2 µM ZnSO$_4$), excess Zn (60 µM) and excess Cd (2 µM) treatment, qRT-PCR was carried out. Total RNA was extracted by RNeasy® Plant Mini kit (Qiagen). On-column DNase digestion was performed to eliminate genomic DNA contamination. RNA concentration and quality were measured by a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific). All RNA samples had A260/230 ratios of over 2.1, and a A260/A280 ratio of 2.1, so that the RNA quality was good enough for further qRT-PCR analysis. The first strand cDNA was synthesized from 1 µg RNA by using the iScript™ cDNA Synthesis Kit (Bio-Rad). Clathrine (Gendre et al., 2006) was used as reference gene for the normalization of *NcZNT1* expression in *N. caerulescens* and *AtUBP6* (At1g51710) for the normalization of *AtBHLH100, AtIRT1, AtIRT2, AtFRO2, AtHMA4, AtHMA3, AtNRAMP3, AtYSL3, AtFRD3, AtMTP1* and *NcZNT1* in *p35S::NcZNT1* expressing *A. thaliana* lines. Primers used for this gene expression analyses are shown in Table 2. Samples to which no reverse-transcriptase enzyme was added (NRT) were used as control to ensure the absence of genomic DNA in every sample.
Table 2. Primers used for q-PCR analyses.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Primer Sequence (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward primer for NcZNT1</td>
<td>GATCTTCGTCGATGTCTTTTGG</td>
</tr>
<tr>
<td>Reverse primer for NcZNT1</td>
<td>TGAGAGGTATGGCTACACCACGAGC</td>
</tr>
<tr>
<td>Forward primer for Clathrin</td>
<td>AGCATACACTGCGTGCAAAG</td>
</tr>
<tr>
<td>Reverse primer for Clathrin</td>
<td>TCGCCTGTGTCACATATCTC</td>
</tr>
<tr>
<td>Forward primer for AtUBP6</td>
<td>GAAAGTGGATTACCGGCTG</td>
</tr>
<tr>
<td>Reverse primer for AtUBP6</td>
<td>CTCTAAGTCTCGCGGAGGAG</td>
</tr>
<tr>
<td>Forward primer for AtIRT1</td>
<td>AAGCTTTGATCAGGTTTG</td>
</tr>
<tr>
<td>Reverse primer for AtIRT1</td>
<td>TTAGGTCCTGACTAGTCCG</td>
</tr>
<tr>
<td>Forward primer for AtIRT2</td>
<td>ATGGCTACTACAGCTGGTC</td>
</tr>
<tr>
<td>Reverse primer for AtIRT2</td>
<td>CTAGACCGGACATCATAGCG</td>
</tr>
<tr>
<td>Forward primer for AtFRO2</td>
<td>CTTGCTCATTCCGTCGAG</td>
</tr>
<tr>
<td>Reverse primer for AtFRO2</td>
<td>AAGATGTTGGAGATGGACCG</td>
</tr>
<tr>
<td>Forward primer for AtBHLH100</td>
<td>AAGTCAAGAGGAGGGTTACA</td>
</tr>
<tr>
<td>Reverse primer for AtBHLH100</td>
<td>GATGCATAAGTAAAGAGTCGCT</td>
</tr>
<tr>
<td>Forward primer for AtFRD3</td>
<td>CGAGTTGCATCTCCTTCTTCT</td>
</tr>
<tr>
<td>Reverse primer for AtFRD3</td>
<td>TGATAACGGTCTCTCGAACA</td>
</tr>
<tr>
<td>Forward primer for AtMTP1</td>
<td>ACGGCCATTACACACGACAA</td>
</tr>
<tr>
<td>Reverse primer for AtMTP1</td>
<td>TGCTTGTCTCATTCAAGCA</td>
</tr>
<tr>
<td>Forward primer for AtYSL3</td>
<td>GAATTGAGAGACTAGTTATTC</td>
</tr>
<tr>
<td>Reverse primer for AtYSL3</td>
<td>CGAGTTTTTACTTTTTGTAGCG</td>
</tr>
<tr>
<td>Forward primer for AtNRAMP3</td>
<td>ACAATGGGAGTCATTTACGC</td>
</tr>
<tr>
<td>Reverse primer for AtNRAMP3</td>
<td>ATGCAACCACAACAGCTCAAC</td>
</tr>
<tr>
<td>Forward primer for AtHMA4</td>
<td>ATGCGGTATACAAAAACAAAG</td>
</tr>
<tr>
<td>Reverse primer for AtHMA4</td>
<td>GAGATTTGTTTACTGCTCTGAGC</td>
</tr>
<tr>
<td>Forward primer for AtHMA3</td>
<td>TTAAGCTGGAGAAAGTTACCGA</td>
</tr>
<tr>
<td>Reverse primer for AtHMA3</td>
<td>GCTAGAGCTGATTTTTCACCT</td>
</tr>
</tbody>
</table>

qRT-PCR was performed by using the kit of iQTMSYBR® Green Supermix (Bio-Rad), including 12.5 µL of iQ SYBR Green Supermix, 5 pmol of forward and reverse primers, and 5 µL of 10 times diluted cDNAs (corresponding to 5 ng/µL RNA) in a total volume of 25 µL. The qRT-PCR conditions were 3 min
at 95°C, followed by 40 cycles of 10 sec at 95°C and 1 min at 62°C. The fluorescence signal was detected by a CFX96TM Real-Time Detection System (Bio-Rad). Melting curves were analysed to confirm the absence of primer dimers and nonspecific products. Three biological repeats per genotype or treatment and two technical repeats per biological repeat were used for the qRT-PCR analysis. The difference between technical repeats was less than 0.2 cycles. Relative transcript levels (RLT) were calculated by the \(2^{-\Delta\Delta Ct}\) method (Livak & Schmittgen, 2001). NcZNT1 expression of leaves under excess Zn (1000 µM ZnSO₄) was used as the calibrator in N. caerulescens plants, which means its RLT value is 1. In case of p35S::NcZNT1 expressing A. thaliana and WT lines, expression of NcZNT1, AtIRT1, AtIRT2, AtFRO2, AtBHLH100, AFRD3, AtMTP1, AtYSL3, AtNRAMP3, AtHMA4, and AtHMA3 genes in transgenic and WT lines was normalized to the respective WT shoot under sufficient Zn condition (2 µM Zn). NcZNT1 expression in the p35S::NcZNT1 expressing A. thaliana line was normalized to its AtHMA4 shoot expression grown in sufficient Zn. All other gene transcripts were normalized to their respective WT shoot transcripts under sufficient Zn exposure. Each point is the average of two technical repeats of three biological samples (each biological sample was again a pool of 3 biological samples). The stability of housekeeping genes was calculated by geNorm in qBasePLUS software (Biogazelle) and the reference genes with a geNorm M value lower than 0.5 were taken as stable genes (Hellemans et al., 2007).

The qPCR analysis of various genes in p35S::NcZNT1 and wild-type lines was depicted as heat map by using online BAR HeatMapper Plus Tool (http://bbc.botany.utoronto.ca/ntools/cgi-bin/ntools_heatmapper_plus.cgi).

**Root and shoot metal accumulation assay**

Shoot and root samples were collected at the end of the metal exposure experiments and minerals like Zn, Cd, Fe, and Mn were measured spectrophotometrically as described by Assunção et al. (2003b).

**GUS staining assay**

For GUS staining, the tissues were cut off and incubated at 37 °C for 3 hours in GUS staining solutions (pH 7.4) having 50 mM sodium phosphate, and 1 mg/ml 5-bromo-4-chloro-3-indolyl b-D-glucuronide (X-Gluc). The stained tissues were washed thrice with 70% ethanol. A Nikon Eclipse 80i microscope was used to
visualize the GUS expression in tissues and images were captured by NIS Elements D3.1 software.

For quantitative kinetic GUS analysis, assays were performed as described by Nap et al. (1992).

**GFP visualization**

Transgenic roots identified by a Leica MZ FLIII Fluorescence Stereo Microscope were either or not immersed in 1 µg/mL propidium iodide for 1-5 minutes, and then washed with deionized water before imaging. Propidium iodide was used to stain red and identify the cell walls. Images were acquired with an inverted laser scanning confocal microscope (LSCM) system, Zeiss LSM 510 Meta (Carl Zeiss, Jena, Germany) or Zeiss LSM 5 PASCAL. The eGFP (green) signal was visualized with an excitation wave length set at 488 nm and assembling emission signals between 505 to 530 nm. The signal for plant cell wall was visualized with excitation wave length set at 543 nm and assembling emission signals at 560 nm. A ×63 Plan Apochromate/ 1.4 oil DIC objective was used for the observation of *A. thaliana* transgenic roots, and EC Plan-Neofluar 20x or LD Plan-Neofluar 40x objectives were used for *N. caerulescens* transgenic roots. Digital images were processed using LSM 510 3.5 or LSM 5 Image Examiner software.

**Confirmation of the 5’ region of NcZNT1 in three N. caerulescens accessions**

To verify the presence of an ATG start codon in *NcZNT1* cDNA 5’ to the one indicated by Milner et al. (2012), in *N. caerulescens* accessions, a forward primer (5’-GCTTTCTGCTCCTTGATCC-3’) and reverse primer (5’-CGATGAGAGGTATGGCTACA-3’) were designed according to the *NcZNT1* cDNA sequence of La Calamine (Assunção et al., 2001) and q-PCR was performed for *NcZNT1* in accessions La Calamine (LC), Prayon (PY), and Ganges (GA). The forward primer was designed at the corresponding 5’-UTR region of the orthologous *A. thaliana* ZIP4 gene. RNA of LC, PY, and GA were isolated using the RNeasy® Plant Mini kit (Qiagen), and cDNA was synthesized by using the M-MLV Reverse Transcriptase (Invitrogen). The amplification of *NcZNT1* cDNA fragments was performed by Pfu DNA polymerase (Fermentas). The amplified fragments were then cloned into the pGEMT-easy vector (Promega) for sequencing. The whole DNA sequence of *NcZNT1* was also amplified by using the same forward primer (5’-GCTTTCTGCTCCTTGATCC-3’), but different reverse primer,
(5’-CTAAGCCCAAATGGCGA-3’) designed at the 3’ end of the NcZNT1 coding sequence. DNA of LC, PY, and GA were extracted by a modified nuclear extraction protocol (Aarts et al., 2000). The amplified fragments were also cloned into the pGEMT-easy vector (Promega) for sequencing. cDNA, DNA and predicted protein sequence results were compared by using MultAlin software (Multiple sequence alignment by Florence Corpet http://multalin.toulouse.inra.fr/multalin).

**Sequence comparison among the ZIP4 promoters from A. thaliana, A. lyrata, A. halleri, C. pyrenaica and N. caerulescens**

In order to look for conserved sequences, that can confirm the importance of identified cis elements the ZIP4 promoters from three different species, C. pyrenaica, A. halleri and A. lyrata were isolated. The same primers used for amplification of the NcZNT1 promoter were used for PCR amplification on the genomic DNA of these plants. All these PCR amplifications gave single bands of different sizes from around 600 to 1200 bp, except for A. halleri. For A. halleri, two bands were obtained by PCR amplification for each primer pairs: between (a) 500 and 750 bp (primer pairs P1 and P3; Table 1) (b) between 750 and 1000 bp (primer pair P2 and P3; Table 1). The PCR amplification from C. pyrenaica resulted in single bands around 600 bp for both the primer sets. The PCR amplification from A. lyrata resulted in single bands around 700 bp (primer pairs P1 and P3) and 1200 bp (primer pair P2 and P3). All these fragments were cloned into the pGEMT-easy vector and sequenced. The sequences of these fragments from each species were overlapping and the larger fragments were used for further analysis.

**Statistical analyses**

Where needed, data were analysed for significance at p<0.05 by using Student’s t-test, ANOVA and two-way ANOVA (Least Significance Difference Test) in the SPSS v. 12 software package for MS Windows.
RESULTS

*NcZNT1* expression is up-regulated under Zn deficiency

Previously, van de Mortel *et al.* (2006) showed that the *NcZNT1* gene is expressed constitutively in roots of *N. caerulescens*, almost irrespective of Zn supply status, in contrast to its *A. thaliana* orthologue, *AtZIP4*, which expression is strongly up-regulated in roots under Zn deficiency. To quantify *NcZNT1* expression in response to different Zn status in roots and shoots of *N. caerulescens*, we performed a quantitative RT-PCR using material from plants exposed to four different Zn concentrations. Compared to Zn sufficient conditions (Zn2 and Zn10), *NcZNT1* expression was significantly induced under Zn deficient conditions (Zn0.05), especially in shoots (8.8 fold), but also in roots (up to 2 fold); while *NcZNT1* expression was repressed by excess Zn treatment (Zn1000) (Fig. 1A). There was no significant difference in expression when comparing both Zn sufficient conditions.

Fig 1. Quantitative reverse transcriptase PCR analysis of *NcZNT1* expression in response to Zn in *N. caerulescens* and in *p35S::NcZNT1* expressing *A. thaliana* (A) *N. caerulescens* plants were grown in ½ Hoagland’s nutrient solution with different Zn conditions i.e. Zn deficiency (0.05 µM ZnSO4), Zn supply (2 µM ZnSO4), Zn sufficient (10 µM ZnSO4) and Zn excess (1000 µM ZnSO4) for four weeks. * in (A) indicates a significant difference (p< 0.05, Student’s t test,) relative to Zn sufficient condition in shoots or roots. (mean ± SE of four replicates). (B) *NcZNT1* expression in *p35S::NcZNT1* expressing *A. thaliana* grown in sufficient Zn (2 µM ZnSO4) conditions. The gene expression was normalized to the *AtUBP6* housekeeping gene in *p35S::NcZNT1* line and wild-type (WT) line. * in (B) indicates a significant difference (p< 0.05, Student’s t test) between transgenic and respective WT line (mean ± SE of four replicates).
NcZNT1 localized to the plant plasma membrane

Based on the NcZNT1 cDNA we cloned previously (Assunção et al., 2001), the predicted protein it encodes (GenBank acc. no. AAK69429.1) consists of 408 amino acids. ZIP proteins are thought to be mainly functioning as plasma membrane metal transporters. Previously, Pence et al. (2000) showed that the NcZNT1 protein can transport Zn and Cd into yeast (Saccharomyces cerevisiae) and thus we expected it to localize to the plasma membrane. Transient expression of pNcZNT1::eGFP construct in cowpea protoplasts showed that the GFP fluorescence was indeed localized to the plasma membrane, even though the strong CaMV35S promoter caused improperly targeted GFP to build up in the cytoplasm (Fig 2). So indeed, NcZNT1 localizes to the plasma membrane, which is in line with a role of Zn uptake into the cell.

Fig 2. A p35S::NcZNT1-GFP chimeric fusion protein transiently expressed in cowpea protoplasts localizes to the plasma membrane (green arrow). Although there is some cytoplasmic fluorescence, probably due to the very high expression caused by the CaMV 35S promoter, the GFP signal was clearly not seen in the vacuolar or chloroplast membranes.
Heterologous expression of NcZNT1 confers early flowering and increased Zn tolerance and accumulation in A. thaliana

NcZNT1 was found to be expressed at higher levels in N. caerulescens, when compared to its orthologues from A. thaliana or T. arvense (Pence et al., 2000; Assunção et al., 2001; Hammond et al., 2006; van de Mortel et al., 2006), corresponding with high Zn uptake, suggesting the involvement of this gene in high Zn uptake from the growth medium. In order to investigate if high expression of the gene as such would be sufficient to increase Zn uptake in plants, we expressed the NcZNT1 gene to high levels in A. thaliana under control of the CaMV 35S promoter (Fig. 1B). Homozygous p35S::NcZNT1 A. thaliana lines were grown on modified half Hoagland’s nutrient solution containing sufficient Zn (2 μM ZnSO₄). All independent transgenic lines flowered two to four days earlier than wild-type Col-0 (WT) plants (Fig. 3). No additional abnormal visible phenotype was discerned at this normal Zn treatment. When the same lines were grown on Zn deficient (0 μM ZnSO₄) medium, they were more sensitive to zinc deficiency than the WT plants (Fig 4A). The experiment was repeated to determine the zinc concentration in the plants, but this time the plants were grown on Zn

---

**Fig 3.** Flowering time of transgenic p35S::NcZNT1 and Col wild-type (Col-WT) A. thaliana lines grown on no Zn (0 μM ZnSO₄). Three independently transformed lines (NcZNT1-1, NcZNT1-2, NcZNT1-3) and Col-WT line were grown hydroponically on half Hoagland’s media and the picture was taken during the fifth week. (A) Visible phenotype of WT plants grown in no Zn (B) Visible phenotype of p35S::NcZNT1 plants grown in no Zn. (C) Flowering time in days of the three independent transformed of p35S::NcZNT1 lines compared to Col-WT. * indicates that the difference between the value for p35S::NcZNT1 is significantly different from Col-WT (p<0.05, Two way ANOVA) (mean ± SE of 4 replicates).
deficient medium with a very low zinc concentration (0.05 µM ZnSO₄) to allow some Zn uptake. The Zn concentration in both shoots and roots of the p35S::NcZNT1 plants was significantly higher than in the WT plants (Fig. 4B). The same was the case when plants were grown at sufficient Zn supply (Fig 7A, B). No abnormal visible phenotype was discerned at the sufficient Zn treatment in transgenic lines supply (Fig 5A). At sufficient Zn supply, the transgenic lines showed reduced dry shoot and root weight compared to the WT line, although morphologically there was no difference (Fig 6A, B). The reverse was seen when plants were exposed to excess Zn. This high Zn supply affected growth of the WT line much more than that of the transgenic lines, also leading to significantly higher shoot and root dry weights in those lines (Fig. 5B; Fig 6A, B). The WT line also displayed purple anthocyanin pigmentation of the older leaves while transgenic plants did not show this and generally appeared to be more green. The p35S::NcZNT1 lines grown on excess Zn had a markedly higher Zn concentration in shoots and roots compared to the WT (Fig. 7A, B). Transgenic lines also showed enhanced Mn, but not Fe, in shoots while there was no difference in Fe and Mn concentration in roots of transgenic and WT lines (Fig. 8, A-D).

**Fig 4.** Phenotypic and Zn concentration analysis of transgenic p35S::NcZNT1 and Col wild-type (Col-WT) A. thaliana lines grown in Zn deficient (0.05 µM ZnSO₄) media. Three independently transformed lines (NcZNT1-1, NcZNT1-2, NcZNT1-3) and wild-type (WT) line were grown hydroponically in half Hoagland’s media supplemented with 0.05 µM ZnSO₄ for four weeks. (A) Visible phenotype of p35S::NcZNT1 and Col-WT plants. (B) Zn concentration in shoot and root (µmoles/g DW). * indicates that the difference between the value for p35S::NcZNT1 is significantly different from Col WT in respective tissue (p<0.05, Two way ANOVA) (mean ± SE of 4 replica).
Fig 5. The phenotypic response of transgenic p35S::NcZNT1 and wild-type (WT) A. thaliana lines to sufficient Zn (2 μM ZnSO₄), excess Zn (60 μM ZnSO₄) and excess Cd (2 μM CdSO₄ and 2 μM ZnSO₄). Three independently transformed lines (NcZNT1-1, NcZNT1-2, NcZNT1-3) and Col-WT line were grown hydroponically on half Hoagland’s media supplemented with respective metal treatment for four weeks. (A) Visible phenotype of p35S::NcZNT1 and Col-WT plants grown on sufficient Zn medium (B) Visible phenotype of p35S::NcZNT1 and Col-WT plants grown on excess Zn medium (C) Visible phenotype of p35S::NcZNT1 and Col-WT plants grown on excess Cd medium.
Fig 6. The dry biomass of transgenic p35S::NcZNTI and Col wild-type (Col-WT) A. thaliana lines to sufficient Zn (2 µM ZnSO₄), excess Zn (60 µM ZnSO₄) and excess Cd (2 µM CdSO₄). Three independently transformed lines (NcZNT1-1, NcZNT1-2, NcZNT1-3) and Col-WT line were grown hydroponically on half Hoagland’s media supplemented with respective metal treatment for four weeks (A) Dry shoot weight (g) (B) Dry root weight (g).
Fig 7. Zn and Cd concentration analysis of transgenic p35S::NcZNT1 and Col wild-type (Col-WT) A. thaliana lines grown in sufficient Zn (2 µM ZnSO₄), excess Zn (60 µM ZnSO₄) and excess Cd (2 µM CdSO₄). Three independently transformed lines (NcZNT1-1, NcZNT1-2, NcZNT1-3) and Col-WT line were grown hydroponically on half Hoagland’s media supplemented with respective metal treatment for four weeks. (A) Zn concentration in shoots (µmoles/g DW) (B) and in roots (µmoles/g DW) (C) shoot and root Cd concentration (µmoles/g DW). * indicates that the difference between the value for p35S::NcZNT1 is significantly different from Col-WT in respective metal treatment (p<0.05, Two way ANOVA) (mean ± SE of 4 replicates).
**NcZNT1** over-expressing *A. thaliana* lines showed enhanced Cd tolerance and accumulation

Since NcZNT1 was reported to be able to transport Cd, in addition to Zn (Pence *et al.*, 2000), we also determined the response of the p35S::NcZNT1 plants to excess Cd exposure. The same three transformed lines and the WT line were grown hydroponically on modified half Hoagland’s solutions containing sufficient Zn (2 μM ZnSO4) or excess Cd (2 μM CdSO4 + 2 μM ZnSO4). Like for excess Zn exposure, the transgenic lines were more tolerant to Cd exposure than WT, with larger rosette size and leaves with less chlorosis (Fig. 5C). Also the shoot and root dry weights of the transgenic lines were significantly higher than those of the WT line (Fig. 6A, B). Cd exposure also increased the shoot Zn and Cd but same Zn concentrations in roots of the transgenic lines. However, the shoot Cd concentrations remained much lower than the root Cd concentrations (Fig. 7C). Shoot Fe concentrations decreased in the transgenic lines upon Cd exposure while the shoot Mn concentrations of the transgenic lines were same as for WT (Fig. 8A,C) There was no difference in the Fe and Mn concentration in roots of the transgenic and WT lines (Fig B, D).

![Graphs showing iron and manganese concentrations](image)

**Fig 8.** Fe and Mn concentration analysis of transgenic p35S::NcZNT1 and Col wild-type (Col-WT) *A. thaliana* lines grown in sufficient Zn (2 μM ZnSO4), excess Zn (60 μM ZnSO4) and excess Cd (2 μM CdSO4). Three independently transformed lines (NcZNT1-1, NcZNT1-2, NcZNT1-3) and Col-WT line were grown hydroponically on half Hoagland’s media supplemented with respective metal treatment for four weeks. (A) Fe concentration in shoots (μmoles/g DW) (B) and in roots (μmoles/g DW) (C) Mn concentration in shoots (μmoles/g DW) (D) and in roots (μmoles/g DW). * indicates that the difference between the value for p35S::NcZNT1 is significantly different from Col-WT in respective metal treatment (p<0.05, Two way ANOVA) (mean ± SE of 4 replicates).
To determine if there is a competition effect of Cd with Zn, Fe and Mn, one p35S::NcZNT1 line was compared with the WT line on half Hoagland’s media supplemented with either 2 μM ZnSO₄ (2 Zn 0 Cd), 5 μM CdSO₄ + 2 μM ZnSO₄ (2 Zn 5 Cd) or 5 μM CdSO₄ - 0 μM ZnSO₄ (0 Zn 5 Cd), keeping the other minerals constant in the media. In all given treatments, due to the much higher Cd exposure, the p35S::NcZNT1 line showed no visible phenotypic difference compared to the WT, with both lines being equally affected, as displayed by severe chlorosis and stunted growth (data not shown). Nevertheless, the transgenic plants contained a significantly higher shoot Cd concentration compared to WT (Fig. 9A). Like in the previous experiments, shoot Zn concentrations were higher in transgenic than WT plants, except for plants grown in the absence of Zn. There is competition between Zn and Cd, as the Zn concentration in shoots of transgenic plants grown on 2 Zn 5 Cd was lower than in plants grown on 2 Zn 0 Cd. This acts through NcZNT1, as WT plants showed similar Zn concentrations in both treatments. In the absence of Zn, both shoot and root Cd concentrations reach high levels, much higher than when Zn is present (Fig. 9A, B). The Fe contents in the roots and Mn content in shoots of the transgenic plants in Cd exposed wild type and transgenic plants were significantly higher than in the non-exposed ones (Fig. 9C, D).

**Fig 9.** Zn, Cd, Fe and Mn concentration analysis of transgenic p35S::NcZNT1 and Col wild-type (Col-WT) A. thaliana lines grown on 2 μM Zn, 5 μM Cd- 2 μM Zn and 5 μM Cd-0 μM Zn. Three independently transformed lines (NcZNT1-1, NcZNT1-2, NcZNT1-3) and Col-WT line were grown hydroponically on half Hoagland’s media supplemented with respective metal treatment for four weeks. (A) Zn and Cd concentration in shoot (μmoles/g DW) (B) and in roots (C) Fe and Mn concentration in shoot (μmoles/g DW) (D) and in roots. * indicates that the difference between the value for p35S::NcZNT1 is significantly different from Col-WT in respective metal treatment (p<0.05, Two way ANOVA) (mean ± SE of 4 replicates).
Expression of NcZNT1 alters the expression of other metal homeostasis genes in A. thaliana

Considering that expression of NcZNT1 alters the A. thaliana Zn and Cd accumulation and tolerance, we determined the expression of Zn and Fe homeostasis genes AtBHLH100, AtIRT1, AtIRT2, AtFRO2 (involved in Fe uptake), AtNRAMP3 (involved in Fe remobilization), AtHMA4 (involved in Zn/Cd translocation), AtYSL3, AFRD3 (involved in Zn/Fe translocation) and AtMTP1 (involved in Zn excess tolerance) upon exposure of WT and transgenic plants to sufficient Zn (2 µM ZnSO₄), excess Zn (60 µM ZnSO₄) and excess Cd (2 µM CdSO₄ + 2 µM ZnSO₄).

Of the genes involved in Fe uptake, AtIRT1, AtIRT2 and AtFRO2 are mainly expressed in roots, whereas AtbHLH100 is expressed both in roots and shoots. Transcription levels of all of these genes go up upon excess Zn and excess Cd exposure in roots of A. thaliana and often also in shoots, even if expression levels are low (Fig. 10; Table 4). This supports the idea that A. thaliana will induce a Fe deficiency response upon excess Zn and Cd exposure (van de Mortel et al., 2006). The p35S::NcZNT1 line exhibited reduced AtbHLH100 expression in shoot and root tissues under excess Zn treatment while showing significantly higher expression upon excess Cd treatment. Expression of AtIRT1, AtIRT2 and AtFRO2 in roots of transgenic plants was significantly reduced upon excess Zn treatment. Excess Cd had this effect only on AtIRT1 expression, while it increased expression of AtIRT2. Similar effects were seen for gene expression in shoot, though the biological relevance of this appears little considering the low expression of these genes in shoots. Also the expression of AtFRD3 was significantly decreased in roots of transgenic compared to WT plants under sufficient and excess Zn but not excess Cd (Fig. 10; Table 4). AtMTP1 is expressed in both roots and shoots, with higher expression in both tissues of the transgenic line than WT upon excess Zn and Cd, while at sufficient Zn, only the expression in shoots is reduced compared to the WT line (Fig. 10; Table 4). Both AtYSL3 and AtNRAMP3 show only increased expression in roots of transgenic plants upon excess Zn, when compared to WT plants, while expression upon other treatments and tissues was similar in both genotypes (Fig. 10; Table 4). AtHMA4 expression levels were found to be higher in root tissues of both lines compared to the shoot tissues (Fig. 10; Table 4). The transgenic line showed increased AtHMA4 expression in roots upon sufficient Zn supply, but decreased expression in roots upon excess Zn supply compared to the WT line. In general, HMA3 was higher expressed in roots than in the shoot and
higher in the shoot and root tissues of the WT line in the excess Zn condition (Fig. 10; Table 4). The transgenic line had higher HMA3 expression in roots under sufficient Zn supply.

Fig 10. Relative gene expression analysis of a known metal transcription factor gene, AtBHLH100, and transporter genes AtIRT1, AtIRT2, AtFRO2, AtHMA4, AtHMA3, AtFRD3, AtMTP1, AtNRAMP3 and AtYSL3 in shoot and root of p35S::NcZNT1 and Col wild-type (Col-WT) in response to sufficient Zn (2 µM ZnSO₄), excess Zn (60 µM ZnSO₄) and excess Cd (2 µM CdSO₄) after two weeks of metal exposure. Expression is relative to the shoot in respective Col-WT in sufficient Zn treatment, used as RTL 1. AtUBP6 (At1g51710) was used as housekeeping gene to normalize the data.
Table 4. Relative gene expression analysis of a known metal transcription factor gene, *AtBHLH100*, and transporter genes *AtIRT1*, *AtIRT2*, *AtFRO2*, *AtHMA4*, *AtHMA3*, *AtFRD3*, *AtMTP1*, *AtNRAMP3* and *AtYSL3* in shoot and root of *p35S::NcZNT1* and Col wild-type (WT) in response to sufficient Zn (2 µM ZnSO$_4$), excess Zn (60 µM ZnSO$_4$) and excess Cd (2 µM CdSO$_4$) after two weeks of metal exposure. Expression is relative to the shoot in respective wild-type in sufficient Zn treatment, used as RTL 1. *AtUBP6* (At1g51710) was used as housekeeping gene to normalize the data. ZNT1 represents *p35S::NcZNT1* line, WT is for wild-type and “suff” represents sufficient. Different letters indicate the significant difference in gene expression of the respective gene between lines grown in given treatments (p<0.05, ANOVA, Least Significant Difference Test) (mean ± SE of 4 replica).

<table>
<thead>
<tr>
<th>Gene</th>
<th>WT Suf Zn Shoot</th>
<th>ZNT1 Suf Zn Shoot</th>
<th>WT Excess Zn Shoot</th>
<th>ZNT1 Excess Zn Shoot</th>
<th>WT Excess Cd Shoot</th>
<th>ZNT1 Excess Cd Shoot</th>
<th>WT Suf Zn Root</th>
<th>ZNT1 Suf Zn Root</th>
<th>WT Excess Zn Root</th>
<th>ZNT1 Excess Zn Root</th>
<th>WT Excess Cd Root</th>
<th>ZNT1 Excess Cd Root</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>AtBHLH100</em></td>
<td>a±0.1</td>
<td>b±0.3</td>
<td>c±0.8</td>
<td>d±1.3</td>
<td>e±3.6</td>
<td>f±5.1</td>
<td>b±0.7</td>
<td>c±0.2</td>
<td>d±1.3</td>
<td>e±3.6</td>
<td>f±5.1</td>
<td>g±1.2</td>
</tr>
<tr>
<td><em>AtIRT1</em></td>
<td>ab±0.2</td>
<td>a±0.3</td>
<td>b±0.5</td>
<td>c±0.6</td>
<td>d±0.1</td>
<td>e±0.2</td>
<td>f±0.1</td>
<td>g±0.2</td>
<td>h±0.1</td>
<td>i±0.2</td>
<td>j±0.1</td>
<td>k±0.1</td>
</tr>
<tr>
<td><em>AtIRT2</em></td>
<td>a±0.2</td>
<td>b±0.3</td>
<td>c±0.4</td>
<td>d±0.5</td>
<td>e±0.6</td>
<td>f±0.7</td>
<td>g±0.8</td>
<td>h±0.9</td>
<td>i±0.1</td>
<td>j±0.2</td>
<td>k±0.1</td>
<td>l±0.1</td>
</tr>
<tr>
<td><em>AtFRO2</em></td>
<td>a±0.1</td>
<td>b±0.2</td>
<td>c±0.3</td>
<td>d±0.4</td>
<td>e±0.5</td>
<td>f±0.6</td>
<td>g±0.7</td>
<td>h±0.8</td>
<td>i±0.1</td>
<td>j±0.2</td>
<td>k±0.1</td>
<td>l±0.1</td>
</tr>
<tr>
<td><em>AtHMA4</em></td>
<td>a±0.1</td>
<td>b±0.2</td>
<td>c±0.3</td>
<td>d±0.4</td>
<td>e±0.5</td>
<td>f±0.6</td>
<td>g±0.7</td>
<td>h±0.8</td>
<td>i±0.1</td>
<td>j±0.2</td>
<td>k±0.1</td>
<td>l±0.1</td>
</tr>
<tr>
<td><em>AtHMA3</em></td>
<td>a±0.1</td>
<td>b±0.2</td>
<td>c±0.3</td>
<td>d±0.4</td>
<td>e±0.5</td>
<td>f±0.6</td>
<td>g±0.7</td>
<td>h±0.8</td>
<td>i±0.1</td>
<td>j±0.2</td>
<td>k±0.1</td>
<td>l±0.1</td>
</tr>
<tr>
<td><em>AtFRD3</em></td>
<td>a±0.1</td>
<td>b±0.2</td>
<td>c±0.3</td>
<td>d±0.4</td>
<td>e±0.5</td>
<td>f±0.6</td>
<td>g±0.7</td>
<td>h±0.8</td>
<td>i±0.1</td>
<td>j±0.2</td>
<td>k±0.1</td>
<td>l±0.1</td>
</tr>
<tr>
<td><em>AtMTP1</em></td>
<td>a±0.1</td>
<td>b±0.2</td>
<td>c±0.3</td>
<td>d±0.4</td>
<td>e±0.5</td>
<td>f±0.6</td>
<td>g±0.7</td>
<td>h±0.8</td>
<td>i±0.1</td>
<td>j±0.2</td>
<td>k±0.1</td>
<td>l±0.1</td>
</tr>
<tr>
<td><em>AtNRAMP3</em></td>
<td>a±0.1</td>
<td>b±0.2</td>
<td>c±0.3</td>
<td>d±0.4</td>
<td>e±0.5</td>
<td>f±0.6</td>
<td>g±0.7</td>
<td>h±0.8</td>
<td>i±0.1</td>
<td>j±0.2</td>
<td>k±0.1</td>
<td>l±0.1</td>
</tr>
<tr>
<td><em>AtYSL3</em></td>
<td>a±0.1</td>
<td>b±0.2</td>
<td>c±0.3</td>
<td>d±0.4</td>
<td>e±0.5</td>
<td>f±0.6</td>
<td>g±0.7</td>
<td>h±0.8</td>
<td>i±0.1</td>
<td>j±0.2</td>
<td>k±0.1</td>
<td>l±0.1</td>
</tr>
</tbody>
</table>
Fig 11. pAtZIP4::GUS expression in A. thaliana plant tissues under Zn deficiency. GUS staining in (A) a detached lateral root; (B) close-up of (A) showing endodermis and inner cortex staining; (C) root tip and root hair zone; (D) Zn deficient leaf; (E) close up of (D) showing trichomes; (F) young inflorescence (G) close up of (F) flower (H) siliques.

Fig 12. pNcZNT1::GUS expression in A. thaliana plant tissues under Zn deficiency. GUS staining in (A) a detached lateral root (B) close-up of (A) showing endodermis and inner cortex staining; (C) root tip and root hair zone; (D) leaf (E) flowers (F) siliques.
Comparison of \textit{AtZIP4} and \textit{NcZNT1} promoter GUS activity in \textit{A. thaliana} and \textit{N. caerulescens}

In order to analyse the differential regulation of \textit{AtZIP4} and \textit{NcZNT1}, we transformed promoter::GUS constructs for both genes into \textit{A. thaliana}. Plants from a homozygous, \textit{pAtZIP4::GUS}-transformed \textit{A. thaliana} line were subsequently grown under Zn deficiency and control conditions. GUS staining was observed only under Zn deficient conditions. The same experiment was also performed with a homozygous, \textit{pNcZNT1::GUS} transformed \textit{A. thaliana} line, which only showed GUS staining in plants subjected to Zn deficiency. GUS staining was examined in more detail with more plants of these transgenic lines throughout their development. The roots of Zn deficient transgenic plants showed strong staining both in the tap root and lateral roots, specifically in the endodermis (Fig. 11A-C). A similar expression pattern was observed in \textit{pNcZNT1::GUS} transformed \textit{A. thaliana} plants in Zn-deficient roots (Fig. 12A-C). For both transgenic lines of \textit{pAtZIP4::GUS} and \textit{pNcZNT1::GUS}, intense staining was observed at the root tip and in the region of maturation of the root, in the rest of the root. No staining was observed in the elongation zone of the root. All these observations were similar for all six independent transformants for both constructs that were used. When examining later stages of development, \textit{pAtZIP4::GUS} expression was also observed in Zn deficient leaves, with most intense staining at the leaf edges (Fig. 11D) and in the trichomes (Fig. 11E). Similar expression patterns in the Zn-deficient leaves were also observed in \textit{pNcZNT1::GUS} transformed plants (Fig. 12D).

The staining was only observed in the Zn-deficient plants, whereas the Zn-sufficient tissues did not give any staining, neither in \textit{pAtZIP4::GUS} nor in \textit{pNcZNT1::GUS} transformants. In case of \textit{pAtZIP4::GUS} expressing plants, the highest staining intensity was observed in the young buds attached to the main stem (Fig. 11F). Strong staining was also observed in the pistils of older buds, in all developmental stages, until just before opening. In the open flower (Fig. 11G), the flower base is stained intensely in addition to the anther filaments. After fertilization, the siliques are typically stained at the distal ends; most prominently in young siliques (Fig. 11H). Also the pedicles are stained. In case of \textit{pNcZNT1::GUS}, a very similar expression pattern was observed in the Zn deficient \textit{A. thaliana} plants in all independent transformants (Fig. 12E,F).
When expressed in *A. thaliana*, both promoters acted very similarly. However, endogenous *NcZNT1* expression is thought to be higher in *N. caerulescens* than the endogenous expression of *AtZIP4* in *A. thaliana*. To confirm that both promoters reflect the endogenous expression, we also transformed both GUS constructs into *N. caerulescens*. Since there is no efficient stable transformation system available for *N. caerulescens*, we used the *A. rhizogenes*-mediated root transformation method modified from Limpens *et al.* (2004). This results in chimeric plants, with a transgenic root system supporting a non-transgenic rosette. Both *AtZIP4* and *NcZNT1* promoters were up-regulated under Zn deficiency, but repressed by Zn supply. The expression of *pAtZIP4::GUS* in *N. caerulescens* roots under Zn deficiency was restricted to root cap and stele (Fig. 13A, C). After 3 hours of GUS staining, *pAtZIP4::GUS* transgenic roots grown in sufficient Zn did not show any GUS expression (Fig. 13B, D). However, if the roots were stained overnight, they had very low expression in the stele (data not shown). When *pAtZIP4::GUS* expressing *N. caerulescens* roots were transferred from Zn deficient to sufficient condition, reduced GUS activity was observed supporting the low expression of *pAtZIP4::GUS* under sufficient Zn (data not shown). Under Zn deficient condition, *pNcZNT1::GUS* was ubiquitously expressed in root tissues, including root tips, root hairs, epidermis, cortex, pericycle and xylem (Fig. 13E, G). The strongest expression was found in the pericycle. With a slight increase of Zn supply (to 10 µM ZnSO4), the expression of the *NcZNT1* promoter at the root tip did not change, but the GUS staining was weaker than under Zn deficiency (Fig. 13H). In contrast to root tips, the expression in mature roots was limited to the stele, particularly in the pericycle and the xylem (Fig. 13F). In general, the *NcZNT1* promoter exhibited stronger GUS staining activities in roots of *N. caerulescens* than the *AtZIP4* promoter whether grown either in Zn deficient or sufficient conditions. Thus, it looks like *N. caerulescens* has a transcription factor binding to the *NcZNT1* promoter which may not bind very well to the *AtZIP4* promoter or that *N. caerulescens* might have a transcription factor binding to the *NcZNT1* promoter which is not present in *A. thaliana*.
Fig 13. pAtZIP4::GUS and pNCZNT1::GUS expression in N. caerulescens roots. Differential expression of both AtZIP4 and NcZNT1 promoters in response to Zn is presented. Three-hour-GUS staining of transgenic pAtZIP4::GUS (A-D) and pNCZNT1::GUS (E-H) roots under Zn deficient (0.05 µM ZnSO₄) conditions (A,E,C,G) and Zn sufficient (10 µM ZnSO₄) conditions (B,F,D,H) condition is shown. Mature roots (A,B,E,F) and root tips (C,D,G,H) are displayed respectively.

Quantitative comparison of pAtZIP4::GUS and pNCZNT1::GUS in A. thaliana show highest expression in Zn deficient roots

To investigate the response of the AtZIP4 and NcZNT1 promoters to Zn in A. thaliana in a quantitative way, the pAtZIP4::GUS and pNCZNT1::GUS expressing A. thaliana plants were grown for eight weeks in half Hoagland solution containing a concentration series of Zn (0, 0.1, 0.2, 0.5, 1, 2 µM) and the GUS activity was measured quantitatively every week (wk) starting from the first week of exposure (Fig. 14). Both in pAtZIP4::GUS and pNCZNT1::GUS lines, the GUS activity in roots were induced by Zn deficiency. The GUS activity was found to increase gradually in wk1 and wk2 upon transfer to 0 µM Zn (Fig. 14A). The highest values were observed in the wk3, and from wk4 onwards the expression was gradually decreasing on 0 µM Zn. The GUS activity at wk3 was increased by 31-fold compared to 2 µM Zn plants. The plants grown on 0.1 µM Zn showed some GUS activity, although to a lesser extent than on 0 µM Zn. In plants grown on 0.2, 0.5, 1 and 2 µM Zn hardly any GUS activity could be detected. A similar expression pattern was observed for pNCZNT1::GUS transgenic plants (Fig. 14B).
Fig 14. Quantitative comparison of GUS expression in (A) pAtZIP4::GUS and (B) pNcZNT1::GUS in A. thaliana roots. A time-course GUS assay was performed at weekly intervals on the transgenic plants grown on half Hoagland’s media containing 0, 0.1, 0.2, 0.5, 1, 2 µM ZnSO₄.

GFP expression of AtZIP4 and NcZNT1 promoters in A. thaliana and N. caerulescens

Although the GUS assay is a quick and convenient method to detect the localization of the promoter, the diffusion of the GUS protein when GUS expression is high, is still a limiting factor for precise localization analyses. Therefore, we used GFP visualization for the analysis of the NcZNT1 and AtZIP4 promoters, both in A. thaliana and in N. caerulescens. pAtZIP4::eGFP was expressed in the pericycle and cortex in A. thaliana roots (Fig. 15A-C); whereas, it was expressed throughout the N. caerulescens roots (Fig. 15D-F). pNcZNT1::eGFP expression was not changed in response to Zn excess in A. thaliana and it was localized in the pericycle (Fig. 15G-N). However, its localization in N. caerulescens was in the xylem and pericycle cells (Fig. 15O-W).
**Fig 15.** *pAtZIP4::eGFP* and *pNcZNT1::eGFP* expression in *A. thaliana* and *N. caerulescens* roots under Zn deficient (-Zn) or sufficient (+Zn) conditions. *A. thaliana* transformed with *pAtZIP4::GFP* (A-C) and *pNcZNT1::GFP* (G-N) constructs through *Agrobacterium tumefaciens* mediated *pAtZIP4* methods were grown on half Hoagland solution containing Zn deficiency (0 μM ZnSO₄) (A-C, G-J) or Zn supply (2 μM ZnSO₄) (K-N). *Agrobacterium rhizogenes* mediated transgenic *pAtZIP4::GFP* *N. caerulescens* (D-F) and *pNcZNT1::GFP* *N. caerulescens* (O-W) were grown on half Hoagland solution containing Zn deficiency (0.05 μM ZnSO₄) (O-S), or Zn supply (100 μM ZnSO₄) (T-W). GFP images of transgenic roots were acquire by an inverted laser scanning confocal microscope (LSCM) system, Zeiss LSM 5 PASCAL. (A), (D), (G), (K), (O), and (T) show the GFP fluorescence; (H), (L), (P), and (U) show the roots stained by Propidium iodide; (B), (E), (I), (M), (Q), and (V) are DIC microscopy; (C), (F), (J), (N), (S), and (W) are merging images. Root tissues are denoted by alphabets i.e. x (xylem), p (pericycle), e (endodermis), c (cortex) and ep (epidermis). Scale bar is 50μm.

**5’ deletion analysis in *A. thaliana* roots under Zn deficient and Zn sufficient conditions**

A 5’ deletion analysis of the *AtZIP4* and *NcZNT1* promoters was performed to identify the *cis* elements in these two promoters involved in the regulation of these genes. Deleted promoters were fused to GUS again. The GUS activity was reduced significantly by ~200 fold in the fragment F14 compared to the values obtained from F5 through F12 (Fig. 16). This indicates that there are at least one or more *cis* elements in the region of -263 bp to -115 bp from the start codon of *AtZIP4*. GUS activity is reduced by around ~25 fold from F12 to F13 and ~9 fold from F13 to F14. This suggests that, there is a *cis* element located between -263 bp and -232 bp, which is essential for transcription of *AtZIP4*. Sequence analysis of the two
promoters *AtZIP4* and *NcZNT1* shows two conserved palindromic sequences. The first palindrome is 5’-ATGTCGACAT-3’ which is a true palindromic sequence and the second sequence is 5’-ATGTCGACAC-3’ which is an incomplete palindrome because of the last base (a C instead of T). Interestingly, the proposed first palindrome is located in this region of -263 bp and -232 bp. However, transcription is not completely diminished as indicated by the residual GUS activity in the transgenic plants. This suggests that there is another *cis* element in between -232 bp and -115 bp. This fragment contains the second conserved palindromic sequence in the promoter. The fragment F15 gives the GUS activity similar to the control. This fragment contains only part of the second palindromic sequence but lacks the first three bases of it. In this fragment GUS activity is reduced compared with F13 but it is the same as F14, suggesting the *cis* element requires the complete palindrome for the transcription. This strongly suggests that the conserved palindromic sequences in the *AtZIP4* promoter are potential *cis* elements. Similarly, the 5’ deletions of the *NcZNT1* promoter shows that the GUS activity is reduced 16 fold from the fragment F23 of 220 bp to fragment F24 of 166 bp. GUS activity is further reduced by two fold from the fragment F24 of 166 bp to fragment F25 of 98 bp. The fragments F26 of 103 bp (including only part of the second palindrome) and F25 of 98 bp (in which the second palindrome is present except for the first three bases) give the same GUS activity as the empty vector transformed control values. In summary, the 5’ deletion analysis suggests the *cis* element(s) are localised in between -263 bp and -115 bp from the start codon for *AtZIP4*; between -223 bp and -105 bp for *NcZNT1*.

**Sequence comparison among the ZIP4 promoters from *Arabidopsis thaliana, Arabidopsis lyrata, Arabidopsis halleri, Cochleria pyrenaica* and *Noccaea caerulescens***

In order to look for conserved sequences, that can confirm the importance of identified *cis* elements, ZIP4 promoter fragments from *A. thaliana, A. lyrata, A. halleri, C. pyrenaica* and *N. caerulescens* were isolated. The length of these fragments were respectively 1048, 902, 571, 905 and 746 and 1189 bp long (Table 3). The short *A. halleri* promoter has a deletion of 154 bp compared to the long one, between -538 and -384 bp, whereas the rest of the sequence is identical, except for a few mismatches and a single nucleotide polymorphisms, suggesting that these two promoters represent two different alleles in this heterozygous species. The sequence analysis shows small segments of sequence similarity among these promoter fragments, whereas the rest of the promoter sequences do not have
significant similarity. Interestingly, the two palindromic sequences (5’-ATGTCGACAT-3’ and 5’-ATGTCGACAC-3’) were conserved in the ZIP4 promoters from all five species (Fig. 17). The location of the palindromes in ZIP4 each species and the predicted TATA boxes are summarised in Fig 17.

**Fig 16.** The effect of 5’ deletions of the AtZIP4 and NcZNT1 promoters on GUS expression in transgenic A. thaliana plants. The GUS activity (pmole/min/μg protein) was tested in roots of transgenic (A) pAtZIP4::GUS and (B) pNcZNT1::GUS plants exposed to 0 μM Zn.
Fig 17. Schematic representation of the ZIP4-like promoters of A. thaliana, N. caerulescens, C. pyrenaica, A. halleri and A. lyrata. The black diagram represents the promoter sequence for each species. Two palindromic sequences (dark grey boxes) (5’-ATGTCGACAT-3’ and 5’-ATGTCGACAC-3’) were found to be conserved in all the promoters from five species. These two palindromes are shown in two boxes on each bar. On the 5’ on the bar the predicted upstream genes are shown with distinct colour: dark grey around 80-100% similarity and light grey 50-80% similarity with A. thaliana for each fragment from different species.

Table 3. Summary of the ZIP4 promoters isolated

<table>
<thead>
<tr>
<th>Species name</th>
<th>Length of the promoter fragment isolated (bp)</th>
<th>1st palindrome (bp) [5’-ATGTCGACAT-3’]</th>
<th>2nd palindrome (bp) [5’-ATGTCGACAC3’]</th>
<th>Predicted TATA box (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. thaliana</td>
<td>1048</td>
<td>-246 to -236</td>
<td>-118 to -108</td>
<td>-59 to -55</td>
</tr>
<tr>
<td>N. caerulescens</td>
<td>902</td>
<td>-189 to -179</td>
<td>-107 to -97</td>
<td>-70 to -66</td>
</tr>
<tr>
<td>C. pyrenaica</td>
<td>571</td>
<td>-235 to -225</td>
<td>-116 to -106</td>
<td>-59 to -55</td>
</tr>
<tr>
<td>A. halleri(long)</td>
<td>905</td>
<td>-221 to -211</td>
<td>-115 to -105</td>
<td>-58 to -54</td>
</tr>
<tr>
<td>A. halleri(short)</td>
<td>746</td>
<td>-221 to -211</td>
<td>-115 to -105</td>
<td>-58 to -54</td>
</tr>
<tr>
<td>A. lyrata</td>
<td>1189</td>
<td>-235 to -225</td>
<td>-116 to -106</td>
<td>-59 to -55</td>
</tr>
</tbody>
</table>
**ZNT1 sequence comparison**

When comparing the coding sequence (CDS) of *AtZIP4* (At1g10970.1), *NcZNT1*-LC (Genbank: AF275751.1, from *N. caerulescens* accession La Calamine, LC), and NcZNT1-PY (Genbank: AF133267.1, *N. caerulescens* accession Prayon, PY), we noticed that the first ~90 bp were missing in the published NcZNT1 CDS from PY (Milner *et al.*, 2012), when comparing to NcZNT1-LC. This results in a N-terminal deletion of 30 amino acids which may very well change the functional properties of the NcZNT1 protein. To clarify if this first ATG in LC and *A. thaliana* also exists in PY, we amplified the 5’ regions of NcZNT1 from three *N. caerulescens* accessions, LC, PY, and GA, using qRT-PCR. Our sequencing results showed that all three accessions contain the first 90 bp at their 5’ ends of CDS (Fig. 18). We thus conclude that the NcZNT1 CDS sequence (Genbank AF133267.1) and published by Pence *et al.* (2000) and by Milner *et al.* (2012) is incomplete, which results in the production of N-terminally truncated protein. The expression of this N-terminally truncated protein is likely to affect the results of the functional studies described by Milner *et al.* (2012). To further clarify the gene structure, we also amplified the full length NcZNT1 coding genomic DNA regions from these three accessions. The DNA sequence length (from start codon to stop codon) of NcZNT1 from LC is 1589 bp; from GA it is 1582 bp; and from PY it is 1591 bp. NcZNT1 DNA contains four exons and three introns in all these accessions, with a coding region of 1227 bp, which translates into 408 amino acids (Fig. 19). We also performed a predicted protein comparison of NcZNT1 between LC, PY and AtZIP4 from *A. thaliana*, which showed there are no stop codons or frameshifts suggesting that not the second (used by Milner *et al.*, 2012), but the first ATG is used for translation of these coding sequences (Fig. 20).

**Fig 18.** Comparison of 5’ *NcZNT1* cDNA between *N. caerulescens* accessions La Calamine, Prayon, and Ganges. The translational start site (ATG) is underlined blue (used as start codon in current study). The second ATG, used as a start codon by Milner *et al.* (2012), is underlined black. Forward and reverse primer pairs used for the amplification of this 5’ region are shown as arrows. The alignment was performed by MultAlin software (Multiple sequence alignment by Florence Corpet; [http://multalin.toulouse.inra.fr/multalin](http://multalin.toulouse.inra.fr/multalin)).
Fig 19. Comparison of *NcZNT1* DNA sequences between *N. caerulescens* accessions La Calamine, Prayon, and Ganges. *NcZNT1* DNA fragment were amplified from La Calamine (LC), Prayon (PY), and Ganges (GA) by using the primer pairs shown as black lines (___). Blue arrow indicate part of 5'UTR sequences. Red arrow is the translational start site (ATG) and blue line is the translational stop site (TAG). Four exons were shown in black arrows and three introns were indicated by dotted lines. The alignment was performed by MultAlin software (Multiple sequence alignment by Florence Corpet; [http://multalin.toulouse.inra.fr/multalin](http://multalin.toulouse.inra.fr/multalin)).
Fig 20. Comparison of predicted protein sequences of \textit{NcZNT1} between \textit{N. caerulescens} accessions La Calamine (Long; L), Prayon (Short; S) and \textit{A. thaliana ZIP4}. The first amino acid Methionine (M) (used as first amino acid in current study) is underlined by letter L. The second Methionine (M) (used as first amino acid by Milner et al., 2012) at position 35 is underlined by letter S. The alignment was performed by MultAlin software (Multiple sequence alignment by Florence Corpet; \url{http://multalin.toulouse.inra.fr/multalin}).
DISCUSSION

Understanding plant metal hypertolerance and hyperaccumulation has drawn a lot of interest to unravel its adaptive evolution and to apply this knowledge for phytoremediation and biofortification purposes. A number of genes were analysed for their role in metal tolerance and accumulation. Previous studies have shown the possible role of a ZIP family member gene \textit{NcZNT1} from \textit{N. caerulescens} as a Zn and Cd uptake transporter (Pence et al., 2000; Assunção et al., 2001). Furthermore, it was reported that \textit{NcZNT1} is predominantly expressed in roots and less in shoots in \textit{N. caerulescens} compared to non-accumulator \textit{T. arvense}, but expression of this gene is barely responsive to changes in Zn supply (Assunção et al., 2001). Only at very high Zn concentrations, the expression is somewhat reduced (Pence et al., 2000; van de Mortel et al., 2006). \textit{AtZIP4} is the orthologue of \textit{NcZNT1} in \textit{A. thaliana}, which is also known to be expressed in root and shoot tissues mainly under Zn deficiency (Grotz et al., 1998; van de Mortel et al., 2006).

In the present study, \textit{NcZNT1} gene was found to be highly expressed in both shoots and roots of \textit{N. caerulescens} under Zn deficiency and repressed under Zn excess. This elucidates the important role of this gene in Zn transport. We consider our analysis to be more reliable because we performed a quantitative \textit{NcZNT1} gene expression analyses while a less quantitative RNA blot analyses was carried out in the previous study (Assunção et al., 2001). Recently, \textit{NcZNT1} was found to be highly expressed under Zn deficient conditions and under Cd exposure, particularly in root tissues (Milner et al., 2012). These observations point out that \textit{NcZNT1} is a Zn and Cd uptake transporter mainly in root tissues. Previously, It was found that \textit{NcZNT1} is able to efficiently transport Zn and to a lesser extent Cd, in yeast system (Pence et al., 2000). However, Milner and colleagues (2012) suggest that \textit{NcZNT1} can only transport Zn but not Cd, which is contradictory to what was previously published (Pence et al., 2000). Our transgenic \textit{A. thaliana} lines expressing \textit{p35S::NcZNT1} accumulated higher Zn and Cd than WT, which is consistent with a Zn and Cd transport ability of \textit{NcZNT1}. Thus the Zn and Cd responsiveness of \textit{NcZNT1} expression, and its Zn and Cd accumulation in transgenic \textit{p35S::NcZNT1} expressing \textit{A. thaliana} still suggest it to be a Zn and Cd transporter. The Zn and Cd transport ability of \textit{NcZNT1} is in agreement with it homologues \textit{TjZNT1}, which was able to transport Zn, Cd and even Mn (Mizuno et al., 2005). The yeast cells expressing \textit{NjZNT1} had a higher Ni tolerance. Our transgenic also lines exhibited enhanced shoot Mn accumulation in Zn excess, which urges to analyse the Mn transport ability of \textit{NcZNT1} in future studies.
It is known that under Zn and Cd excess, Fe uptake is compromised in *A. thaliana* (van de Mortel, 2006). As some of the known Fe responsive genes can transport Zn and Cd (Korshunova *et al.*, 1999) they could possibly play a role in indirect Zn and Cd accumulation. Since we found a reduced Fe accumulation in our transgenic lines exposed to Cd excess, we analysed the gene expression of known Fe transporters to find their possible role in indirect Zn and Cd accumulation in our transgenic lines. Fe deficiency responsive genes like *AtbHLH100*, *AtIRT1*, *AtIRT2* and *AtFRO2* were highly upregulated in both transgenic and WT lines, although with some differences, under Zn and Cd excess clearly showing that these lines were experiencing Fe deficiency (Fig. 10; Table 4). Furthermore, higher expression of these known transporters could possibly mediate Zn and Cd accumulation. Particularly *AtIRT1* was previously shown to transport Zn and Cd in addition to Fe in root epidermal cells (Eide *et al*. 1996; Korshunova *et al*. 1999; Rogers *et al*. 2000; Vert *et al*. 2002). Thus there is likely to be some indirect Zn and Cd uptake in *p35S::NcZNT1* *A. thaliana* lines due to upregulation of Fe deficiency responsive machinery. These observations of upregulation of Fe transporters and Cd accumulation in the transgenic lines are in agreement with the previously known Cd uptake by up regulation of *NcIRT1* in *N. caerulescens* under Fe deficient conditions (Lombi *et al.*, 2002). However, as NcZNT1 is a known Cd transporter (Pence *et al*. 2000), we consider it to be a major player in the Cd accumulation in our transgenic *A. thaliana* lines since WT line also had the higher expression of Fe transporters but it could not accumulate enhanced Cd (Fig. 10; Table 4).

Most Zn and Cd metals are stored in vacuoles. Members of the CDF protein family are involved in the sequestration of metals into vacuoles. Some CDF family members, like *AtMTP1*, *PtdMTP1*, *AtMTP3* and *TgMTP1*, were shown to cause increased Zn tolerance and accumulation when ectopically or heterologously expressed in *A. thaliana* (van der Zaal *et al.*, 1999; Blaudez *et al.*, 2003; Arrivault *et al.*, 2006; Gustin *et al.*, 2009), suggesting that their normal function is most likely to create a sink for Zn in the vacuoles of plant cells in case of intracellular Zn excess or as buffer in case of Zn deficiency. Since expression of *AtMTP1* was found to be higher in the shoots and roots of *p35S::NcZNT1* expressing plants under excess Zn and Cd, this clearly illustrates the important role of this gene in the detoxification of Zn metal in these conditions. Thus the enhanced Zn tolerance exhibited by *NcZNT1* expressing line is likely be due to vacuolar sequestration of metals mediated by *AtMTP1* since it is vacuolar membrane localized transporter (Desbrosses-Fonrouge *et al.*, 2005; Kobae *et al.*, 2004). Vacuolar sequestration of
Cd has been reported to confer Cd hypertolerance in *N. caerulescens* ([Ueno et al., 2011](#)) by the HMA3 gene (Heavy Metal ATPase 3). In order to see if *AtHMA3* could also have the higher expression explaining the Cd tolerance in our transgenic lines, we analysed its expression but we did not find any differential expression of this gene in *p35S::NcZNT1* expressing lines compared to WT (Fig. 10; Table 4). This leads to the conclusion that enhanced Cd tolerance exhibited by our transgenic lines is not mediated by *AtHMA3*. Phytochelatins (PCs), metallothioneins (MTs), amino acids and organic acids are compounds that can chelate metals and thus play a role in detoxification of metals in plants ([Ernst et al., 1992](#)). It was reported that phytochelatins (PCs), which can detoxify Cd, had a long distance root to shoot transport together with Cd in *A. thaliana* ([Gong et al., 2003](#)). Since our transgenic lines had higher Cd accumulation in shoots but not in roots, PCs might be involved in Cd tolerance and long distance Cd transport. However, this needs to be investigated to elucidate the role of PCs in Cd tolerance and accumulation in our transgenic lines.

Previously, it was reported that the *AtZIP4* promoter has little activity under sufficient Zn supply but strongly induces transcription in response to Zn deficiency ([van de Mortel., 2006](#)), which agreed with the first study about this transporter ([Grotz et al., 1998](#)). Both *pAtZIP4::GUS* and *pNcZNT1::GUS* constructs, when transformed into *A. thaliana*, showed GUS expression in endodermis and pericycle in roots and also in leaves, trichomes and even flowers after exposure to Zn deficiency (Fig. 11; Fig. 12). This similar expression pattern of *AtZIP4* and *NcZNT1* in *A. thaliana* demonstrated important roles of these genes in Zn uptake in these organs. Furthermore, the promoter analysis led to the identification of two palindromic cis-regulatory elements which were same in both *NcZNT1* and *AtZIP4* promoters. The deletion of these elements resulted in the reduction of GUS activity in *A. thaliana* (Fig. 16). Our group has reported that these cis elements called Zinc Deficiency Response Element (ZDRE), RTGTCGACAY, are present in *AtZIP4* promoter which are the binding sites for basic-region leucine zipper (bZIP) transcription factors *bZIP19* and *bZIP23* ([Assunção et al., 2010](#)). These transcription factors regulate a set of target genes as a response to Zn deficiency as these genes containe ZDRE in their promoters. The *AtZIP4* promoter contains two ZDRE, at -246 to -236bp and at -118 to -108bp, and *NcZNT1* as well, at -189 to -179 bp and -107 to -97bp (Fig. 16). The conserved Zn deficiency responsive elements (ZDRE) were reported for *AtZIP4* and *NcZNT1*, whereas few other sets of cis elements were identified in few micronutrient
responsive promoters (Assunção et al., 2010; Kobyashi et al., 2003). We have cloned and expressed both promoters in N. caerulescens roots and have found that like in A. thaliana, both AtZIP4 and NcZNT1 promoter activities are induced under Zn deficiency in N. caerulescens roots, but repressed by Zn supply (Fig. 13). This implies that a similar regulatory mechanism is involved in regulating these promoters in response to Zn. However, there are also differences. The NcZNT1 promoter showed stronger GUS expression than AtZIP4 promoter in N. caerulescens at low Zn supply, and the former promoter is also active at higher Zn supply levels, while the AtZIP4 promoter is not (Fig. 13). This differential expression of NcZNT1 and AtZIP4 in N. caerulescens compared to their similar expression in A. thaliana points out that there is likely to be an additional cis-element present in the NcZNT1 promoter recognized by a N. caerulescens specific transcription factor, to ensure the higher promoter activity even at higher Zn supply levels compared to AtZIP4 promoter. Recently, another putative cis-region of NcZNT1 promoter, upstream of the known ZDRE, was identified, which was not influenced by the Zn supply of N. caerulescens and could possibly be involved in higher expression of this gene (Milner et al., 2012). From this information and due to the presence of similar known ZDRE elements in the NcZNT1 and ATZIP4 promoters, with similar expression in A. thaliana but different expression pattern in N. caerulescens, we can deduce that a new cis element might have evolved in NcZNT1 promoter as well as a new cis-trans interaction. Of course, only altering the cis element, so that it recruits an existing transcription factor, would be sufficient for hyperexpression in N. caerulescens. However, it would not be unlikely that the regulation of this N. caerulescens transcription factor is also different from A. thaliana. A remarkable characteristic of Zn hyperaccumulators is the high expression of Zn homeostasis genes (Becher et al., 2004; Weber et al., 2004; Hammond et al., 2006; Talke et al., 2006; van de Mortel et al., 2006, 2008). In non-hyperaccumulators these genes are mainly induced upon Zn deficiency and bZIP19 and bZIP23 are the known regulators of Zn deficiency responsive genes (Assunção et al., 2010). Therefore, the transcription factors controlling the Zn deficiency response are likely to be important regulators of hyperaccumulation traits. It will be interesting to look for bZIP19 and bZIP23 in N. caerulescens in order to understand the altered regulation of Zn hyperaccumulation related genes.

The pNcZNT1::eGFP expression in pericycle cells in N. caerulescens and pAtZIP4::eGFP expression in pericycle and cortical cells in A. thaliana implies their role in Zn uptake in root endodermal and pericycle cells (Fig. 15). Their lack
of expression in epidermis is unexpected for transporters involved in root uptake from the soil. For instance, the *AtIRT1* gene involved in Fe-uptake from the soil, was found to be localised in the epidermis of *A. thaliana* roots (Vert *et al.*, 2002). We propose the role of *NcZNT1* and *AtZIP4* in Zn transport from pericycle cells into cells associated with xylem loading, ultimately for long distance transport, rather than their direct involvement in Zn uptake from the soil. During Zn deficiency, creating a strong loading of apoplastic Zn into the stele may be sufficient and no additional, epidermal Zn uptake and subsequent symplastic transport to the stele may be needed (Kramer, 2012). *NcZNT1* and *AtZIP4* expression in stele of older root tissues with less mineral uptake, would prevent Zn leakage and will ensure Zn availability for xylem loading and ultimately Zn supply to the shoot tissues. The role of these genes, particularly of *NcZNT1* in root tissues enabling long distance metal transport is consistent with the known shoot metal hyperaccumulation controlled by root processes in *N. caerulescens* (Guimaraes *et al.*, 2009).

Recently, *NcZNT1* was proposed to be involved in Zn uptake in root tissues and long distance transport (Milner *et al.*, 2012). Our data is consistent with the role of *NcZNT1* in keeping higher influx into cells associated with xylem loading for shoot translocation but rejects its involvement in Zn uptake in root tissues. Milner *et al.* (2012) have also carried out the functional analysis of the *NcZNT1* gene in yeast and in *A. thaliana*. Their p35S::*NcZNT1* expressing *A. thaliana* lines were sensitive to excess Zn but not to excess Cd. A major difference in both studies which might explain some of the difference lies in the fact that they have used a shorter *NcZNT1* cDNA from accession Prayon (*NcZNT1*-PR), missing the first ATG start codon thus resulting in N terminal truncated protein missing 30 amino acids (Fig. 18; Fig. 19; Fig. 20). We have isolated and cloned the full length *NcZNT1* cDNA from accession La Calamine (*NcZNT1*-LC), which population is only some 30 km distant from Prayon (Assunção *et al.*, 2001). The same ATG start codon is found for the *NcZNT1/AtZIP4* orthologue of *A. lyrata* (GenBank acc. no. XM_002892566). Omitting the first ATG of a cDNA is expected to influence its functional analysis and also tissue localization as performed by Milner *et al.* (2012). It is known from literature that N terminal can affect localization of proteins. It was reported that the deletion of 33 amino acids in N terminal of RGS4 protein, a GTPase activating protein, results in loss of plasma membrane localization (Srinivasa *et al.*, 1998). Although, the functional analysis of N terminal of *NcZNT1* for plasma membrane localization is not known, it is plausible that the truncated
NcZNT1 protein may be mislocalized. This is consistent with the fact that NcZNT1 expressing A. thaliana had higher sensitivity to excess Zn (Milner et al., 2012) while we have observed Zn tolerance and accumulation in our p35S::NcZNT1 expressing A. thaliana lines. The possible reason could be the mislocalization of NcZNT1 protein to the organelles which are susceptible to excess metals. The analyses performed by using the shorter version of NcZNT1 are thus questionable. Therefore, we consider our data to be more reliable, however, we suggest to analyse both NcZNT1-LC and N terminal truncated NcZNT1-PR cDNAs together in future to find out their possible functional differences.

ACKNOWLEDGMENTS

This research was supported by the Higher Education Commission of Pakistan (HEC), and the EU-METALHOME project (HPRN-CT-2002-00243) (ST, MA). We would like to thank Dr. R. Jefferson, CAMBIA, Australia, for providing us with the CAMBIA1301 vector. We are grateful to Dr. Erik Limpens, Molecular Biology, Wageningen University, The Netherlands for the kind provision of the pKGWFS-RR vector. We thank Henk Kieft and Dr. Norbert de Ruijter, Plant Cell Biology, Wageningen University, The Netherlands for their help in microscopic analysis and to Mohamed El Soda for help in statistical analysis. We are grateful to Dr. Maarten Koornneef for critical reading and suggestions for the manuscript preparation.
Chapter 4

The selective advantage of a mutant 

*nicotianamine synthase* allele in a natural 

*Noccaea caerulescens* population

Ya-Fen Lin\(^a\), Joop van Loon\(^b\), Henk Schat\(^c\), Holger Schmidt\(^d\), Stephan Clemens\(^d\), Judith van de Mortel\(^a\), Jian Wu\(^a\), Nguyen Duy\(^a\), Xu Cheng\(^a\) and Mark G. M. Aarts\(^a\)*

\(^a\) Laboratory of Genetics, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands;
\(^b\) Laboratory of Entomology, Wageningen University, Wageningen, The Netherlands;
\(^c\) Department of Clinical Genetics, Vrije University, Amsterdam, The Netherlands;
\(^d\) Department of Plant Physiology, University of Bayreuth, Bayreuth, Germany;

*Correspondence: Mark.Aarts@wur.nl*
ABSTRACT

Noccaea caerulescens, a Zn/Ni/Cd hyperaccumulator, is a good model plant species in which to study the evolution of heavy metal hyperaccumulation and tolerance. Several metal homeostasis genes are expressed at higher levels than in related metal non-accumulator species, including genes encoding the enzyme nicotianamine synthase (NAS). This enzyme synthesizes the metal chelator nicotianamine (NA), which is essential for the correct distribution of Fe and Cu. However, modification of NAS activity also affects Zn, Mn and Ni homeostasis. Despite the apparent importance of NAS for metal tolerance, we identified a transposon insertion mutant nas1 allele in a natural population of N. caerulescens growing near a zinc mine smelter deposit in La Calamine (Belgium). The transposon disrupts the normal reading frame so although the gene is transcribed, the protein is predicted to be terminated prematurely at amino acids position 228 instead of 323, which renders the enzyme non-functional. The transposon insertion appears to be stable, and both the mutant and wild-type (WT) alleles are found in the population, suggesting neither is superior in terms of fitness or that there is some form of balancing selection. In this chapter, we investigate the selective advantage of the nas1 allele, given that homozygous nas1/nas1 transposon mutants are clearly more sensitive to Zn, Cd and Ni than near-isogenic NAS1/NAS1 WT plants. The nas1 mutant also accumulated more Zn and Cd in the shoots compared to WT plants, which may boost the concentration of NA in the roots thus assisting the transport of Zn and Cd to above-ground tissues. The NAS3 and NAS4 genes were overexpressed in the nas1 mutant, suggesting functional compensation for the loss of NAS1 protein. Furthermore, when Pieris rapae caterpillars were raised on nas1/nas1 mutants growing in high-Zn or high-Cd environments, their growth and development was more severely disrupted than caterpillars raised on NAS1/NAS1 WT plants. We concluded that the nas1 mutant is advantageous because it promotes the accumulation of more metals and thus offers protection against herbivores. However, the mutant plants were also smaller and less metal tolerant. Therefore, local conditions such as the presence of herbivores and the abundance of metals in the soil will determine which of these properties has a more prominent impact.

KEYWORDS

Nicotianamine, NAS1, Noccaea caerulescens, inorganic defense hypothesis, herbivory.
INTRODUCTION

Human activities such as industry and mining introduce metals into the environment and threaten the survival of plants. Certain plants have adapted to metalliferous and serpentine soils by evolving into hyperaccumulators that can store high concentrations of metals, and such plants are also metal hypertolerant (Kramer, 2010). More than 500 hyperaccumulator plant species have been described, which are tolerant to one or more metals such as Zn, Ni, Cd, Cu, Mn, Se, Pb, As, Co and Tl (van der Ent et al., 2013). Understanding the molecular and biochemical mechanisms that allow such species to accumulate heavy metals at 2–3 times the magnitude of non-accumulators could provide insight into the natural selection and adaptation of plants to survive in metalliferous environments.

Noccaea caerulescens (J. & C. Presl) F. K. Mey is a well-known Zn/Ni/Cd hyperaccumulator, which shares 88% sequence identity with Arabidopsis thaliana at the transcriptome level, making it an ideal model for the analysis of metal hyperaccumulation mechanisms. It is also a suitable model in which to investigate the evolutionary basis and ecological impact of metal hyperaccumulation because it grows in soils with diverse characteristics in terms of the types and quantities of metals, and exists as a range of genotypes that differ in their hyperaccumulation and metal tolerance properties (Assunção et al., 2003a, Milner&Kochian, 2008, Peer et al., 2003, Rigola et al., 2006a).

Four hypotheses have been put forward to explain the advantages of hyperaccumulation: (1) the tolerance/disposal hypothesis; (2) the interference hypothesis; (3) the drought resistance hypothesis; and (4) the elemental defense hypothesis (Boyd&Martens, 1992, Boyd, 2004). Only the last of these, more recently described as the inorganic defense hypothesis (Boyd, 2012), is supported by experimental evidence (Hörger et al., 2013). The defense hypothesis proposes that higher concentrations of metals in edible organs benefit plants by conferring a fitness penalty on herbivores and pathogens (Boyd, 2007, Rascio&Navari-Izzo, 2011). For example, higher Zn levels have been shown to benefit N. caerulescens by suppressing the symptoms of mildew caused by Erysiphe spp. (Fones et al., 2010) whereas moderate levels of Zn can modulate the feeding behavior of Schistocerca gregaria (forskal) (Behmer et al., 2005). Similarly, thrips (Frankliniella occidentalis) favors the N. caerulescens accession Prayon, which accumulates low levels of Cd, over Ganges, which is a Cd hyperaccumulator (Jiang et al., 2005). However, the putative protective role of Cd in N. caerulescens has not been studied in detail and that of Zn is still a matter of debate.
Metal ions cannot be transported among plant tissues in the form of free ions (Callahan et al., 2006), instead forming complexes with ligands such as amino acids (histidine and glutamine), proteins (metallothioneins), organic acids (citrate and malate) and oligopeptides (glutathione, phytochelatins and nicotianamine) that act as intercellular chelators (Haydon & Cobbett, 2007, Verbruggen et al., 2009). The symplastic chelator nicotianamine (NA) is synthesized from three molecules of S-adenosylmethionine in a reaction catalyzed by nicotianamine synthase (NAS). This molecule can bind several different heavy metals such as Ni, Cu, Fe and Zn, and is required for the long-distance transport and intercellular translocation of micronutrients in plants (Clemens et al., 2013, Curie et al., 2009). Cytosolic NA levels are regulated by zinc-induced facilitator 1 (ZIF1), a member of the major facilitator superfamily (MFS), which is localized on the vacuolar membrane and translocates NA into the vacuole (Haydon et al., 2012). Precise ZIF1 expression is necessary for correct Zn and Fe homeostasis in plants because ZIF1 can enhance Zn sequestration and Fe translocation (Clemens et al., 2013, Hofmann, 2012).

The tomato (Solanum lycopersicum) chloronerva mutant was the first NA-free mutant to be described, and its residual NAS activity was shown to cause Fe deficiency symptoms compared to wild-type plants (Ling et al., 1999). NAS genes have been identified in many dicot and monocot species, including barley, rice, maize, sorghum, tobacco, legumes, grapevine and A. thaliana (Bauer et al., 2004, Clemens et al., 2013, Higuchi et al., 1999, Inoue et al., 2003, Ling et al., 1999, Mizuno et al., 2003). The A. thaliana genome contains four unlinked NAS genes, which are partially functionally redundant as indicated by the individual mutant phenotypes (Bauer & Schuler, 2011, Klatte et al., 2009). Comparative sequence analysis also identified four NAS genes in the metal hyperaccumulators A. halleri and N. caerulescens, and these genes were found to be overexpressed compared to their A. thaliana orthologs (van de Mortel et al., 2006, Weber et al., 2004). This implied that NAS genes may play an important role in metal hyperaccumulation and tolerance, and this was supported by the loss of Zn hyperaccumulation in A. halleri plants by suppressing NAS2 expression, eliminating its root-to-shoot Zn translocation capability (Deinlein et al., 2012). Thus far, the functions of the remaining three NAS genes have not been studied in detail in metal hyperaccumulators.

Given that NAS genes appear to play a pivotal role in mineral homeostasis, it was surprising to find a transposon insertion allele in the NAS1 cDNA isolated from a N. caerulescens specimen collected near a zinc mine smelter deposit at La Calamine, close to Liège in Belgium. The transposon interrupts the NAS1 open
NAS1 mutant and its selective advantages

reading frame causing a frameshift and premature stop codon. We set out to establish the frequency of this mutant allele in the population and its impact, if any, on the Zn/Cd hyperaccumulation/hypertolerance phenotype. With the inorganic defense hypothesis in mind, we also tested for differences in the growth and development of *Pieris rapae* larvae, especially when plants were raised on soils containing excess Zn, Cd or Ni. Based on the results, we speculate on the potential advantages of this mutant allele in a natural population.
MATERIALS AND METHODS

Plant material and growth condition

*N. caerulescens* leaves and inflorescence samples were collected from several villages in Belgium, Kelmis/La Calamine (LC) in 2009 (50 samples), Plombières (PB) in 2009 (73 samples) and 2011 (50 samples), and Prayon (PR) in 2013 (40 samples). The origin of each sample was positioned using a global positioning system (GPS). The GPS location and number of plants is summarized in Table 1 and Fig 1C.

A NAS1/nas1 plant collected from La Calamine was selfed to yield near-isogenic plants differing only in the NAS1 genotype. Homozygous seeds were obtained by selfing selected NAS1/NAS1 (WT) and nas1/nas1 (nas1) plants. To obtain genetically homogenous seeds, the WT and nas1 plants were cultivated in the greenhouse, and experiments were carried out on the subsequent generation of seeds.

*N. caerulescens* seeds were germinated in half-strength Murashige and Skoog (½MS) medium with 0.8% agar (MS salts including vitamins, Duchefa Biochemie, cat no. M0222.0025). They were grown for 2 weeks and then transferred to half-strength Hoagland’s solution (½HS) containing 10 μM ZnSO₄ as the normal Zn supply, and left for another week (Schat et al., 1996). Seedlings were transferred into ½HS supplemented with four different metal concentrations: (1) normal (½HS with 10 μM ZnSO₄); (2) Ni treatment (½HS with 2 μM ZnSO₄ and 10 μM NiSO₄); (3) Cd treatment (½HS with 2 μM ZnSO₄ and 5 μM CdSO₄); and (4) Zn treatment (½HS with 100 μM ZnSO₄). The nutrient solution was replaced every week. After 3 weeks, the *N. caerulescens* plants were exposed to high metal concentrations: (1) normal conditions (½HS with 10μM ZnSO₄); (2) high Ni treatment (½HS with 10 μM ZnSO₄ and 100 μM NiSO₄); (3) high Cd treatment (½HS with 10 μM ZnSO₄ and 50 μM CdSO₄); and (4) high Zn treatment (½HS with 1000 μM ZnSO₄). All plants were grown in a climate chamber with a 12-h photoperiod, a 20/15°C day/night temperature regime and 70% relative humidity. After another 3 weeks, roots and shoots were collected separately for the analysis of gene expression, metal concentration, NA concentration and protein concentration.
Table 1. Summary of genotyping experiments showing the prevalence of transposon-insertion nas1 alleles in natural habitats

The distribution of the transposon-insertion allele nas1 was studied in the N. caerulescens samples collected from La Calamine (LC), Plombières (PB) and Prayon (PR), Belgium, in 2009, 2011 and 2013. Table shows the number of plants found representing each genotype and the allelic frequency.

<table>
<thead>
<tr>
<th>Year</th>
<th>Site</th>
<th>Latitude/longitude</th>
<th>Number of plants found</th>
<th>NAS1 allele frequency</th>
<th>nas1 allele frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>La Calamine</td>
<td>50°42'19.04&quot;N/6°0'21.42&quot;E</td>
<td>12/21/17</td>
<td>0.41</td>
<td>0.59</td>
</tr>
<tr>
<td>2009</td>
<td>Plombières</td>
<td>50°44'2.94&quot;N/5°58'5.04&quot;E</td>
<td>21/25/27</td>
<td>0.45</td>
<td>0.55</td>
</tr>
<tr>
<td>2011</td>
<td>Plombières</td>
<td>50°34'54.2&quot;N/005°39'52.0&quot;E</td>
<td>10/15/25</td>
<td>0.09</td>
<td>0.91</td>
</tr>
<tr>
<td>2013</td>
<td>Prayon</td>
<td></td>
<td>3/1/36</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Library screening

N. caerulescens root and shoot cDNA libraries (Assunção et al., 2001) were used for full-length cDNA cloning. The shoots cDNA library was screened using a NcNAS1 probe fragment and the roots cDNA library was screened using a NcNAS2 probe fragment. DNA fragments were amplified from N. caerulescens accession ‘La Calamine’ genomic DNA using primers based on the A. thaliana NAS1 sequence (forward primer 5′-ACA TGG GTT GCC AAG ACG-3′, reverse primer 5′-ATG AGA AGG CAC GAG ACT CC-3′) and NAS2 sequence (forward primer 5′-ATG GCT TGC GAA AA-3′, reverse primer 5′-CTT CTC CAT CAT ACC ACG ATT GT-3′). The reaction was heated to 94°C for 2 min, followed by 35 cycles at 94°C for 30 s, 55°C for 30 s and 72°C for 45 s, with a final extension at 72°C for 2 min. A single 0.9-kbp product was generated for NAS1 and a single 1.0-kbp product for NAS2. We then labeled 1 µl of each product with [α-32P]dATP using a Hexalabel™ DNA labeling kit (Fermentas, Loughborough, UK). Nylon membranes containing the cDNA library plaques were hybridized as described by Assunção et al. (2001). The isolated cDNA clones were sequenced using ABI PRISM BigDye terminator cycle sequencing technology v2.0 on an ABI3700 DNA analyzer, according to the manufacturer’s recommendations (Applied Biosystems, Foster City, CA).
Genotyping

*N. caerulescens* DNA was extracted using a modified nuclear extraction protocol (Aarts et al., 2000). One primer set was designed with primers Xu002 and Xu003 flanking the *NAS1* transposon insertion to generate products of 261 and 610 bp for the WT and insertion mutant, respectively (Fig. 1, Table 2). The PCR was heated to 94°C for 2 min, followed by 30 cycles at 94°C for 30 s, 55°C for 30 s and 68°C for 1 min, with a final extension at 68°C for 10 min. The products were separated by 1% agarose gel electrophoresis.

To design additional PCR primers to genotype InDel and CAPS neutral genetic markers, ESTs that might be involved in metal homeostasis were selected from an EST library of *N. caerulescens* accession ‘La Calamine’ (Rigola et al., 2006). The cDNA sequences were compared to the Arabidopsis ‘AGI whole genome’ database at TAIR (The Arabidopsis Information Resource, www.Arabidopsis.org) to locate putative introns and coding regions in *N. caerulescens*. Primer pairs were designed for InDel and CAPS markers to amplify the intron or coding region of metal homeostasis related genes from *N. caerulescens* (Table 2). Products distinguishing large InDels were analyzed by 1.2% agarose gel electrophoresis whereas those distinguishing small InDels were analyzed using 3% Metaphor (Tebubio, cat no. 50180) agarose gels. The products generated using CAPS primers were digested with the appropriate enzyme and analyzed by 2% agarose gel electrophoresis. Heterozygosity was calculated by dividing the number of heterozygous plants by the total number of tested plants.

Measurement of metal concentrations

Roots removed from WT and *nas1* plants were dipped into ice-cold 5 mM Pb(NO$_3$)$_2$ for 30 min, and shoots were rinsed in Milli-Q water. The samples were dried at 65°C, homogenized, and 100 mg of sample was wet-ashed in a 4:1 mixture of 2 mL HNO$_3$ (65%) : HCl (37%) in Teflon bombs at 140°C for 7 h. The concentration of Zn, Fe, Ni, Cu, Mn and Cd was determined using an AAnalyst 100 atomic absorption spectrometer (Perkin Elmer). The metal concentration was calculated as a proportion of plant dry weight.
### Table 2. Primer list for heterozygosity analysis

We chose 24 primer pairs for heterozygosity analysis in the Noctaaea population from Plombière. Three types of primers were used in the experiments, including normal PCR primers, InDel primers, and CAPS primers. Results: + indicates the primer pairs showed heterozygosity in the population, — indicates there was no heterozygosity detected in the population. N.D. means no PCR product was detected. Heterozygosity was calculated using the following formula: detected number of heterozygous samples (X) / number of total detectable samples (Y), shown in the table as X/Y.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Primer sequence</th>
<th>Type of marker</th>
<th>Polymorphic</th>
<th>Gene Annotation</th>
<th>Heterozygosity (genotype)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>C 1</td>
<td>5'-ACCGAGCTTCTCTTCAATCTCCTTCG-3'</td>
<td>+</td>
<td>Metal transporter, putative (ZIP5)</td>
<td>0.30 (10:15:25)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>5'-TGAAGGGAAACTTGCGCCCC-3'</td>
<td>—</td>
<td>germin-like protein (GLP4) (GLP5)</td>
<td>0.09 (3:4:39)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>5'-ACCGAGCTTCTCTTCAATCTCCTTCG-3'</td>
<td>Indels</td>
<td>Metal transporter, putative (NRAMP3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5'-TGAGGTGAGGGATCATAGGC-3'</td>
<td>Indels</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5'-TGGTTCCATAACCACAAGAG-3'</td>
<td>Indels</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>5'-GGGCTTTGTTGAATGGAATG-3'</td>
<td>Indels</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>5'-AGAAATCTTCTGCTTTCGCT-3'</td>
<td>Indels</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>5'-CAGACTCCAGTCATTGTTCG-3'</td>
<td>Indels</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>5'-TCCTTGCTCTCTTGTCCTGGC-3'</td>
<td>Indels</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>5'-CTCTTGCTCTCTTGTCCTGGC-3'</td>
<td>Indels</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>5'-TAAGAAGGTCGAAATGCGCGC-3'</td>
<td>Indels</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>5'-ATGTCATCTTCCTCTTCCTCTC-3'</td>
<td>Indels</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>5'-CCTGAGGCTAAGAAAGGAAG-3'</td>
<td>Indels</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>5'-TAAAGGAAACTCCCAATCTC-3'</td>
<td>Indels</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>5'-AACTGGCGGCTACGGAAGAG-3'</td>
<td>Indels</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>5'-TTACAATCCAGCCCAGAGAAGAATC-3'</td>
<td>Indels</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>5'-AAGGAGAGATCCTCAAGCCCAAC-3'</td>
<td>Indels</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>5'-AAGGACATACACCCAACCACCGAGAAG-3'</td>
<td>Indels</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>5'-GGCTCTCTCACTTCAATTTCTG-3'</td>
<td>Indels</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>5'-AACTTCGCTGATGCACTTTCGC-3'</td>
<td>Indels</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>5'-TCCTGCTCTCTTCCTCTCTTCG-3'</td>
<td>Indels</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>5'-CTCTTGCTCTCTTGTCCTGGC-3'</td>
<td>Indels</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>5'-AAGAAGGAAACTCCCAATCTC-3'</td>
<td>Indels</td>
<td>—</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: Heterozygosity (genotype) is calculated as the number of heterozygous samples (X) divided by the number of total detectable samples (Y), shown in the table as X/Y.
Transcript analysis

Total RNA was extracted from roots and shoots using the RNaseasy® Plant Mini kit (Qiagen, cat. no. 74124) followed by on-column DNase treatment to remove genomic DNA. The RNA quality and concentration were determined using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific) to ensure that each RNA sample had an A260:A280 ratio of 2.1 and an A260:A230 ratio ≥2.1. First cDNA strand was synthesized from 1 µg of RNA using the iScript™ cDNA Synthesis Kit (Bio-Rad, cat. no. 170-8890). Quantitative real-time PCR (qPCR) was carried out using iQ™SYBR® Green Supermix (Bio-Rad, cat. no. 170-8880) in reactions comprising 12.5 µL of iQ SYBR Green supermix, 5 pmol of the forward and reverse primers, and 5 µL of five-fold diluted cDNAs (corresponding to 10 ng/µL RNA) in a total volume of 25 µL. Primer pairs (Table 3) representing four reference genes (Clathrine, EF1-α, TcUBC, and TcTublin) and four target genes (NAS1, NAS2, NAS3 and NAS4) were designed and tested (Gendre et al., 2007, van de Mortel et al., 2006). PCR was carried out by heating the reactions to 95°C for 3 min, followed by 40 cycles of 95°C for 10 s and 62 °C for 1 min. Melting curves were analyzed to confirm the absence of nonspecific products and primer dimers. No-reverse-transcription controls were included to confirm the absence of genomic DNA. The fluorescence signals were detected using a CFX96™ Real-Time Detection System (Bio-Rad). Ct values from four biological repeats (and two technical repeats per biological repeat) were included in the analysis. The difference between technical repeats was less than 0.2 cycles. The relative transcript levels (RTLs) were calculated using the 2^∆∆Ct method (Livak&Schmittgen, 2001). Gene expressions in WT shoots under normal exposure conditions (Zn10) were used for calibration (RTL = 1). The stability of housekeeping gene expression was calculated using geNorm in qBasePLUS software (Biogazelle) and reference genes with a geNorm M value lower than 0.5 were regarded as stable (Hellemans et al., 2007).

Measurement of nicotianamine concentration

Nicotianamine concentrations were determined as described by Deinlein et al. (2012), using 100 mg of well-mixed sample powder to measure Fmoc-derivatized nicotianamine (three injections per sample).
Table 3. Primer list for NAS gene expression analysis by qRT-PCR

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer</th>
<th>Sequence (5′-3′)</th>
<th>Amplicon size (bp)/citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>NcNAS1</td>
<td>Forward</td>
<td>GACACTTTGTTCCGGGAACATCG</td>
<td>160</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CCAAGATTGTTGGGAGAAGTGGTGC</td>
<td></td>
</tr>
<tr>
<td>NcNAS2</td>
<td>Forward</td>
<td>GACAATCGTGCTCCACGTGTTACC</td>
<td>521</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>GCCTATGATTGTTGGGAGAAGTGGTGC</td>
<td></td>
</tr>
<tr>
<td>NcNAS3</td>
<td>Forward</td>
<td>TCTCGAAGCTCGAGAGTCTGAAACC</td>
<td>146</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CTTGATGAGTTTTTGTGGAATCTCC</td>
<td></td>
</tr>
<tr>
<td>NcNAS4</td>
<td>Forward</td>
<td>GTGACAGCATCTGCGATCTGACG</td>
<td>147</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>GCTTTGACAGCATCTTTGTGAGCTGC</td>
<td></td>
</tr>
<tr>
<td>NcTubulin</td>
<td>Forward</td>
<td>CCTACGCACCAGTCATCTCTGC</td>
<td>151 (Gendre et al., 2007)</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CCACGTCACATCGACAGCAAGC</td>
<td></td>
</tr>
<tr>
<td>Clathrine</td>
<td>Forward</td>
<td>AGCATAACATCGTGCAAAAGGC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>TCGCCTGTGTCACATATCTC</td>
<td></td>
</tr>
<tr>
<td>Ef1 α</td>
<td>Forward</td>
<td>GTGACGATTCTGGAAGTGGACC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>AATGCTAATGGTGATACCACGC</td>
<td>(Gendre et al., 2007)</td>
</tr>
<tr>
<td>UBC</td>
<td>Forward</td>
<td>GGAGCCCCCGCTTGGAC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>CGGCGAGGGCGAGTGTA</td>
<td></td>
</tr>
</tbody>
</table>

Pieris rapae herbivory experiments

The life cycle of *P. rapae* includes four stages: egg, caterpillar, pupa and butterfly (http://www.fastplants.org/). The caterpillar passes through five instars, numbered L1 to L5. Choice and no-choice trials were carried out using plants treated with metals for 6 weeks. Choice experiments were carried out with *P. rapae* L4 caterpillars and butterflies. Non-choice experiments were performed using L1 caterpillars.

Choice experiments with caterpillars were set up by cutting WT and *nas1* leaves into segments and placing them together with one large caterpillar (L4) in a 5-cm diameter circular plate (14–15 replicates per treatment). After 17 h, photographs were taken of all the plates and the remaining leaf area was calculated using Image J software (Abramoff et al., 2004, Schneider et al., 2012). The consumed area was determined by subtracting the final leaf area from the original leaf area and expressing this as a percentage.
For no-choice experiments two setups were used. In the first setup, three L1 caterpillars that had hatched during the last 12 h were inoculated on the same leaf of each plant and covered with a clip cage. Plants, together with caterpillars, were both placed in a climatic chamber under the growth conditions described above. To prevent caterpillar escape, pots were placed on iron trays surrounded by water. The experiment was terminated after 6 days and surviving caterpillars were removed from the plants. The plants were allowed to recover for 2 weeks for subsequent experiments. In the second setup, five L1 caterpillars were inoculated on WT and nasl N. caerulescens plants. To avoid caterpillars falling and escaping, the pots were covered with transparent film and surrounded with water. After 15 days, caterpillars and plants were collected to allow the fresh and dry weights to be determined.

**Protein concentrations**

Protein concentrations were determined using the Dumas method (Rhee, 2001). We sealed 10–15 mg of dried, ground plant material in an aluminum cup and measured the nitrogen concentration by FlashEA® 1112 N/Protein (Thermo Scientific) with Eager 300 software. The measurement started with a blank measurement (cellulose, 10 mg) and was followed by five samples of standards (methionine at 3, 6, 9, 12 and 15 mg). To confirm the stability of the measurement, one blank sample was analyzed after every 10 samples, and one positive sample (methionine, 10 mg) was measured after every 20 samples. The protein concentration was calculated from the nitrogen concentration by multiplying by a nitrogen-to-protein conversion factor, 6.25.

**Statistical analysis**

Two-way ANOVA followed by the post-hoc Tukey’s test was applied using PASW Statistics 17 (SPSS Inc) to show the difference between treatments (normal Zn, high Zn, high Ni and high Cd). The Simple Main Effect test was also applied with the same PASW Statistics 17 software to compare the differences between genotypes (WT and nasl).
RESULTS

Identification, frequency and distribution of a mutant transposon-insertion nas1 allele in a natural N. caerulescens population

We initiated the analysis of NAS genes in the metal hyperaccumulator N. caerulescens by screening cDNA libraries prepared from the accession ‘La Calamine’ (Assunção et al., 2001) and cloning a full-length cDNA for the N. caerulescens orthologs of AtNAS1 (At5g04950), AtNAS2 (At5g56080), AtNAS3 (At1g09240) and AtNAS4 (At56430). The N. caerulescens NAS1, NAS2, NAS3, and NAS4 genes showed 91.3%, 91%, 87.6% and 90.2% identity to their A. thaliana counterparts (van de Mortel, 2007). As well as the normal full-length clone, a longer NAS1 cDNA was isolated and found to contain a 349-bp insertion within the open reading frame (Fig 1A, Appendix 1), which was not present in NAS2, NAS3 or NAS4. The insert comprised two 118-bp long terminal repeats (LTRs) and the 6-bp target site duplication (TSD) ACACAT. The presence of the TSD and the absence of homology to reverse transcriptase or any other protein sequence suggested the insertion is a non-autonomous transposon.

DNA blot hybridization against N. caerulescens genomic DNA revealed ~30 copies of this transposon in N. caerulescens plants collected at Plombières, Belgium (calamine accession, found very close to where ‘La Calamine’ was collected) and serpentine accessions from Monte Prinzera, Italy (Assunção et al., 2003b). Furthermore, ~15 copies were found in plants collected from an old Zn-smelter deposit at Durfort, France (calamine accession). BlastN analysis indicated that related sequences (80–85% identity) are present in the A. thaliana genome (data not shown). The transposon insertion causes a frame-shift in the predicted NAS1 protein sequence, introducing a premature stop codon at amino acid position 228 instead of 323 (Fig 1A, Appendix 2).

To explore the distribution of this nas1 allele in natural N. caerulescens populations, a PCR assay was developed to distinguish homozygous wild type (NAS1/NAS1, NN), heterozygous (NAS1/nas1, Nn) and homozygous mutant (nas1/nas1, nn) plants (Fig 1B). Leaves and flowers were collected from plants growing at three sites near Liège, Belgium, around the villages of La Calamine
(LC), Plombières (PB) and Prayon (PR) (Fig 1C). The site at Plombières was sampled for two years, from approximately the same locations (Fig 1D). The analysis of genotypes and allele frequencies showed that in all of these populations, the nas1 allele was more prominent than the wild-type NAS1 allele (Table 1).
Fig. 1. Identification and distribution of the transposon-insertion mutant nas1

(A) Map of the NAS1 cDNA showing the transposon insertion. The transposon contains two 118-bp long terminal repeats (LTRs) and 6-bp target site duplications (TSDs) flanking the transposon. The corresponding protein is terminated prematurely at amino acid position 228 instead of 323. The Xu002/Xu003 primer pair was designed to identify the NAS1 and mutated nas1 allele. (B) PCR to identify homozygous wild type plants (WT, NAS1/NAS1, NN), heterozygous plants (NAS1/nas1, Nn), and homozygous nas1 mutants (nas1/nas1, nn) using primers Xu002 and Xu003. NN shows a single band representing NAS1 (261 bp), and nn shows a 610-bp band including the integrated transposon. Nn displays both bands. (C) Natural N. caerulescens populations were collected from the sites shown, labeled using Google Earth. The yellow line represents national borders. The distance between Plombières (PB) and La Calamine (LC) is ~6 km, with Prayon (PR) ~35 km further. The detailed PB collection site (2009 and 2011) is shown in panel (D).

In addition to the N. caerulescens accessions from Plombières, La Calamine and Prayon, we also screened for the nas1 allele in a variety of other locations and found that it was not present in plants collected from Lellingen/Wilwerwilz, Luxemburg (LE); Monte Prinzera, Italy (MP); St. Felix de Pallières (SF), St. Laurent Le Minier/Ganges (GA), Col du Mas de l’Ayre (CMA), Durfort, (DF), France; the Krušné hory mountains, Bohemia (KH), or the Moravsko slezské Beskydy mountains, Moravia, Czech Republic (MB). It was also absent from N. praecox accessions from Wulfen, Mezica Slovenia (SLO) and N. minima accessions from Hochobir, Austria (data not shown).

Heterozygosity is a good indicator of genetic variability and may indicate natural selection in favor of allelic diversity. To determine whether heterozygosity at the NAS1 locus is similar to that of other genes in this predominantly self-fertilizing species, we also analyzed a number of neutral genetic markers in the PB population collected in 2011. Among a total of 23 markers, only four were polymorphic in this population, indicating the population is genetically not very heterogeneous, and allowing the degree of heterozygosity to be quantified (Table 2). The NAS1 locus appeared to show significantly (P-vale <0.05) more heterozygosity (0.30) than the other three loci, represented by primer set 1 (0.09), primer set 5 (0.18) and primer set 16 (0.14) based on chi-square analysis, but no difference to the locus represented by primer set 6 (0.24). This indicates that natural selection maintains heterozygosity at the NAS1 locus in this N. caerulescens
population. Even though the *NAS1* gene is thought to be important for proper Zn homeostasis, the prominence of the *nas1* allele in populations growing on metalliferous soils suggests there must be an important reason for its selection.

**The *nas1* mutant is more sensitive to heavy metals than wild type plants and accumulates more metal in the leaves**

The prominence of *nas1* in natural *N. caerulescens* populations was initially investigated by studying plant phenotypes under different metal exposure conditions. Genetically homogenous, but near isogenic, wild-type (*NAS1/NAS1*) and *nas1* mutant (*nas1/nas1*) plants were grown hydroponically, and exposed to different heavy metal environments: control (10 μM ZnSO₄), high Zn (1 mM ZnSO₄), high Cd (50 μM CdSO₄) and high Ni (100 μM NiSO₄). WT plants performed better than mutants (Fig. 2A) under all metal treatments. The *nas1* mutants were always smaller than WT plants, and accumulated less biomass as reflected by the dry weight of the shoots (Fig. 2B). The stunting of the *nas1* mutants was already evident after 2 weeks exposure to low concentrations of metals. The dry weight of the roots was significantly reduced in the *nas1* mutant compared to WT roots under control and high Ni conditions, but was only slightly lower under high Zn and high Cd conditions (Fig. 2C, right). Furthermore, the *nas1* mutants were also visually more sensitive than WT plants to excess Zn, Cd and Ni treatments, producing yellow and/or curled leaves (Fig. 2A).

The sensitivity of the *nas1* mutant to excess Zn, Cd, and Ni, was possibly due to the accumulation of more metal in the mutant plants. To test this hypothesis, we measured Zn, Fe, Cu, Mn, Ni and Cd levels in the roots and shoots of WT and *nas1* plants exposed to the four different metal treatments. Overall, there was little difference in metal concentration when we compared WT and *nas1* mutants, except that higher levels of Zn were observed in mutant shoots under the Zn1000 regime, and higher levels of Cd and Cu were observed in mutant shoots under the Cd50 regime. We also observed lower levels of Zn, Mn and Cu in the roots of *nas1* mutants. However, the *nas1* mutant showed a significantly higher capacity for the translocation of Zn, Cd and Mn compared to WT plants when exposed to excess Zn and Cd (Figs 3 and 4). There was no significant difference in the efficiency of Ni and Fe translocation between WT and *nas1* plants (Fig. 5).
Fig. 2. Metal tolerance of WT and nas1 mutants in response to different metal treatments
Nine-week-old wild-type (WT, black) and nas1 mutant (white) plants exposed to different metal treatments: normal (10 µM ZnSO₄, Zn10), high Zn (1000 µM ZnSO₄, Zn1000), high Cd (50 µM CdSO₄, Cd50) and high Ni (100µM NiSO₄, Ni100) (A). The dry weights of the shoots (B) and roots (C) are shown in mg (mean ± SE). Letters shows differences between treatments analyzed by two-way ANOVA followed by Tukey’s post-hoc test and a simple main effect test.
Fig. 3. Zn and Cd accumulation in WT and nas1 mutants under different metal treatments
Nine-week-old wild-type (WT, black) and nas1 mutant (white) plants exposed to different metal treatments: normal (10 µM ZnSO4, Zn10), high Zn (1000 µM ZnSO4, Zn1000), high Cd (50 µM CdSO4, Cd50) and high Ni (100µM NiSO4, Ni100). Zn concentration in the shoots (A) and roots (C) under the four treatments, shown as mean ± SE (mg kg⁻¹ dry weight). Cd concentration in the shoots (B) and roots (D) under the Cd50 treatment. The Zn translocation efficiency (E) and Cd translocation efficiency (F) are calculated as shoot/root concentration ratios and shown as mean ± SE (C). Letters shows differences between treatments analyzed by two-way ANOVA followed by Tukey’s post-hoc test and a simple main effect test. A start represents significant difference between WT and nas1 mutant under Cd excess by Students’-t Test.
Fig. 4. Mn and Cu accumulation in WT and nas1 mutants under different metal treatments. Nine-week-old wild-type (WT, black) and nas1 mutant (white) plants exposed to different metal treatments: normal (10 µM ZnSO₄, Zn10), high Zn (1000 µM ZnSO₄, Zn1000), high Cd (50 µM CdSO₄, Cd50) and high Ni (100 µM NiSO₄, Ni100). Mn concentration in the shoots (A) and roots (C) under the four treatments, shown as mean ± SE (mg kg⁻¹ dry weight). Cu concentration in the shoots (B) and roots (D) under the four treatments, shown as mean ± SE (mg kg⁻¹ dry weight). The Mn translocation efficiency (E) and Cu translocation efficiency (F) are calculated as shoot/root concentration ratios and shown as mean ± SE (C). Letters show differences between treatments analyzed by two-way ANOVA followed by Tukey’s post-hoc test and a simple main effect test.
Fig. 5. Fe and Ni accumulation in WT and nas1 mutants under different metal treatments

Nine-week-old wild-type (WT, black) and nas1 mutant (white) plants exposed to different metal treatments: normal (10 µM ZnSO₄, Zn10), high Zn (1000 µM ZnSO₄, Zn1000), high Cd (50 µM CdSO₄, Cd50) and high Ni (100 µM NiSO₄, Ni100). Fe concentration in the shoots (A) and roots (C) under the four treatments, shown as mean ± SE (mg kg⁻¹ dry weight). Ni concentration in the shoots (B) and roots (D) under the Ni100 treatment. The Zn translocation efficiency (E) and Cd translocation efficiency (F) are calculated as shoot/root concentration ratios and shown as mean ± SE (C). Letters show differences between treatments analyzed by two-way ANOVA followed by Tukey’s post-hoc test and a simple main effect test.
Because NA is known to facilitate metal transport in plants, we proposed that the enhanced metal translocation in the nas1 mutants may reflect a higher level of NA compared to WT plants. We therefore measured the NA concentrations in WT and nas1 plants exposed to different metal treatments (Fig. 6). There was no significant difference in NA levels when we compared the shoots of WT and nas1 plants under normal conditions (Fig 6A) but NA levels in nas1 roots were significantly higher under high Zn condition (Fig 6B). A similar albeit non-significant increase in NA levels was observed in nas1 roots exposed to excess Cd.

Fig. 6. Nicotianamine concentration in WT and nas1 mutant plants in response to different metal treatments
Nine-week-old wild-type (WT, black) and nas1 mutant (white) plants exposed to different metal treatments: normal (10 µM ZnSO₄, Zn10), high Zn (1000 µM ZnSO₄, Zn1000) and high Cd (50 µM CdSO₄, Cd50). NA concentrations (µM) in the shoots (A) and roots (B) are shown as mean ± SE. Letters shows differences between treatments analyzed by two-way ANOVA followed by Tukey’s post-hoc test and a simple main effect test.
Based on the nucleotide information, the NAS1 protein in nas1 mutants is predicted as truncated and nonfunctional protein due to the transposon insertion, so the higher levels of NA (and more efficient metal translocation) in nas1 mutant plants must reflect functional complementation by other NAS genes. We therefore tested the expression of the NAS1, NAS2, NAS3 and NAS4 genes by quantitative RT-PCR. In WT plants, NAS4 was the most strongly expressed gene in the shoots whereas NAS2 was hardly expressed at all in this tissue, but NAS2 was the most strongly expressed gene in the roots in which NAS3 was hardly expressed (Fig. 7A, 7B). The overall impact of the NAS1 function was determined by summing the expression of all active NAS genes in each genotype, i.e. NAS1, NAS2, NAS3 and NAS4 in WT plants, but NAS2, NAS3 and NAS4 in nas1 mutants (Fig. 7C). We found that the sum of NAS gene expression was similar in both genotypes under normal conditions and high Cd conditions, and that the sum of NAS gene expression in the nas1 mutant was higher than WT in the presence of excess Zn, especially in shoots (Fig. 7C). Therefore, despite the lack of functional NAS1 enzyme in the nas1 mutant, the overall expression profile suggested functional compensation by the other NAS genes, mostly contributed by NAS3 and NAS4 in shoots (Fig. 7A, 7C) and by NAS2 and NAS4 in roots (Fig. 7B, 7C). We also noted that NAS1 was strongly expressed in the shoots in all treatments (but only at high Zn in the roots) of the nas1 mutant, despite producing a non-functional protein (Fig, 7A). This suggests that NAS1 plays an essential role in the shoots, and that the absence of the functional protein will be the target for compensation by feedback regulation.

The nas1 mutant plants are better protected against herbivory

The observed selective advantage of the nas1 mutant in natural populations may be explained by the defense hypothesis, i.e. the higher metal concentrations present in the nas1 plants protect them against herbivory. We tested this hypothesis by carrying out choice and no-choice experiments using Small Cabbage White (Pieris rapae) caterpillars and butterflies, a cosmopolitic insect species.

Among the four P. rapae life stages (egg, larval/caterpillar, pupa and butterfly), the butterfly has the greatest movement and selection ability. One butterfly can travel more than 20 m before deciding where to lay eggs (Raworth, 2005). Accordingly, we used choice experiments to find out whether the butterflies showed any oviposition preference when presented with WT or nas1 plants. The results indicated no laying preference based on average egg numbers per plant (Fig.
Hence, the butterfly showed no evidence of discrimination between the two genotypes.

Although a *P. rapae* caterpillar cannot move as far as a butterfly, it can still travel several meters while foraging. We therefore carried out a choice experiment to determine whether caterpillars displayed any feeding preference when presented with WT or nas1 leaves. These experiments clearly showed that the caterpillars preferred the nas1 leaves under control conditions (Zn10) but this preference disappeared when the plants were exposed to excess metal treatments (Fig. 9).

Caterpillars feeding on WT and nas1 leaves were also compared by setting up a no-choice experiment, in which five L1 caterpillars were placed onto the leaves of WT or nas1 plants grown under the four different conditions. After 15 days, it was already clear that caterpillar growth was affected by both the treatments and the plant genotype as interacting factors. Caterpillars feeding on WT plants exposed to excess Zn or Cd were smaller than those feeding on WT plants under normal conditions (Fig. 10A), while the effects of excess Ni were lethal (data not shown). Although there was no significant difference between the caterpillars feeding on WT and mutant leaves under normal conditions, those fed on mutant leaves exposed to excess Zn or Cd were even smaller than those feeding on WT plants under similar conditions (Fig. 10B) suggesting that feeding on the mutant genotype amplified the effects of metal toxicity.

The experiment was had a look in more detail by considering developmental stages L1 to L5 in parallel exposures (Fig. 10C-E). There was no difference in development between WT and nas1 leaves under normal conditions and most had reached stage L5 after 15 days. But when plants were exposed to excess Zn, 85% of caterpillars feeding on WT leaves were still in stage L4 after 15 days with only 6% progressing to L5, whereas only 60% of caterpillars feeding on the nas1 mutant leaves had reached stage L4 by the same time, with 30% still in stage L3. Similar results were observed when caterpillars were fed on leaves exposed to excess Cd: 53% of caterpillars feeding on WT had reached stage L4, 32% were moulting to stage L5 and 14% were in L5, whereas 71% of caterpillars feeding on the nas1 mutant leaves were still in stage L4 and 10% were still at stage L3. The presence of high metal concentrations therefore caused developmental delay but this was exacerbated in the case of the nas1 mutant.
Fig. 7. NAS gene expression in WT and nas1 mutant plants in response to different metal treatments

Nine-week-old wild-type (WT, black) and nas1 mutant (white) plants exposed to different metal treatments: normal (10 µM ZnSO₄, Zn10), high Zn (1000 µM ZnSO₄, Zn1000), high Cd (50 µM CdSO₄, Cd50) and high Ni (100µM NiSO₄, Ni100). Combined NAS1-4 gene expression levels were measured in shoots and roots by real-time RT-PCR and are shown as RTL (relative transcript levels) ± SE. The individual gene expression levels are also shown in shoots (A) and roots (B), and the combined gene expression in (C). NAS1 levels in the WT shoot under Zn10 treatment is used as the calibrator (RTL = 1). Letters shows differences between treatments analyzed by two-way ANOVA followed by Tukey’s post-hoc test and a simple main effect test.

Fig. 8. Choice experiments with butterflies in response to different metal treated N. caerulescens plants

Wild-type (WT, black) and nas1 mutant (white) plants exposed to different metal treatments: normal (10 µM ZnSO₄, Zn10), high Zn (1000 µM ZnSO₄, Zn1000), high Cd (50 µM CdSO₄, Cd50) and high Ni (100µM NiSO₄, Ni100) were used for choice experiments with Pieris rapa butterflies. Average egg numbers in each plant were calculated from 12 repeats and shown as mean ± SE. There was no significant difference between WT and nas1 mutant plants for any of the treatments.
Fig. 9. Choice experiments with caterpillars in response to different metal treated *N. caerulescens* plants

Wild-type (WT, black) and *nas1* mutant (white) plants exposed to different metal treatments: normal (10 µM ZnSO$_4$, Zn10), high Zn (1000 µM ZnSO$_4$, Zn100), high Cd (50 µM CdSO$_4$, Cd50) and high Ni (100 µM NiSO$_4$, Ni100) were used for choice experiments with *Pieris rapa* L4 caterpillars for 27 hours. Average leaf area consumed by caterpillars was calculated in mm$^2$ (A) and as a percentage of total leaf area (B). Values are mean ± SE. Photos of remaining leaves (C). Letters shows differences between treatments analyzed by two-way ANOVA followed by Tukey’s post-hoc test and a simple main effect test.
Fig. 10. Weights and stages of caterpillars in non-choice experiments in response to different metal-treated *N. caerulescens* plants  L1 caterpillars were fed on wild type (WT, black) or nas1 mutant (white) plants treated with 10 µM ZnSO$_4$ (Zn10), 1000 µM ZnSO$_4$ (Zn1000) and 50 µM CdSO$_4$ (Cd50). Photos (A) and fresh weights (B) of caterpillars were taken after 15 days. Weights are shown in mg as mean ± SE. Percentages of caterpillars at (L1 to L5) under different treatments, Zn10 (C), Zn1000 (D), and Cd50 (E) are also identified after 15 days. Letters show differences between treatments analyzed by two-way ANOVA followed by Tukey’s post-hoc test and a simple main effect test.
Chapter 4

*N. caerulescens* plants before and after the no-choice experiments are shown in Fig. 11A. The dry weights of the *nas1* roots and shoots were always lower than those of the wild type plants (Fig 11B), as seen before (Fig 2).

A.

<table>
<thead>
<tr>
<th></th>
<th>Zn10 Before</th>
<th>Zn10 After</th>
<th>Zn1000 Before</th>
<th>Zn1000 After</th>
<th>Cd50 Before</th>
<th>Cd50 After</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>nas1</td>
<td>WT</td>
<td>nas1</td>
<td>WT</td>
<td>nas1</td>
<td>WT</td>
</tr>
</tbody>
</table>

![Images of plants before and after experiments](image)

B.

![Bar chart of shoot dry weight](image)

C.

![Bar chart of root dry weight](image)

**Fig. 11. N. caerulescens plants before and after non-choice experiments**

Wild type (WT) and *nas1* mutants (nn) were grown in ½ Hoagland’s solution containing low metal levels (10 µM ZnSO₄, 100 µM ZnSO₄ and 5 µM CdSO₄) for 3 weeks, and for another 3 weeks exposed to high metal concentration: normal Zn (10 µM ZnSO₄, Zn10), high Zn (1000 µM ZnSO₄, Zn1000), and high Cd (50 µM CdSO₄, Cd50). Five L1 caterpillars are applied to the plants for 15 days. The photos of plants (A) were taken before and after the caterpillar experiments. The shoot (B, left) and root (C, right) dry weights of wild-type (WT, black) and *nas1* mutant (white) plants after 15 days non-choice experiments were also measured in mg and shown as mean ± SE. Letters shows differences between treatments analyzed by two-way ANOVA followed by Tukey’s post-hoc test and a simple main effect test.
Heavy metals can influence herbivores directly (i.e. the metals are consumed by the caterpillars and exert direct toxic effects) or indirectly (i.e. they change the nutrient composition of the host plants and affect caterpillar nutrition, e.g. by reducing the amount of protein available). We therefore measured the protein concentrations of WT and nas1 mutant shoots using the Dumas method. This indicated that the protein concentration increased when the plants were exposed to excess Zn and Cd compared to plants growing under normal conditions, and that the mutant plants had a higher protein content than WT plants exposed to excess Zn and Cd (Fig. 12).

![Protein content graph](image)

**Fig. 12. Protein concentration in WT and nas1 mutant plants in response to different metal treatments**

Nine-week-old wild-type (WT, black) and nas1 mutant (white) plants exposed to different metal treatments: normal (10 µM ZnSO₄, Zn10), high Zn (1000 µM ZnSO₄, Zn1000), high Cd (50 µM CdSO₄, Cd50) and high Ni (100µM NiSO₄, Ni100) were collected for the DOMUS analysis. Protein concentrations in dried shoots are shown as mean ± SE. Letters shows differences between treatments analyzed by two-way ANOVA followed by Tukey’s post-hoc test and a simple main effect test.
DISCUSSION

Metal bioavailability and transportation to plant cells is regulated by transcription factors and the interaction between metal transporters and chelators. Such chelators are generally small molecules like nicotianamine (NA), which is synthesized by the enzyme nicotianamine synthase (NAS). At least four NAS genes are expressed in the hyperaccumulator *N. caerulescens* (van de Mortel *et al.*, 2006). However, we identified a *nas1* mutant allele caused by a transposon insertion, and the allele was found to be prevalent in several natural *N. caerulescens* populations from the same region in Belgium.

The potential fitness advantage of this mutated *nas1* allele was initially investigated by phenotypic analysis in plants exposed to excess metals. We found that the *nas1* mutant was more sensitive to Zn, Cd and Ni than WT plants, thus conferring a selective disadvantage. Although NAS1 is functionally incomplete in the *nas1* mutant due to a premature stop codon at position 228, the loss of functional NAS1 protein is balanced by higher levels of *NAS2, NAS3* and *NAS4* expression, suggesting these achieve functional compensation to boost the levels of NA in the roots. This improves the ability of the mutant plants to translocate Zn/Cd and indeed they are more efficient than WT plants under the same conditions. This improved translocation efficiency may explain the delayed growth and development of *Pieris rapae* caterpillars, which is more severe (and in some cases lethal) when they feed on *nas1* mutant leaves compared to WT leaves. This suggests the *nas1* mutant may confer an adaptive advantage (Fig 13).

The presence of a transposon insertion in the *NAS1* gene of certain *N. caerulescens* populations raises questions about its origin and whether its potential selective advantage can be inferred from its frequency. The presence of the *nas1* allele in Plombières (PB), La Calamine (LC) and Prayon (PR) and its absence elsewhere suggests either a recent origin or that the allele is only advantageous in three neighboring sites around Belgium. However, the allele frequency differs significantly among these sites: PB (0.59), LC (0.55) and PR (0.91). This implies the populations may be exposed to different selective pressures or represent different stages on the way toward allelic fixation (Babbitt *et al.*, 2010, Tian *et al.*, 2002), or potentially reflect a founder effect, e.g. when the PR population is established from the PB or LC populations, but by chance with predominantly *nas1*
Fig. 13. Defense mechanism of WT and nas1 mutants

Although the NAS1 protein is truncated in the nas1 mutant, enhanced NAS3 and NAS4 expression in the shoot and NAS2 expression in roots complements the production loss of nicotianamine by the disfunctional NAS1 protein, which is produced at higher levels than in WT plants at high Zn exposure. This has a positive effect on metal translocation from root to shoot in the mutant compared to the WT, and the presence of excess heavy metals causes reduced growth and development (sometimes lethal) in caterpillars.
plants. Genetic polymorphisms are often maintained in a population if the heterozygote confers a selective advantage (overdominance), if there is frequency-dependent fluctuation in selection or in cases of pleiotropy (Andrés et al., 2010). If the genetic variation can be resolved to a single nucleotide polymorphism (SNP) within or linked to the allele, this can be used to determine the signature of selection (Nielsen, 2005). For example, if the advantageous genetic diversity in and/or nearby the *NAS1* locus was maintained in the population, such a locus would experience balancing selection, which promotes genetic and phenotypic variability (Tian et al., 2002). In contrast, if variation in the regions neighboring the *NAS1* locus was reduced, as in the case of positive and purifying selection, only the fittest allele would be maintained in the population (Kim&Stephan, 2002).

In the case of *NAS1* we favor the hypothesis that the locus is under positive, purifying selection. We have tested the phenotypes of *NAS1/NAS1* and *NAS1/nas1* plants and found no phenotypic difference for metal related traits between these two genotypes, thus making the overdominance theory a less likely reason for the high *nas1* allele frequency. Although the population at PR is of later origin than those of PB or LC (W.H.O. Ernst, pers. communication), the high *nas1* allele frequency in the PR population is not likely to be due to a founder effect, with so few *nas1/nas1* and especially *NAS1/nas1* plants still present (Fig 1). That leaves a positive selection for the *nas1* allele as the most likely explanation for the high allele frequency, and it would mean that in all three populations, the *nas1* allele is progressing towards a selective sweep.

The most convincing case of inorganic defense was found in the studies of Ni and Se (Boyd, 2007). We also witnessed the strong and lethal effect of elevated Ni levels in *N. caerulescens* plants on Small Cabbage White caterpillars. Since only few studies have considered the effects of Zn and Cd, with different results, the potential inorganic defense role of Zn in hyperaccumulators remains controversial. Our results and earlier reports (Behmer et al., 2005, Jhee et al., 1999, Pollard&Baker, 1997) indicate that Zn plays a role in inorganic defense. However, other studies have provided no support for this hypothesis (Huitson&Macnair, 2003, Noret et al., 2005, Noret et al., 2007). The potential inorganic defense role of Zn may depend on the plant species and its natural enemies. *N. caerulescens* displays this defensive effect against different butterflies (leaf chewers), such as *Pieris napi oleracea* Harris (Jhee et al., 1999), *P. brassicae* (Pollard&Baker, 1997) and *P. rapae* (in this study). We clearly found that increasing the concentration of Zn or Cd in *N. caerulescens* leaves causes delayed caterpillar growth and
development, and the effect is exacerbated in the *nas1* mutant. The effect of Cd is more likely to be influenced by the mechanism of herbivory, e.g. chewing insects tend to cause more severe damage to plants than sucking ones (Konopka *et al.*, 2013), and the herbivore’s own tolerance to heavy metal exposure.

A potential defense mechanism incorporating our data and based on the defensive enrichment hypothesis described by Boyd (2012) is shown in Fig. 14. The proposed threshold protective benefit concentration (TPBC) of Zn was initially below 1000 mg kg\(^{-1}\) dry weight. The TPBC corresponding to the evolutionarily effective defense threshold (EEDT), where natural selection begins to enhance the accumulation of heavy metals, will need to be determined in further experiments. Based on this trend, we propose that the *nas1* mutant confers greater protection than WT plants at a higher metal concentration. The Zn concentration in the shoots of WT plants and *nas1* mutants from Plombières, Belgium was 15,000–20,000 mg kg\(^{-1}\) dry weight, so the *nas1* mutant may be fitter in this environment, explaining its selective advantage.

Although the *nas1* mutant has an advantage, because it is protected from herbivores, its small size and sensitivity to heavy metals (due to the accumulation of higher levels of metal in the leaves) means that local conditions will determine whether the allele confers a fitness advantage or penalty. This may explain why we still find WT and *nas1* mutants coexisting in La Calamine and Plombières. These sites are older than at Prayon and therefore more diverse in terms of soil metal concentrations. Thus, selection there may not go very quickly, as the selective advantage for *nas1* mutant plants compared to WTs will be much less prominent on sites with less extreme Zn and Cd levels. There, the higher metal accumulation of *nas1* will not outweigh the smaller plant size and higher nutritional value (higher protein concentration) in comparison to WT plants. However with decreasing metal exposure, also WT plants will be likely to experience an increasing selective disadvantage compared to other plant species less tolerant to metal exposure, which can also be observed at La Calamine and Plombières. It will be very interesting to monitor these three populations over the next decades to see how a single mutant allele can maintain its selective advantage in nature by regulating the metal homeostasis mechanism of the whole plant.
Fig. 14. The Defensive Enhancement Hypothesis for NASI

The Zn concentration (x-axis) of WT plants (■) and nasl mutants (■) from control (10 μM ZnSO₄), high Cd (50 μM CdSO₄ and 10 μM ZnSO₄), and high Zn (1 mM ZnSO₄) treatments were plotted against the Protective Benefit (y-axis) from the formula K x Zn concentration (mg kg⁻¹ plant dry weight)/(caterpillar fresh weight x surviving percentage of caterpillars), with K representing a constant coefficient. The logarithmic trend line describes the relationship between Zn concentration and the potential evolutionary benefit. The Zn hyperaccumulation threshold was set at 3000 mg kg⁻¹ plant dry weight (Ent et al., 2013). The threshold protective benefit concentration (TPBC), which corresponds to a certain evolutionarily effective defense threshold (EEDT), is probably 0–800 mg kg⁻¹ dry weight. Figure based on (Boyd, 2012).

In terms of evolutionary and ecological impact, heavy metal hyperaccumulation has evolved many times and is involved in the elemental defense of plants. The evolution of metal-dependent defense mechanisms can even be achieved at lower metal concentrations (Cheruiyot et al., 2013). A possible scenario for the evolution of hyperaccumulation was recently proposed by (Hörger et al., 2013) involving selection pressure from pathogens or herbivores that had evolved resistance, in association with gene flow from non-metallicolous populations to increase the possibility that a hyperaccumulation trait would become established in N. caerulescens. Therefore, whether nasl mutants that promote metal accumulation and protection from herbivores remain stable in the identified N. caerulescens populations will reflect gene-flow with N. caerulescens accessions from non-metalliferous sites and the selection pressure to maintain the allele under different environmental conditions.
ACKNOWLEDGEMENTS

The authors acknowledge Niels van der Last, Annemiek Stegehuis, Diede Maas, Allison Luger, Yanli Wang and Zeshan Hassan for collecting *N. caerulescens* samples from Belgium and conducting the genotyping experiments; Andrea Saccani, Giovanna Visioli, Andrea Pirondini, Bettine Aigner, Radim Vasut, Maarten Koornneef and Paula Pongrac for providing the seeds of the different populations and species; Wilco Ligterink and Bas Dekkers for assistance with qRT-PCR; Peter de Gijsel and Liya Yi for help with nitrogen measurements; Corrie Hanhart for help with *N. caerulescens* seed propagation; and Maarten Koornneef and Bas Zwaan for their comments and suggestions on the manuscript. Funding was provided by the Graduate School Experimental Plant Sciences and the EU COST Action FA0905 on *Mineral-improved crop production for healthy food and feed.*
**Appendix 1**

**NASI sequences including transposon insert**

```plaintext
atggcttgcaacaacaatatggtgttgaagcgaatcatcgacttacccaagctcagaggttt
aaacactttccaaaaatgtgcagacattggtgctgggcaactctggtccacggttcgtctacccaagggactatc
gagatgctcagcagatgtctgaagaagtaataatctacatcagcagctgtgtggatttggtgaagcagaa
gttatatgtgagcaacaactctccacaatcctgggtcttttacaagagaaagcgaaccgcacacttgacacttttac
catcattactacagaacactacactaagcagttgagttgtgttggtgttggttgaagcagaa
```

___: present the NASI cDNA

☐: 6-bp target site duplication (TSD) ACACAT

☐: two 118-bp long terminal repeats (LTRs)

____: present the stop codon in the nasI::tp allele
Appendix 2

NAS1 and transposon-inserted NAS1 predicted protein sequences

NAS1 predicted protein (322 amino acids)
MACNNNFVVKRIIDLYDQISKLESKLPSKNVDTLFGQLVSTCLPTDTNIDVT
KMSEEVKDMRSNLKLCGEAEGYLEQHFSTILGSQEDGNPLDHIFPYY
DNYLKLKIEFDLSSQHTTHVPTKIAFVGSGPMPLTSIVLAKFHPNTTFHN
FDIDSHANTLASLVSRLPSKRMIHFHTTDVLNAGELDQYDVFLALV
GMDKEAKVRAIEHLEKHMAPGAVLMLRSAHALRAFLYPIVDSDDLKGF
QLLTIIHPTDDVVNSVVIARKLGSTTAGINGTRGMFCNCSKVAIHN
NNRCKKKKMIEEFAIIE*

transposon-inserted NAS1 predicted protein (227 amino acids)
MACNNNFVVKRIIDLYDQISKLESKLPSKNVDTLFGQLVSTCLPTDTNIDVT
KMSEEVKDMRSNLKLCGEAEGYLEQHFSTILGSQEDGNPLDHIFPYY
DNYLKLKIEFDLSSQHTTHVPTKIAFVGSGPMPLTSIVLAKFHPNTTFHN
FDIDSHANTLASLVSRLPSKRMIHFHTTDVLNAGELDQYDVFLALV
GMDKEAKVRAIEHLEKHIIEI*
Chapter 5

A comprehensive set of transcript sequences of the heavy metal hyperaccumulator *Noccaea caerulescens*

Ya-Fen Lin¹#, Edouard Severing¹,²,#, Bas te Lintel Hekkert³, Elio Schijlen³, and Mark G.M. Aarts¹

¹ Laboratory of Genetics and
² Laboratory of Bioinformatics, Wageningen University, Wageningen, The Netherlands;

#These authors contributed equally to this work.
ABSTRACT

Noccaea caerulescens is an extremophile plant species belonging to the Brassicaceae family. It has adapted to grow on soils containing high, normally toxic, concentrations of metals such as nickel, zinc and cadmium. Next to being extremely tolerant to these metals, it is one of the few species known to hyperaccumulate these metals to extremely high concentrations in their aboveground biomass. In order to provide additional molecular resources for this model metal hyperaccumulator species to study and understand the mechanism of heavy metal exposure adaptation, we aimed to provide a comprehensive database of transcript sequences for N. caerulescens. In this study, 23830 transcript sequences (isotigs) with an average length of 1025 bps were determined for roots, shoots and inflorescences of N. caerulescens accession ‘Ganges’ by Roche GS-FLEX 454 pyrosequencing. These isotigs were grouped into 20,378 isogroups, representing potential genes. This is a large expansion of the existing N. caerulescens transcriptome set consisting of 3705 unigenes. When compared to a Brassicaceae proteome set, 22,232 (93.2%) of the N. caerulescens isotigs (corresponding to 19191 isogroups) had a significant match and could be annotated accordingly. Of the remaining sequences, 98 isotigs resembled non-plant sequences and 1386 had no significant similarity to any sequence in the GenBank database. Among the annotated set there were many isotigs with similarity to metal homeostasis genes or genes for glucosinolate biosynthesis. Only for transcripts similar to Metallothionein3 (MT3), clear evidence for an additional copy was found. This comprehensive set of transcripts is expected to further contribute to the discovery of mechanisms used by N. caerulescens to adapt to heavy metal exposure.

KEY WORDS
Metal hyperaccumulation, metal hypertolerance, phytoremediation, zinc, cadmium, gene expression, Brassicaceae
INTRODUCTION

*Noccaea caerulescens* (J. & C. Presl) F.K. Mey., formerly named *Thlaspi caerulescens*, is an outstanding model plant species to study heavy metal hyperaccumulation (Assunção *et al.*, 2003, Milner&Kochian, 2008, Peer *et al.*, 2003). It is one of the few plant species of which genotypes are known that are adapted to grow on soil containing high levels of zinc (Zn), cadmium (Cd), nickel (Ni) and/or lead (Pb)(Mohtadi *et al.*, 2012). Not only are these genotypes extremely tolerant to the different heavy metals they are exposed to, *N. caerulescens* can also hyperaccumulate these metals to high concentrations in shoots. Accumulations of Zn to 30,000 μg g⁻¹, Cd to 2700 μg g⁻¹, and Ni to 4000 μg g⁻¹ levels are reported on a shoot dry weight (dw) base. These are two orders of magnitude higher than other, non-accumulating, species generally accumulate (Brown *et al.*, 1995, Eijsackers *et al.*, 1993, Lombi *et al.*, 2000, Reeves&Brooks, 1983). Next to *N. caerulescens*, *Arabidopsis halleri* is developed as model metal hyperaccumulator (Meyer&Verbruggen, 2012). This species is also hypertolerant to Zn and Cd, and a strong Zn hyperaccumulator, but less of a Cd hyperaccumulator and not known to be adapted to Ni. Unlike *N. caerulescens* it is self-incompatible, which complicates genetic analysis, but it is much closer to the general plant model species *Arabidopsis thaliana*, with which is shares more sequence synteny. Adaptation to heavy metals seems to be more commonly occurring in the Brassicaceae family than in other families, with other *Noccaea* species known to hyperaccumulate Ni and Zn, and one, *N. praecox*, also hyperaccumulating Cd, as recently reviewed by Koch&German (2013). Other genera with metal hyperaccumulating species are *Alyssum* and *Streptanthus* (both mostly Ni adapted). Thus, it is fortunate that the first plant species for which genomic data became available, *A. thaliana*, belongs to the Brassicaceae family, as this triggered further interest in generating genome sequence information from several members of this family (http://www.brassica.info/resource/sequencing/bmap.php). For neither *A. halleri* nor *N. caerulescens* the genome sequence has been determined, which is why most of the gene expression research on these species so far has relied on heterologous micro-array analysis using the available *A. thaliana* micro-arrays (Becher *et al.*, 2004, Hammond *et al.*, 2006, Talke *et al.*, 2006, van de Mortel *et al.*, 2006, van de Mortel *et al.*, 2008, Weber *et al.*, 2004). This revealed that both species seem to have evolved similar strategies for dealing with the high metal exposure, typically by modifying the expression of several genes normally involved in Zn and Fe.
mineral homeostasis. A striking example is the copy number expansion of the HMA4 gene, observed in both species, which increased expression of the gene when compared to non-accumulating species (Craciu et al., 2012, Hanikenne et al., 2008, Ó Lochlainn et al., 2011).

One reason to investigate the remarkable metal adaptation properties of N. caerulescens is that these extremophile plants are interesting target species to develop for metal phytoextraction purposes, in which plants are used to remediate soils contaminated with toxic metals (Anjum, 2012, Peer et al., 2006). Another reason is that the rare extremophile nature of metal hyperaccumulators, to have adapted to otherwise hostile environments, makes them interesting models for plant evolutionary genomics studies (Hanikenne & Nouet, 2011). For both purposes, unravelling the evolutionary and physiological mechanisms that allowed their adaptation, at the molecular level, will be needed. The main approaches that have been followed so far in identifying genes involved in metal adaptation involve comparisons of hyperaccumulating and related non-hyperaccumulating plant species, either by using genetic crosses, or by transcriptomics or proteomics comparisons (reviewed by HassanandAarts (2011), Krämer et al. (2007), LinandAarts (2012)). However, so far, comprehensive sets of genome and transcriptome sequences of metal hyperaccumulator species are not available, which is seriously limiting the progress in molecular analysis of metal adaptation.

Molecular analysis of N. caerulescens has been largely performed based on its close relationship to the well-known model species A. thaliana, allowing the use of molecular tools and genomic databases developed for this species. Although the lineages of both species separated probably some 20 Mya (Clauss & Koch, 2006), there still is substantial genome sequence similarity between both species, estimated at 87-88% sequence identity in intergenic transcribed spacer regions and 88.5% sequence identity in transcribed regions (Peer et al., 2003, Rigola et al., 2006). This high level of conservation was sufficient to use heterologous, A. thaliana, micro-arrays for comparative transcriptome analyses (Hammond et al., 2006, Plessl et al., 2005, Plessl et al., 2010, van de Mortel et al., 2006, van de Mortel et al., 2008). The first attempt to obtain sequence information of N. caerulescens was performed by Rigola et al. (2006), who generated an Expressed Sequence Tag (EST) database of little over 3700 transcript sequences. This resource has been used to generate a cDNA-based micro-array, which has been used for transcript profiling of N. caerulescens, but due to the limited number of probes on the array, the information that could be gathered was limited (Plessl et
al., 2010a). Since then, the rapid development of high-throughput sequencing techniques has made whole genome and transcriptome sequencing a lot more efficient and affordable (Morozova et al., 2009). The 454/Roche pyrosequencing method takes advantages of long read lengths and a fast running time (Metzker, 2010, Niedringhaus et al., 2011) and is a suitable method for whole transcriptome sequencing, when the main purpose is to obtain a comprehensive set of transcript sequences of substantial length to be used as a reference database for future transcriptome profiling studies.

Here we present such a comprehensive set of N. caerulescens transcripts, which largely exceeds the previous dataset in number of identified genes (Rigola et al., 2006). We identified many transcripts involved in mineral accumulation and homeostasis, which may be relevant for metal hyperaccumulation and hypertolerance. In addition we listed genes involved in the biosynthesis of glucosinolates. These are secondary metabolites conferring resistance to herbivores, especially prominent in Brassicaceae (Bones & Rossiter, 1996). This new transcript sequence information will facilitate the further analysis of the extremophile traits of N. caerulescens and is expected to contribute to similar studies in related metal hyperaccumulators, such as Arabidopsis halleri, and also less studied species like the Zn/Cd hyperaccumulator Noccaea praecox or the Ni-hyperaccumulator Noccaea goesingense (Koch & German, 2013). It will also contribute to more efficient functional studies of heavy metal related genes, to establish their role in metal adaptation or for possible applications in metal phytoextraction, or genes related to synthesis of glucosinolates.
MATERIAL AND METHODS

Plant materials and RNA preparation

An inbred line of *N. caeruleus* accessions ‘Ganges’ (kindly obtained from Dr. Henk Schat, Free University, Amsterdam, The Netherlands) was used. Roots and shoots were collected separately from five-week-old plants, grown in half-strength Hoagland’s solutions (Assunção *et al.*., 2001) containing 10 µM ZnSO₄, in a climate controlled growth chamber (set at 20°C/15°C day/night temperature; 70% relative humidity; 12 hours day time). Inflorescences were collected from 22-week-old plants grown in soil. RNA of these three plant parts was extracted separately by using the RNeasy® Plant Mini kit (Qiagen, cat. no. 74904) with on-column RNase-Free DNase set digestions (Qiagen, cat. no. 79254). The RNA concentrations were quantified using the highly selective Qubit™ RNA BR Assay kit (Invitrogen™, cat. no. Q10210) with a Qubit® 2.0 Fluorometer.

cDNA library construction

Similar amounts of RNA from roots, shoots, and flowers were pooled and used for preparation of a normalized and random-primed cDNA library for Roche/454 sequencing (Vertis Biotechnologie AG). From the total RNA sample, poly(A)+ RNA was isolated, and used for cDNA synthesis. First strand cDNA synthesis was primed using random hexamer primers. To the 5’ and 3’ ends of the cDNA, Roche/454 adapters A and B, as provided by the manufacturer, were ligated and cDNA was finally amplified using 12 PCR cycles and proofreading DNA polymerase. Normalization was performed by denaturation and re-association of cDNA. Re-associated double strand cDNA was separated from remaining (normalized) single strand cDNA (ss-cDNA) over a hydroxylapatite column. After separation, ss-DNA was PCR-amplified using six PCR cycles. Finally cDNA in the size range of 500-850 bp was eluted from a preparative agarose gel. The final normalized cDNA library contains double stranded fragments of between 500-850 bp, consisting of the following sequence structure: 5’-454-Adapter A (CCA-TCT-CAT-CCC-TGC-GTG-TCT-CCG-ACT-CAG), 5’- barcode (CACACG), 5’ adapter (GAC-CTT-GGC-TGT-CAC-TCA-GTT), cDNA insert (400-750bp), 3’ adapter (TCG-CAG-TGA-GTG-ACA-GGC-CA), 3’-454-Adapter B (CTG-AGA-CTG-CCA-AGG-CAC-ACA-GGC-GAT-AGG).
**De novo sequencing and assembly of *N. caerulescens* transcript sequences**

Prior to Roche/454 sequencing, cDNA library molecules were clonally amplified using a 'two copies per bead' ratio for one Large Volume Emulsion PCR, following the manufacturer’s protocol (Roche, Genome Sequencer FLX Titanium Series). Two million DNA carrying beads (from a 23% enrichment) were loaded on half a picotiter plate equivalent, divided over two regions. Sequencing was done on a 454 GS FLX Titanium instrument using XLR70 chemistry and 200 flow cycles. The raw reads were assembled after adaptor trimming using Newbler (454 life Sciences Corporation) version v 2.6.

**Sequence annotation and characterization**

The predicted proteomes and genome assemblies of *A. thaliana*, *Arabidopsis lyrata*, *Brassica rapa*, *Capsella rubella* and *Thellungiella halophila* (recently suggested to be named *Eutrema halophila* (Koch&German, 2013)) were downloaded from the phytozome repository version 9.1 (Goodstein et al., 2012). The protein function annotation of *A. thaliana* version TAIR10.0 was obtained from the TAIR website (www.arabidopsis.org). The gene ontology annotation for *A. thaliana* was downloaded from the gene ontology webpage (www.geneontology.org). The protein annotations (including GO-terms) of the four other species were downloaded from the phytozome repository. Pathway and Enzyme Code annotations for *A. thaliana* and *B. rapa* were downloaded from the PlantCyc database (www.plantcyc.org) version 8.0. The annotated GO terms were summarized using the plant GOSlim Set of the GOSlim viewer at the AgBase website (http://www.agbase.msstate.edu/cgi-bin/tools/goslimviewer_select.pl) (McCarthy et al., 2006).

The predicted Brassicaceae proteins were clustered into groups of orthologous proteins by first performing all 25 possible pairwise similarity searches between the five predicted proteomes using BlastP version 2.26 (Altschul et al., 1990). Pairwise orthologs were determined for all ten possible species-pairs using Inparanoid version 4.1 (Remm et al., 2001). Finally, multi-species-ortholog clusters were built by feeding the Inparanoid output files to the multiparanoid Perl-script (Alexeyenko et al., 2006). Whenever possible, the annotation of orthologous protein clusters was inherited from *A. thaliana* or *B. rapa* members (in that order). Clusters without *A. thaliana* or *B. rapa* members remained un-annotated.
The transcriptome of *N. caerulescens* was first searched against the predicted proteomes using BlastX with an e-value cut-off of $10^{-5}$. Sequences for which the best hit was a member of an annotated cluster inherited the annotation from that cluster. The remaining sequences inherited their annotation from their best hit. Sequences without a significant match against the predicted Brassica proteomes were searched (BlastX with e-value cut-off $10^{-5}$) against the non-redundant protein database (NR) downloaded from the NCBI ftp-site (ftp://ftp.ncbi.nlm.nih.gov/). Proteins without BlastX hits were searched against the genomes of the Brassica species using BLAT (Kent, 2002) version 35. BLAT hits were filtered by requiring 90% of the *N. caerulescens* sequence to be aligned with an average identity of at least 85%. Finally sequences were searched against the non-redundant nucleotide database, downloaded from the NCBI ftp-site (BlastN with e-value cut-off $10^{-5}$). Blast hits were only accepted if the corresponding HSPs had $\geq 85\%$ sequence identity and covered $\geq 90\%$ of the query isoforms.

Protein alignments were constructed using ClustalW2 (Larkin *et al.*, 2007) and regions of low alignment quality were removed by hand using JalView (Waterhouse *et al.*, 2009). Maximum likelihood trees were constructed using PhyML (the following command line parameters were used: `-c 4 -m LG -o lr -v e -a e -f e –b 100`) (Guindon *et al.*, 2010). In brief, the trees were constructed using the LG amino acid substitution model (Le&Gascuel, 2008) with four relative substitution rate categories. The proportion of invariable sites and equilibrium amino-acid frequencies were estimated from the data. The alpha parameter was estimated by maximizing the likelihood of the phylogeny. Branching patterns were validated using bootstrap-analysis; 100 bootstrap-samples were generated for the ML-trees and 1000 for the NJ trees. Trees were displayed using the ETE2 python package (Huerta-Cepas *et al.*, 2010) and dendroscope (Huson&Scornavacca, 2012). To investigate the presence of gene duplications of metal homeostasis related genes, sequence sets were created for each gene family implicated in metal homeostasis by performing blast searches against the proteomes of the five Brassicaceae reference species (with e-value cut-off $10^{-5}$). The exonerate program was used to rapidly compare transcript sequences to genome sequences (Slater&Birney, 2005). Neighbour-joining trees were subsequently constructed for all alignments using ClustalW2. This Transcriptome Shotgun Assembly project has been deposited at DDBJ/EMBL/GenBank under the accession GASZ00000000. The version described in this paper is the first version, GASZ01000000. Raw sequencing data has been deposited in the NCBI short read archive under the accession SRX456668.
RESULTS

De novo sequencing and assembly of *N. caerulescens* isotigs

To maximize the transcript diversity, a pool of RNA from roots, shoots, and inflorescences of *N. caerulescens* was made. A cDNA library synthesized from this RNA pool was used for sequencing. In total 834,911 raw reads with an average length of 401 nucleotides (nt) were obtained and subsequently assembled using Newbler. During the assembly process, two initial contigs can be joined together into an isotig when raw reads are found that map to the ends of both contigs. The resulting transcriptome assembly (Table 1) consisted of 26,785 contigs the vast majority of which were further assembled into a set of 23,836 putative transcripts (isotigs) with an average sequence length of 1025 base pairs (bp). Isotigs sharing at least one contig are grouped into isogroups. Isotigs are likely to be transcript isoforms from the same gene (Isogroup). The isotigs in this study were grouped into 20,378 isogroups. In total 2326 isogroups (11.4%) consisted of more than one isotig. The distributions of the isotigs sizes and the number of isotigs per isogroup are provided in figure 1.

Table 1 Summary of the *N. caerulescens* transcriptome analysis.
The Newbler software was used to assemble raw sequence reads into isotigs (representing putative transcripts) and isogroups (representing putative genes). Quality value 40 (Q40) means the error rate is below 1 in 10,000 nucleotides.

<table>
<thead>
<tr>
<th>transcript assembly</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>total reads</td>
<td>834,911</td>
</tr>
<tr>
<td>total bases</td>
<td>3*10⁸</td>
</tr>
<tr>
<td>percentage aligned reads</td>
<td>86.24</td>
</tr>
<tr>
<td>No. assembled reads</td>
<td>638,451</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Isotigs</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. isotigs</td>
</tr>
<tr>
<td>No. bases</td>
</tr>
<tr>
<td>Average isotig size</td>
</tr>
<tr>
<td>N50 isotig size</td>
</tr>
<tr>
<td>Largest isotig size</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Isogroups</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. isogroups</td>
</tr>
<tr>
<td>Average no. isotigs/isogroup</td>
</tr>
<tr>
<td>Largest isotig count</td>
</tr>
<tr>
<td>No. isogroups with one isotig</td>
</tr>
<tr>
<td>Q40 Plus bases</td>
</tr>
</tbody>
</table>
Fig 1. Analysis of *N. caerulescens* isotigs and isogroups
The frequency distribution of isotigs according to sequence length (a), the frequency distribution of the number of isotigs (isotig count) per isogroup (b) are shown.

**Annotation and functional characterization of *N. caerulescens* genes**

*N. caerulescens* is a member of the Brassicaceae family, which is why we used the annotation of five other Brassicaceae species of which full genome sequence is available (*A. thaliana*, *A. lyrata*, *B. rapa*, *C. rubella* and *T. halophila*) as a reference for annotation. The predicted proteomes of these five reference species were first clustered in 24,172 groups of orthologous proteins, which we called the Brassicaceae proteome (see material and methods). In total 23,691 of these clusters were annotated following the *A. thaliana* or *B. rapa* annotation. We then compared the *N. caerulescens* isotig DNA sequences against the Brassicaceae proteome and listed the best match (Summarized in figure 2; full list in supplemental table S1). A total of 22,232 (93.2%) *N. caerulescens* isotigs (19191 isogroups) had a significant match with the Brassicaceae proteome and we annotated these sequences according to their best match.

Even though five closely related species were used as reference, still 1,604 isotigs were found without any significant hit in the Brassicaceae proteome data set. These isotigs were therefore compared with the NCBI non-redundant (NR) protein database. The search revealed that, although not detected in our initial search, 33 isotigs did have a significant match to a plant protein, including 16 isotigs that strongly resembled an *A. thaliana* protein and seven that resembled proteins from
Figure 2. Classification of *Noccaea caerulescens* isotigs according to best-hits

The number of best-hits of *N. caerulescens* isotigs to the Brassicaceae proteome (comprising those of *Thellungiella* (*Eutrema*) *halophila*, *Arabidopsis thaliana*, *Arabidopsis lyrata*, *Brassica rapa* and *Capsella rubella*) are listed. The 1604 isotigs without any significant hit to Brassicaceae proteome were further compared to the NCBI non-redundant protein database, identifying best-hit similarities to 33 proteins of plant origin and 98 of non-plant origin. The non-hit isotigs were further matched to the available genome sequences of the listed Brassicaceae reference species, identifying another 74 best-hit similarities. Finally, a set of 1399 isotigs remained for which no protein similarity could be found.

Other Brassicaceae (Supplemental table S2). 13 of the 33 isotigs corresponded to (retro)transposons and nearly all of the remaining ones encoded for hypothetical or unknown proteins. Next to these, an additional 98 isotigs showed only significant similarity to non-plant sequences (Supplemental table S3). Based on the high similarity to genes of organisms often found to be associated with plants (fungi, bacteria, viruses, etc.), these are most likely reflecting such associations, rather than actual *N. caerulescens* genes. The remaining sequences with no hit when compared to the NR protein database, were compared against the genome sequences of the five Brassicaceae reference species. This revealed an additional
74 sequences with significant DNA matches to one or more of the reference genomes. Still, there are 1399 isoligs that had no hit to Brassica proteins and proteins in the NR database (Supplemental table S4). When these were compared to the NR DNA database, an additional 13 isoligs were found to be similar to DNA entries in the database. Only two of these were plant genes. Five isoligs showed similarity to *A. thaliana* genomic sequences, one to plant mitochondrial DNA and the remaining five appeared to correspond to microbial sequences.

The list of isoligs with best-hit matches to the Brassicaceae proteome (Table S1) has been used to obtain an estimate of transcript length coverage of the *N. caerulescens* isoligs (Figure 3). This shows that around half of the transcripts cover over 80% of the protein sequence of their Brassicaceae best-hit match. In addition, we compared the full list of *N. caerulescens* isoligs to the previously obtained set of *N. caerulescens* ESTs (Rigola et al., 2006) using BlastN (≥ 95% identity and ≥ 60% coverage of shorter sequences). This retrieved 3773 of the 4289 sequences (88%). Sequence similarities between both datasets were generally around 98%, suggesting SNP frequencies of around 2 per 100 bp.

![Figure 3. Frequency distribution of *N. caerulescens* isoligs according to percentage coverage of their Brassicaceae proteome best-hit.](image)
Gene ontology (GO) annotation of *N. caerulescens* isogroups

Subsequently, we used the *N. caerulescens* isogroup sequences to categorize putative genes according to Gene Ontology (GO), by assigning them to three categories of GO terms: Biological Process, Cellular Compartment, and Molecular Function (see materials and methods) (Supplemental table S5). The annotated GO terms were further classified with the plant set of the GO Slim Viewer (McCarthy *et al.*, 2006) (summarized in Figure 4, full list in supplemental table S6).

Within Biological Process, two major categories are cellular process (25.1%) and metabolic process (19.3%), followed by developmental process (12.3%) and response to stress/stimulus (10.4%). Genes involved in cellular ion homeostasis (involving “cations” and “metal”, but also specifically zinc, copper, calcium, iron, manganese, potassium and phosphate) (0.1%), are likely to be crucial for the regulation of the heavy metal balance in *N. caerulescens* (Figure 4), in addition to genes involved in transport (2.7%), especially ion transport (0.5%) and signal transduction (1.6%). Within Cellular Compartment, the largest category is cytoplasmic component (50.7%), followed by membrane apparatus (12%), and then intracellular components (10.7%), extracellular components (2.1%) and cell wall (0.9%). Genes encoding for proteins localized to the plasma membrane (5%) and the vacuolar membrane (0.6%) will comprise metal transporter genes that are expected to be involved in metal uptake and sequestration (Figure 4). Within Molecular Function, the largest category is binding activity (37.1%) comprising ion binding (4%) as an important sub-category for metal hyperaccumulation traits. Furthermore, the genes categorized for transporter activity (8.1%) are expected to play an important role in metal uptake and metal transportation. To see if this is different from other species, we also compared *N. caerulescens* and *A. thaliana* regarding the 18-24 GO terms with the highest percentages of *N. caerulescens* isogroup counts (Figure 5). This made clear that there are only minor differences between both species.
Figure 4. Gene Ontology (GO) classification of *N. caerulescens* isogroups

*N. caerulescens* isogroups were classified into three major functional GO groups: Biological Process (a), Cellular Component (b) and Molecular Function (c) and subsequently sub-classified as indicated.
Fig 5. Comparative Gene Ontology (GO) classification of *N. caerulescens* and *A. thaliana*

*N. caerulescens* isogroups and *A. thaliana* genes were classified into three major functional groups: Cellular Component, Molecular Function, and Biological Process. Per group, the percentage of all isogroups or genes corresponding with the indicated GO-terms have been determined and the top 18 (Cellular Component) or top 24 (Biological Process and Molecular Function) of GO-terms are displayed.
A total of 3051 isotigs (2610 isogroups) could be assigned to 352 PlantCyc pathways (Supplemental table S7). The 26 most represented pathways (around 20 or more isogroups) are shown in Figure 6. The top six pathways are triacylglycerol degradation, homogalacturonan degradation, glycolysis II (from fructose-6P), tRNA charging, aerobic respiration (alternative oxidase pathway), and betanidin degradation, which have more than 50 isogroups found in the pathway.

Fig 6. Main biological pathways in which *Noccaea caerulescens* genes are involved. The top 26 pathways (A–X) are indicated in which *N. caerulescens* transcripts are found to be involved, upon comparison of the isogroup sequences to the PlantCyc database. A, triacylglycerol degradation; B, homogalacturonan degradation; C, glycolysis II (from fructose-6P); D, tRNA charging; E, aerobic respiration (alternative oxidase pathway); F, betanidin degradation; G, photosynthesis light reactions; H, glutathione-mediated detoxification II; I, phospholipases; J, TCA cycle variation V (plant); K, sucrose degradation III; L, fatty acid β-oxidation II (peroxisome); M, coumarin biosynthesis (via 2-coumarate); N, phosphate acquisition; O, tetrahydrofolate biosynthesis II; P, lactose degradation III; Q, photorespiration; R, CDP-diacylglycerol biosynthesis II; S, 3-phosphoinositide biosynthesis; T, D-myo-inositol (1,4,5)-trisphosphate biosynthesis; U, callose biosynthesis; V, phenylalanine degradation III; W, xylan biosynthesis; X, adenosine nucleotides de novo biosynthesis; Y, starch biosynthesis; Z, trehalose biosynthesis I
Transcripts related to the response to heavy metal exposure

The metal hyperaccumulation and hypertolerance properties of *N. caerulescens* partly rely on a number of genes known in other species, mainly *A. thaliana*, to be involved in mineral homeostasis. A list of such genes, including 87 isogroups, is shown in Table 1. The list includes metal transporter genes belonging to the ZRT/IRT-like Protein (ZIP) gene family (9 isogroups), the Natural Resistance-Associated Macrophage Protein (NRAMP) family (8 isogroups), the Heavy Metal ATPase (HMA) family (9 isogroups), the Metal Tolerance Protein (MTP) family (7 isogroups), and the Calcium exchanger (CAX) family (9 isogroups). In addition, we listed genes belonging to the Plant Cadmium Resistance (PCR) gene family (2 isogroups), the Pleiotropic Drug Resistance protein (PDR) family (14 isogroups) and Plant Defensin (PDF) family (8 isogroups). Furthermore, transcripts related to metal chelator and metal chelator transporter functions were listed, such as genes of the Nicotianamine Synthase (NAS) family (4 isogroups), the Phytochelatin Synthase (PCS) family, (2 isogroups), the Metallothionein (MT) family (4 isogroups), the Yellow Stripe Like protein (YSL) family, (6 isogroups), the Zinc Induced Facilitator (ZIF) family (5 isogroups), and the Multidrug Resistance-associated Protein/ ATP-binding cassette transporter ABCC type (MRP/ABCC) family (26 isogroups). N.B. we did not include the large families of genes encoding for heavy metal-associated isoprenylated plant proteins (HIPP) or heavy metal-associated plant proteins (HPP), of which some members are recently suggested to be relevant for response to Cd (de Abreu-Neto et al., 2013).

For two larger gene families, the ZIP and MTP families, we performed a phylogenetic analysis (Figure 7a, b) to confirm that the Blast analysis (Table 2) indeed identified the most likely *A. thaliana* orthologs. Of the ZIP family, no transcripts were found for orthologs of *AtIRT2, AtZIP3, AtZIP5, AtZIP7, AtZIP8* and *AtZIP12*. For the MTP family, we did not find *N. caerulescens* orthologs of the *AtMTP2, AtMTP3, AtMTP4, AtMTP9* and *AtMTP10* genes. Next we searched for *N. caerulescens* specific gene duplications for all gene classes in table 2. To this end we created for each class multiple sequences alignments with sequences from the five reference species that were similar to the *N. caerulescens* isogroup sequences (see material and methods). Neighbour-joining trees were constructed from these alignments and manually inspected for possible putative *N. caerulescens* specific duplications. Only in the tree of the MT3 class we identified a putative *N. caerulescens* specific duplication. In order to obtain additional support for a recent gene duplication we identified the MT3 orthologs of the five Brassicaceae
reference species and constructed a phylogenetic tree of them, including the MT3 gene of out-group Vitis vinifera (Figure 7c). This confirmed that all inspected genomes had only one copy of the MT3 gene.

Figure 7. Phylogenetic comparison of ZIP, MTP and MT gene related N. caerulescens isotigs.

Maximum Likelihood trees, sowing the longest isotigs (indicated in red) belonging to different isogroups of the ZRT/IRT-like Protein (ZIP)(a) and Metal Tolerance Protein (MTP) gene families (b) as identified in Table 2, and compared on amino acid sequence level to A. thaliana genes to indicate the most likely orthologs. When all proteins listed in Table 2 were manually examined for N. caerulescens specific gene duplications, this appeared to be the case only for Metallothionein 3 (MT3). The Neighbour Joining tree is shown in (c), displaying the comparison of MT3 isotigs to most similar genes of the previously used five Brassicaceae reference gene sets, indicates that N. caerulescens expresses an additional gene copy. Note that all branches in the Maximum Likelihood and Neighbor Joining trees with 50% or less bootstrap support were collapsed.
Next to the mineral homeostasis genes, we also examined the occurrence of transcripts representing genes in the glucosinolate biosynthesis pathway. This pathway is generally prominent in Brassicaceae as it generates glucosinolates that upon interaction with myrosinase enzymes release toxic isothiocyanates, thiocyanates or nitrils, which display strong anti-feeding properties against herbivores (Bones & Rossiter, 1996). Most of the genes in this pathway and their roles have been identified (Sønderby et al., 2010b, Wang et al., 2011). Important genes are a series of MYB and MYC genes encoding for the transcription factors controlling aliphatic and indolic glucosinolate biosynthesis (Gigolashvili et al., 2007, Schweizer et al., 2013, Sønderby et al., 2010a), next to genes involved in core structure formation, side-chain elongation, secondary modification and co-substrate pathways (as summarized by Sønderby et al. (2010b)) and more recently identified genes involved in secondary modification of indolic glucosinolates and genes encoding glucosinolate transporters (Nour-Eldin et al., 2012, Pfalz et al., 2011). N. caerulescens plants originating from metalliferous soils were found to contain lower levels of glucosinolates than plants originating from non-metalliferous soils (Noret et al., 2007), which is why we were interested to see if the glucosinolate genes known from A. thaliana were also present in the N. caerulescens transcriptome. Supplemental table S8 shows the list of genes involved in the glucosinolate biosynthesis pathway, which we classified based on their presumed function as transcription factor genes, genes involved in side-chain elongation, core structure formation and secondary modification or genes of a co-substrate pathway (according to Wang et al. (2011)). Of the 61 genes in the list, 40 were found to be expressed in the N. caerulescens transcriptome.
DISCUSSION

*N. caerulescens* is one of the two heavy metal hyperaccumulator plant model species, together with *A. halleri*, that are studied in detail to understand their adaptation to growing on soil containing extremely high concentrations of Zn and Cd (Verbruggen et al., 2013). As part of this adaptation, they not only tolerate exposure to normally lethal concentrations of both metals, but also hyperaccumulate them in their shoots, probably as protection against herbivory and microbial infection (Hoerger et al., 2013). *N. caerulescens* is the only species to also hyperaccumulates Ni and Pb.

The development of new and cost-effective gene expression analysis methods such as RNA-Seq prompted us to generate a comprehensive reference transcriptome dataset for *N. caerulescens* that will facilitate and support such gene expression analyses, without having to rely on heterologous comparisons. This data set replaces a previous data set we made, comprising only 4289 transcript sequences, representing 3709 unigenes, of accession ‘La Calamine’ (Rigola et al., 2006). We now assembled 23,836 isotigs, further condensed into 20,378 isogroups of accession ‘Ganges’, which is a better Cd-hyperaccumulator than ‘La Calamine’. When assuming that isogroups will best represent genes, we have expanded the available transcriptome sequence information, in terms of sequenced genes, by almost 5.6 fold.

The use of isotigs, rather than contigs for listing transcripts, will account for small sequence differences between contigs corresponding to the same transcript. Such can be caused by sequence errors, allelic differences, or differences between recently duplicated gene copies. The accession ‘Ganges’ was used for transcriptome sequencing. This accession had been inbred for at least seven generations, and as *N. caerulescens* is a self-compatible, readily self-fertilizing species, we expect that only few sequence differences between contigs will be due to allelic variation and that most of the differences between contigs and isotigs will be due to sequencing errors. The difference between isotigs and isogroups is mainly accounted for by differential or alternative splicing or generation of alternative transcripts from the same gene. The 454 sequence technology that was used generates relatively large reads and facilitates easier assembly compared to that of the much shorter reads created by the Illumina sequencing technology. However, despite the long reads, for about 10-15% of the isogroups more than one
isotigs was assembled, reflecting the occurrence of alternative transcripts (Table S1).

When comparing the *N. caerulescens* transcriptome to that of other diploid Brassicaceae species for which transcriptome sequence is available, such as *Thlaspi arvense* (Dorn *et al.*, 2013) or *Thellungiella salsuginea* (*Eutrema salsugineum*) (Lee *et al.*, 2013), the *N. caerulescens* transcriptome set is relatively small. The draft transcriptome of *T. arvense* consists of 33,873 contigs, representing 25,232 Brassicaceae genes, and that of *T. salsuginea* comprises 42,810 unigenes, corresponding to 24,457 *A. thaliana* peptides. The 20,378 *N. caerulescens* isogroups corresponded to 19,191 protein matches in the Brassicaceae reference proteome set. The main reason for these differences will be the representation of more tissues or organs and conditions in the other sequenced libraries. For *N. caerulescens*, only roots, leaves and flowers of plants grown under normal Zn supply conditions were sampled, while material from more organs or conditions were sampled for the other two species. The different sequencing approaches may also account for the differences. *T. arvense* was sequenced using Illumina technology, which generates many more, but shorter, reads, while for *T. salsuginea* both a normalized and non-normalized library were sequenced, using 454 technology, and in the final assembly, previously generated EST sequences present in the NCBI GenBank database were included.

*T. salsuginea* is predicted to contain 26,521 protein coding genes (Yang *et al.*, 2013), of which 18,970 were found to be represented in the 42,810 unigene set (Lee *et al.*, 2013). Similar gene numbers are also found for *Leavenworthia alabamica* (30,343 genes) and *Sisymbrium irio* (28,917 genes) (Haudry *et al.*, 2013). These three species have a genome size comparable to the genome size of *N. caerulescens*, which is expected to be 310-330 Mb as based on 2C DNA content (Peer *et al.*, 2003a). When expecting that *N. caerulescens* will have around 26-30,000 genes, this means around 60% of its transcriptome is covered in the current dataset.

*N. caerulescens* belongs to the Coluteocarpeae tribe of the Brassicaceae that consists mainly of *Noccaea* species, which, when tested, are all found to be metal hyperaccumulators (Koch&German, 2013). Some of the other species in this tribe, e.g. *Raparia bulbosa* and *Thlaspiceras oxyceras*, are also known to accumulate metals. This tribe belongs to the expanded lineage II of the Brassicaceae, to which also the *Brassica* and *Thellungiella* (*Eutrema*) genera belong. This fits well with the observation that most *N. caerulescens* transcript
sequences find their best BLASTX hit with *T. halophila* proteins (Supplemental table S1). Only 98 sequences appeared to be originating from non-plant organisms, potentially corresponding to organisms living in association with the *N. caerulescens* plants we used to generate the cDNA libraries. Considering that for the *T. arvense* transcriptome set, 424 out of 33,873 contigs showed similarity to fungal genes, the contribution of such genes to the *N. caerulescens* dataset is modest. Upon comparing the *N. caerulescens* isotigs to existing sequence information available in the NCBI GenBank, 1386 isotigs (5.8%) were found to have no significant similarity to any other sequence in database (Supplemental table S4). Potentially these could be genes unique to *N. caerulescens*; however, they may also represent miRNA precursors that are much less conserved and hard to detect using Blast analysis. Of course it is possible that these sequences represent genomic DNA sequences (not very likely considering the use of DNAses in the cDNA library construction), or transcripts of plant-associated organisms that have not been sequenced yet. In the previous *N. caerulescens* transcriptome dataset we generated, we found 8% of the unigenes to show no similarity to any other entry in the NCBI database (Rigola *et al.*, 2006). Comparison of the “no-hit” sequences to the *N. caerulescens* whole genome sequence, when available, will be needed to clarify if these are from *N. caerulescens*.

Isogroups rather than isotigs were used for GO annotation and analysis to avoid overrepresentation of certain GO terms due to alternative transcripts generating many isotigs for the same gene. Sometimes the isogrouping appears to be too strict, which has forced transcripts from different paralogs into one isogroup, as was found for some of the mineral homeostasis related genes (Table 2). The GO analysis (Fig 4 and 5, and supplemental tables S5 and S6) showed that *N. caerulescens* isogroups were largely GO-annotated in comparable GO-class representations as for *A. thaliana*.

Many mineral homeostasis genes were identified among the *N. caerulescens* isotigs (Table 2). Although we identified genes for all relevant gene families, not all genes for each family were found to have *N. caerulescens* counterparts, representing their potential orthologs. For instance, for the ZIP gene family of plasma membrane metal importers, no *N. caerulescens* transcript was found for *IRT2* and *ZIP3, 5, 7, 8 and 12*. For *IRT2*, this is not so remarkable, since it appears to be a recent gene duplication of *IRT1* that has occurred in the Arabidopsis lineage and may not be present in *N. caerulescens*. Not finding potential orthologs for *ZIP5* and *ZIP8* is more remarkable, since transcripts for
these genes were found previously (Rigola et al., 2006) and at least the ortholog of ZIP5 was expressed in roots and shoots (Wu et al., 2009). Also for the MTP family of vacuolar metal importers we did not detect transcripts similar to all A. thaliana MTP genes. Phylogenetic analysis of the ZIP and MTP isotigs showed that those were classified correctly (Fig. 7). For the HMA family of plasma membrane metal exporters, we initially did not find a potential ortholog of the NcHMA3 gene (Ueno et al., 2011). However, upon re-examining the isogroups showing similarity to HMA proteins, isogroup02580, which was found to be very similar to a B. rapa protein annotated as HMA2, was actually highly similar (~90% DNA identity) to A. halleri and A. thaliana HMA3 genes.

Only for the MT3 gene we could identify an additional copy in N. caerulescens compared to the single copies found in related species (Fig 7c). This gene has been implicated in Cu homeostasis, and expression in a southern France accession (‘St. Felix’, which is close to the origin of ‘Ganges’) is much higher than in the two other tested accessions and appears to be constitutive, rather than induced by Cu (Roosens et al., 2004). Such could well be the consequence of an additional copy, which would mean the additional copy is accession, rather than species specific, and is worth further investigations. Not finding copy number expansion for other genes is remarkable, as for several of the other genes listed in Table 2, such as HMA3 and HMA4, multiple copies have already been reported in other accessions, including ‘Ganges’ (Craciun et al., 2012, Iqbal et al., 2013, Ó Lochlainn et al., 2011, Ueno et al., 2011). However, for HMA4, for which cDNA sequences of ‘Ganges’ are available, the different cDNA copies are very similar in sequence (Iqbal et al., 2013), only differing in a few bp and three InDels, which will not be distinguished from sequence differences due to sequencing errors by the assembly software we used. This is likely to be the case for more recently duplicated gene copies. Detailed analysis of copy number variation will be much less ambiguous upon availability of a whole genome sequence, where non-coding sequence can be taken into account to distinguish duplicated copies.

When examining genes involved in glucosinolate biosynthesis, we identified isotigs corresponding to all transcription factor genes involved in aliphatic glucosinolates, but not for two of the MYB genes involved in indolic glucosinolates (MYB34 and MYB122) (Supplemental table S8). In contrast, hardly any of the aliphatic side-chain elongation genes as well as the two CYP79F genes involved in aliphatic core structure formation was represented in the N. caerulescens isogroup list. Of course, not finding them in the isotig list does not
mean these genes are not expressed, but it is remarkable that expression of several of the structural genes involved in aliphatic glucosinolate biosynthesis is so low that apparently these genes are more likely to missed by 454 sequencing than the other genes. Low expression of glucosinolate genes would be in line with a previous report that especially metallicolous accessions from the south of France, where also ‘Ganges’ is originating from, are low in total glucosinolate levels (Noret et al., 2007).

The main reason for studying *N. caerulescens* is to learn more on its extraordinary capacities to tolerate exposure to high concentrations of heavy metals and accumulate these to extremely high levels in the leaves. The transcriptome sequence as such will not tell us that much on which genes will be relevant for these traits, but it can be used as an excellent reference for RNA-Seq studies to determine gene expression in organs of different accessions, exposed to a range of metal concentrations. There are many different *N. caerulescens* populations in Europe, which can differ substantially in their ability to tolerate and accumulate metals, as well as differ in their metal preferences. Many populations grow on non-metallicolous soils, which are not enriched in metals. At those sites, they are likely to accumulate mainly Zn. When exposed to Zn, Ni and Cd they can hyperaccumulate them, but as they are generally not metal tolerant, plants will rapidly die upon exposure. Other natural populations grow on (ultramafic) serpentine outcrops, which are generally rich in metals, but not Zn. At those sites, they often hyperaccumulate Ni, and when exposed to Zn or Cd, will also hyperaccumulate these metals in their shoot. However, they are often only tolerant to Ni and will suffer or die from exposure to Cd. Finally, there are several local populations spread over different sites in Europe that have been heavily contaminated by Zn and Cd, and often also Pb (Mohtadi et al., 2012). These calamine populations are Zn and often Cd hyperaccumulating and are also tolerant to these metals. When exposed to Ni, they will also hyperaccumulate it, but they are less tolerant to it than populations from serpentine sites (Assunção et al., 2003). Especially populations found on metal contaminated sites in the Cevennes region, around the town of Ganges in the south of France, are particularly good at hyperaccumulating Cd. Most of the phenotypic differences related to metal tolerance and accumulation between populations are due to genetic differences, often reflected in differences in gene expression (van de Mortel et al., 2008). With the *N. caerulescens* transcriptome dataset we generated, it will be much easier to study such differences. Also, when the whole genome sequence of *N. caerulescens*
will be determined, the transcriptome data can be used to annotate predicted genes and delineate potential transcripts of such genes. This will be useful to determine the function of these genes. Finally, since the previous transcriptome dataset was obtained from accession ‘La Calamine’, and the new one from ‘Ganges’, the comparison of transcripts from the same genes revealed a SNP frequency of around 2 per 100 bp. Although the presence of sequence errors in both sets will have inflated this number a bit, it will be straightforward to convert these differences into genetic markers that can be used in mapping quantitative trait loci in segregating populations of *N. caerulescens* (Assunção et al., 2006, Deniau et al., 2006).

**ACKNOWLEDGMENTS**

The authors acknowledge the Dutch Graduate School for Experimental Plant Sciences and the Zenith-scheme of the Dutch Organization for Scientific Research (NWO) for financially supporting this research. Corrie Hanhart is gratefully acknowledged for providing the *N. caerulescens* inflorescences used for RNA extraction.
Table 2. *Noccaea caerulescens* isogroups corresponding to *A. thaliana* genes involved in mineral homeostasis and heavy metal stress response.

Gene class abbreviations are: *bZIP*, basic-leucine zipper; *ZIP/IRT*, ZRT/IRT-like Protein; *MTP*, Metal Tolerance Protein; *NRAMP*, Natural Resistance-Associated Macrophage Protein; *HMA*, Heavy Metal ATPase; *YSL*, Yellow Stripe Like protein; *ZIF*, Zinc Induced Facilitator; *NAS*, Nicotianamine Synthase; *MRP/ABCC*, Multidrug Resistance-associated Protein; *PCR*, Plant Cadmium Resistance; *PDR*, Pleiotropic Drug Resistance Protein; *CAX*, Calcium exchanger; *PCS*, Phytochelatin Synthase; *MT*, Metallothionein; *PDF*, Plant Defensin.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Isogroup no.</th>
<th>Longest isotig length</th>
<th>Best hit</th>
<th>E-value</th>
<th>Annotation source</th>
<th>Annotation source description</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>bZIP19</em></td>
<td>isogroup05302</td>
<td>1413</td>
<td>Carubvl0005538mPACid:20894480</td>
<td>1.49E-152</td>
<td>AT4G35040</td>
<td>Basic-leucine zipper (bZIP) transcription factor</td>
</tr>
<tr>
<td><em>IRT1</em></td>
<td>isogroup16405</td>
<td>559</td>
<td>AT4G19690.2PACid:19648297</td>
<td>1.31E-110</td>
<td>AT4G19690</td>
<td>iron-regulated transporter 1</td>
</tr>
<tr>
<td><em>IRT3</em></td>
<td>isogroup00440</td>
<td>1428</td>
<td>Thhalvl0023475mPACid:20201432</td>
<td>0</td>
<td>AT1G60960</td>
<td>iron regulated transporter 3</td>
</tr>
<tr>
<td><em>ZIP1</em></td>
<td>isogroup00441</td>
<td>1341</td>
<td>Carubvl0015186mPACid:20899838</td>
<td>2.19E-110</td>
<td>AT3G12750</td>
<td>zinc transporter 1 precursor</td>
</tr>
<tr>
<td><em>ZIP2</em></td>
<td>isogroup07896</td>
<td>1085</td>
<td>Thhalvl0013963mPACid:20204020</td>
<td>0</td>
<td>AT5G59520</td>
<td>ZRT/IRT-like protein 2</td>
</tr>
<tr>
<td><em>ZIP4</em></td>
<td>isogroup06106</td>
<td>1285</td>
<td>AT1G10970.1PACid:19650740</td>
<td>0</td>
<td>AT1G10970</td>
<td>zinc transporter 4 precursor</td>
</tr>
<tr>
<td><em>ZIP6</em></td>
<td>isogroup06071</td>
<td>1293</td>
<td>Thhalvl0016885mPACid:20179820</td>
<td>5.49E-180</td>
<td>AT2G30080</td>
<td>ZIP metal ion transporter family</td>
</tr>
<tr>
<td><em>ZIP9</em></td>
<td>isogroup09981</td>
<td>918</td>
<td>Thhalvl0025584mPACid:20196090</td>
<td>1.39E-169</td>
<td>AT4G33020</td>
<td>ZIP metal ion transporter family</td>
</tr>
<tr>
<td><em>ZIP10</em></td>
<td>isogroup06264</td>
<td>1263</td>
<td>Thhalvl0007930mPACid:20188086</td>
<td>1.39E-165</td>
<td>AT1G31260</td>
<td>zinc transporter 10 precursor</td>
</tr>
<tr>
<td><em>ZIP11</em></td>
<td>isogroup09047</td>
<td>988</td>
<td>AT1G55910.1PACid:19649515</td>
<td>0</td>
<td>AT1G55910</td>
<td>zinc transporter 11 precursor</td>
</tr>
<tr>
<td><em>MTP1/ZAT</em></td>
<td>isogroup00207</td>
<td>2104</td>
<td>AT2G46800.2PACid:19641115</td>
<td>2.92E-85</td>
<td>AT2G46800</td>
<td>zinc transporter of Arabidopsis thaliana</td>
</tr>
<tr>
<td>Gene</td>
<td>Isogroup</td>
<td>Log2</td>
<td>PACid:</td>
<td>Score</td>
<td>TAIR Accession</td>
<td>Function</td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
<td>------</td>
<td>--------</td>
<td>--------</td>
<td>----------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>MTP5</td>
<td>isogroup01278</td>
<td>1466</td>
<td>Tthalv10020889mPACid:20182053</td>
<td>0</td>
<td>AT3G12100</td>
<td>Cation efflux family protein</td>
</tr>
<tr>
<td>MTP6</td>
<td>isogroup10136</td>
<td>907</td>
<td>Tthalv10001431mPACid:20189077</td>
<td>9.03E-167</td>
<td>AT2G47830</td>
<td>Cation efflux family protein</td>
</tr>
<tr>
<td>MTP7</td>
<td>isogroup03594</td>
<td>1841</td>
<td>Tthalv10011469mPACid:20184771</td>
<td>0</td>
<td>AT1G51610</td>
<td>Cation efflux family protein</td>
</tr>
<tr>
<td>MTP8</td>
<td>isogroup05294</td>
<td>1416</td>
<td>Tthalv10006008mPACid:20190151</td>
<td>0</td>
<td>AT3G58060</td>
<td>Cation efflux family protein</td>
</tr>
<tr>
<td>MTP11</td>
<td>isogroup07586</td>
<td>1117</td>
<td>497437PACid:16043531</td>
<td>0</td>
<td>AT2G39450</td>
<td>Cation efflux family protein</td>
</tr>
<tr>
<td>MTP12</td>
<td>isogroup17157</td>
<td>526</td>
<td>Tthalv10027647mPACid:20189919</td>
<td>6.90E-11</td>
<td>AT2G04620</td>
<td>Cation efflux family protein</td>
</tr>
<tr>
<td>NRAMP1</td>
<td>isogroup00017</td>
<td></td>
<td>Tthalv10018389mPACid:20191829</td>
<td>0</td>
<td>AT1G80830</td>
<td>Natural resistance-associated macrophage protein 1</td>
</tr>
<tr>
<td>NRAMP2</td>
<td>isogroup18133</td>
<td>488</td>
<td>Tthalv10011381mPACid:20184824</td>
<td>6.26E-43</td>
<td>AT1G47240</td>
<td>NRAMP metal ion transporter 2</td>
</tr>
<tr>
<td>NRAMP3</td>
<td>isogroup03577</td>
<td>1845</td>
<td>Bra000573PACid:22711711</td>
<td>0</td>
<td>AT2G23150</td>
<td>Natural resistance-associated macrophage protein 3</td>
</tr>
<tr>
<td>NRAMP4</td>
<td>isogroup06857</td>
<td>1197</td>
<td>Tthalv10004033mPACid:20199053</td>
<td>0</td>
<td>AT5G67330</td>
<td>Natural resistance-associated macrophage protein 4</td>
</tr>
<tr>
<td>NRAMP5</td>
<td>isogroup13240</td>
<td>720</td>
<td>Tthalv10026903mPACid:20193643</td>
<td>1.53E-159</td>
<td>AT4G18790</td>
<td>NRAMP metal ion transporter family protein</td>
</tr>
<tr>
<td>EIN2-NRAMP</td>
<td>isogroup05270</td>
<td>1417</td>
<td>AT5G03280.1PACid:19671290</td>
<td>0</td>
<td>AT5G03280</td>
<td>NRAMP metal ion transporter family protein</td>
</tr>
<tr>
<td></td>
<td>isogroup05627</td>
<td>1367</td>
<td>Tthalv10012456mPACid:20203745</td>
<td>0</td>
<td>AT5G03280</td>
<td>NRAMP metal ion transporter family protein</td>
</tr>
</tbody>
</table>

N. caerulescens transcriptome
<table>
<thead>
<tr>
<th>isogroup</th>
<th>1367</th>
<th>Thalav10012456m</th>
<th>PACid:20203745</th>
<th>0</th>
<th>AT5G03280</th>
<th>NRAMP metal ion transporter family protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>isogroup11803</td>
<td>801</td>
<td>Thalav10012456m</td>
<td>PACid:20203745</td>
<td>7.09E-37</td>
<td>AT5G03280</td>
<td>NRAMP metal ion transporter family protein</td>
</tr>
<tr>
<td><strong>HMA1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>isogroup17702</td>
<td>504</td>
<td>Carubv10004142m</td>
<td>PACid:20894021</td>
<td>1.03E-48</td>
<td>AT4G37270</td>
<td>heavy metal atpase 1</td>
</tr>
<tr>
<td>isogroup03139</td>
<td>2048</td>
<td>490893</td>
<td>PACid:16045505</td>
<td>0</td>
<td>AT4G37270</td>
<td>heavy metal atpase 1</td>
</tr>
<tr>
<td><strong>HMA2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>isogroup08856</td>
<td>1001</td>
<td>Thalav10024341m</td>
<td>PACid:20194930</td>
<td>3.36E-139</td>
<td>AT4G30110</td>
<td>heavy metal atpase 2</td>
</tr>
<tr>
<td>isogroup10109</td>
<td>910</td>
<td>Bra024107</td>
<td>PACid:22710111</td>
<td>4.02E-103</td>
<td>Bra024107</td>
<td>heavy metal atpase 2</td>
</tr>
<tr>
<td>isogroup02580</td>
<td>2745</td>
<td>857597</td>
<td>PACid:16066225</td>
<td>0</td>
<td>Bra011165</td>
<td>heavy metal atpase 2</td>
</tr>
<tr>
<td><strong>HMA4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>isogroup00021</td>
<td>4417</td>
<td>Carubv10012822m</td>
<td>PACid:20897962</td>
<td>0</td>
<td>AT2G19110</td>
<td>heavy metal atpase 4</td>
</tr>
<tr>
<td><strong>HMA6/PAAX1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>isogroup07162</td>
<td>1157</td>
<td>AT4G33520.</td>
<td>PACid:19645606</td>
<td>1.41E-96</td>
<td>AT4G33520</td>
<td>P-type ATPase 1</td>
</tr>
<tr>
<td><strong>HMA7/RAN1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>isogroup11167</td>
<td>839</td>
<td>Thalav10000758m</td>
<td>PACid:20202842</td>
<td>3.25E-166</td>
<td>AT5G44790</td>
<td>copper-exporting ATPase (RAN1)</td>
</tr>
<tr>
<td><strong>HMA8/PAAX2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>isogroup02963</td>
<td>2165</td>
<td>Carubv10000184m</td>
<td>PACid:20908511</td>
<td>0</td>
<td>AT5G21930</td>
<td>P-type ATPase of Arabidopsis 2</td>
</tr>
<tr>
<td><strong>YSI1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>isogroup02761</td>
<td>2382</td>
<td>Bra013764</td>
<td>PACid:22706007</td>
<td>0</td>
<td>Bra013764</td>
<td>YELLOW STRIPE like 1</td>
</tr>
<tr>
<td><strong>YSI3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>isogroup01006</td>
<td>2083</td>
<td>Thalav10012886m</td>
<td>PACid:20203959</td>
<td>0</td>
<td>AT5G35550</td>
<td>YELLOW STRIPE like 3</td>
</tr>
<tr>
<td><strong>YSI5</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>isogroup03348</td>
<td>1934</td>
<td>Thalav10020160m</td>
<td>PACid:20182600</td>
<td>0</td>
<td>AT3G17650</td>
<td>YELLOW STRIPE like 5</td>
</tr>
<tr>
<td><strong>YSI6</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>isogroup00783</td>
<td>2244</td>
<td>Thalav10003758m</td>
<td>PACid:20198225</td>
<td>0</td>
<td>AT3G27020</td>
<td>YELLOW STRIPE like 6</td>
</tr>
<tr>
<td><strong>YSI7</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>isogroup17698</td>
<td>497</td>
<td>Thalav10018216m</td>
<td>PACid:20192307</td>
<td>1.52E-84</td>
<td>AT1G65730</td>
<td>YELLOW STRIPE like 7</td>
</tr>
<tr>
<td>isogroup12528</td>
<td>757</td>
<td>Thalav10018216m</td>
<td>PACid:20192307</td>
<td>1.23E-138</td>
<td>AT1G65730</td>
<td>YELLOW STRIPE like 7</td>
</tr>
<tr>
<td><strong>ZIF1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>isogroup05885</td>
<td>1322</td>
<td>Thalav10013388m</td>
<td>PACid:20203275</td>
<td>8.58E-175</td>
<td>AT5G13740</td>
<td>zinc induced facilitator 1</td>
</tr>
<tr>
<td><strong>ZIFL1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>isogroup13593</td>
<td>697</td>
<td>Bra008805</td>
<td>PACid:22694599</td>
<td>3.84E-111</td>
<td>AT5G13750</td>
<td>zinc induced facilitator-like 1</td>
</tr>
<tr>
<td>isogroup14641</td>
<td>646</td>
<td>Thalav10013417m</td>
<td>PACid:20206568</td>
<td>1.30E-136</td>
<td>AT5G13750</td>
<td>zinc induced facilitator-like 1</td>
</tr>
<tr>
<td><strong>ZIFL2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>isogroup02224</td>
<td>478</td>
<td>Bra021053</td>
<td>PACid:22721993</td>
<td>1.81E-60</td>
<td>AT3G43790</td>
<td>zinc induced facilitator-like 2</td>
</tr>
<tr>
<td>Protein</td>
<td>isoform</td>
<td>Genotype</td>
<td>PACid</td>
<td>Expression</td>
<td>accession</td>
<td>Function</td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
<td>----------</td>
<td>---------</td>
<td>-------------</td>
<td>-----------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td><em>NAS1</em></td>
<td>isogroup15107</td>
<td>Thalv10014123m</td>
<td>PACid:20206622</td>
<td>3.97E-119</td>
<td>AT5G04950</td>
<td>nicotianamine synthase 1</td>
</tr>
<tr>
<td><em>NAS2</em></td>
<td>isogroup06206</td>
<td>Thalv10013812m</td>
<td>PACid:20203757</td>
<td>0</td>
<td>AT5G56080</td>
<td>nicotianamine synthase 2</td>
</tr>
<tr>
<td><em>NAS3</em></td>
<td>isogroup09071</td>
<td>Thalv10008206m</td>
<td>PACid:20187253</td>
<td>0</td>
<td>AT1G09240</td>
<td>nicotianamine synthase 3</td>
</tr>
<tr>
<td><em>NAS4</em></td>
<td>isogroup06921</td>
<td>Thalv10024166m</td>
<td>PACid:20201706</td>
<td>0</td>
<td>AT1G56430</td>
<td>nicotianamine synthase 4</td>
</tr>
<tr>
<td><em>MRP1/ABCC1</em></td>
<td>isogroup04168</td>
<td>473263PACid:16034489</td>
<td>1.04E-12</td>
<td>AT1G30400</td>
<td>multidrug resistance-associated protein 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>isotig01105</td>
<td>Thalv10006546m</td>
<td>PACid:20187496</td>
<td>6.64E-96</td>
<td>AT1G30400</td>
<td>multidrug resistance-associated protein 1</td>
</tr>
<tr>
<td><em>MRP2/ABCC2</em></td>
<td>isogroup11349</td>
<td>1975</td>
<td>AT1G30400.2PACid:19655137</td>
<td>3.18E-06</td>
<td>AT1G30400</td>
<td>multidrug resistance-associated protein 1</td>
</tr>
<tr>
<td></td>
<td>isotig01102</td>
<td>Thalv10016133m</td>
<td>PACid:20181262</td>
<td>6.34E-06</td>
<td>AT2G34660</td>
<td>multidrug resistance-associated protein 2</td>
</tr>
<tr>
<td><em>MRP3/ABCC3</em></td>
<td>isogroup13185</td>
<td>724</td>
<td>482430PACid:16055424</td>
<td>3.02E-83</td>
<td>AT2G34660</td>
<td>multidrug resistance-associated protein 2</td>
</tr>
<tr>
<td></td>
<td>isogroup17760</td>
<td>504</td>
<td>Thalv10019893m</td>
<td>PACid:20181977</td>
<td>4.35E-23</td>
<td>AT3G13080</td>
</tr>
<tr>
<td></td>
<td>isogroup02973</td>
<td>2159</td>
<td>Thalv10019893m</td>
<td>PACid:20181977</td>
<td>0</td>
<td>AT3G13080</td>
</tr>
<tr>
<td><em>MRP4/ABCC4</em></td>
<td>isogroup08374</td>
<td>1033</td>
<td>Thalv10019893m</td>
<td>PACid:19658374</td>
<td>3.01E-12</td>
<td>AT3G13080</td>
</tr>
<tr>
<td></td>
<td>isotig00322</td>
<td>Bra018722</td>
<td>PACid:22695420</td>
<td>0</td>
<td>AT2G47800</td>
<td>multidrug resistance-associated protein 4</td>
</tr>
<tr>
<td></td>
<td>isotig00323</td>
<td>Bra018722</td>
<td>PACid:22695420</td>
<td>0</td>
<td>AT2G47800</td>
<td>multidrug resistance-associated protein 4</td>
</tr>
<tr>
<td></td>
<td>isotig00324</td>
<td>Bra018722</td>
<td>PACid:22695420</td>
<td>0</td>
<td>AT2G47800</td>
<td>multidrug resistance-associated protein 4</td>
</tr>
<tr>
<td></td>
<td>isotig00326</td>
<td>Bra018722</td>
<td>PACid:22695420</td>
<td>0</td>
<td>AT2G47800</td>
<td>multidrug resistance-associated protein 4</td>
</tr>
<tr>
<td>Gene</td>
<td>Isoform</td>
<td>MIDI</td>
<td>Description</td>
<td>Chromosome</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>---------</td>
<td>------</td>
<td>-------------</td>
<td>------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRP5/ABCC5</td>
<td>isogroup02478</td>
<td>3126</td>
<td>Carubv10008087m</td>
<td>PACid:20891448</td>
<td>0</td>
<td>AT1G04120 multidrug resistance-associated protein 5</td>
</tr>
<tr>
<td></td>
<td>isogroup10208</td>
<td>901</td>
<td>Thhalv10006549m</td>
<td>PACid:20186680</td>
<td>0</td>
<td>AT1G04120 multidrug resistance-associated protein 5</td>
</tr>
<tr>
<td></td>
<td>isogroup14209</td>
<td>665</td>
<td>Thhalv10006549m</td>
<td>PACid:20186680</td>
<td>3.34E-38</td>
<td>AT1G04120 multidrug resistance-associated protein 5</td>
</tr>
<tr>
<td>MRP6/ABCC8</td>
<td>isogroup15243</td>
<td>615</td>
<td>Thhalv10019895m</td>
<td>PACid:20181972</td>
<td>3.67E-53</td>
<td>AT3G21250 multidrug resistance-associated protein 6</td>
</tr>
<tr>
<td></td>
<td>isogroup07630</td>
<td>1109</td>
<td>Thhalv10019895m</td>
<td>PACid:20181972</td>
<td>3.29E-07</td>
<td>AT3G21250 multidrug resistance-associated protein 6</td>
</tr>
<tr>
<td>MRP7/ABCC7</td>
<td>isogroup20180</td>
<td>277</td>
<td>Bra034706</td>
<td>PACid:22686665</td>
<td>9.37E-37</td>
<td>AT3G13100 multidrug resistance-associated protein 7</td>
</tr>
<tr>
<td></td>
<td>isogroup12703</td>
<td>742</td>
<td>Thhalv10019894m</td>
<td>PACid:20183109</td>
<td>1.81E-11</td>
<td>AT3G13100 multidrug resistance-associated protein 7</td>
</tr>
<tr>
<td></td>
<td>isogroup14335</td>
<td>661</td>
<td>Bra034706</td>
<td>PACid:22686665</td>
<td>3.50E-136</td>
<td>AT3G13100 multidrug resistance-associated protein 7</td>
</tr>
<tr>
<td>MRP9/ABCC9</td>
<td>isogroup04767</td>
<td>1517</td>
<td>Bra003402</td>
<td>PACid:22702084</td>
<td>6.74E-14</td>
<td>AT3G60160 multidrug resistance-associated protein 9</td>
</tr>
<tr>
<td></td>
<td>isogroup09719</td>
<td>937</td>
<td>486482</td>
<td>PACid:16051413</td>
<td>7.50E-07</td>
<td>AT3G60160 multidrug resistance-associated protein 9</td>
</tr>
<tr>
<td>MRP10/ABCC14</td>
<td>isotig00325</td>
<td>5040</td>
<td>Thhalv10005741m</td>
<td>PACid:20190810</td>
<td>0</td>
<td>AT3G62700 multidrug resistance-associated protein 10</td>
</tr>
<tr>
<td></td>
<td>isotig00327</td>
<td>4933</td>
<td>Thhalv10005741m</td>
<td>PACid:20190810</td>
<td>0</td>
<td>AT3G62700 multidrug resistance-associated protein 10</td>
</tr>
<tr>
<td>MRP14/ABCC10</td>
<td>isogroup15495</td>
<td>603</td>
<td>AT3G59140.1</td>
<td>PACid:19663037</td>
<td>6.18E-119</td>
<td>AT3G59140 multidrug resistance-associated protein 14</td>
</tr>
<tr>
<td></td>
<td>isogroup18407</td>
<td>480</td>
<td>486382</td>
<td>PACid:16034938</td>
<td>5.72E-91</td>
<td>AT3G59140 multidrug resistance-associated protein 14</td>
</tr>
<tr>
<td>Gene</td>
<td>Accession</td>
<td>AT1G14870.1PACid:19649994</td>
<td>1.51E-88</td>
<td>AT1G14870</td>
<td>Plant Cadmium Resistance 2</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-----------</td>
<td>---------------------------</td>
<td>---------</td>
<td>------------</td>
<td>-----------------------------</td>
<td></td>
</tr>
<tr>
<td>PCR2</td>
<td>isogroup15934</td>
<td>Bra026796PACid:22713063</td>
<td>1.43E-111</td>
<td>AT1G14870</td>
<td>Plant Cadmium Resistance 2</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Accession</th>
<th>AT1G14870.1PACid:20182879</th>
<th>1.37E-66</th>
<th>AT3G16340</th>
<th>pleiotropic drug resistance 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDR1</td>
<td>isogroup02600</td>
<td>Thhalv10019897mlPACid:20182879</td>
<td>8.02E-38</td>
<td>AT3G16340</td>
<td>pleiotropic drug resistance 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Accession</th>
<th>AT2G29940.1PACid:19639436</th>
<th>6.90E-12</th>
<th>AT2G29940</th>
<th>pleiotropic drug resistance 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDR3</td>
<td>isogroup02662</td>
<td>Thhalv10001880mlPACid:20200369</td>
<td>4.23E-33</td>
<td>AT2G26910</td>
<td>pleiotropic drug resistance 4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Accession</th>
<th>AT1G161980mlPACid:20200369</th>
<th>3.89E-17</th>
<th>AT2G26910</th>
<th>pleiotropic drug resistance 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDR5</td>
<td>isogroup03404</td>
<td>Thhalv10001880mlPACid:20200369</td>
<td>3.89E-17</td>
<td>AT2G26910</td>
<td>pleiotropic drug resistance 4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Accession</th>
<th>AT2G36380.1PACid:19640612</th>
<th>6.09E-16</th>
<th>AT2G36380</th>
<th>pleiotropic drug resistance 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDR6</td>
<td>isogroup11286</td>
<td>Thhalv10016141mlPACid:20179582</td>
<td>4.60E-08</td>
<td>AT2G36380</td>
<td>pleiotropic drug resistance 5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Accession</th>
<th>AT2G36380.1PACid:20180787</th>
<th>2.61E-09</th>
<th>AT2G36380</th>
<th>pleiotropic drug resistance 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDR7</td>
<td>isogroup13734</td>
<td>Thhalv10016140mlPACid:20180787</td>
<td>2.58E-10</td>
<td>AT2G36380</td>
<td>pleiotropic drug resistance 6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Accession</th>
<th>AT1G162100mlPACid:20180787</th>
<th>4.48E-14</th>
<th>AT1G15210</th>
<th>pleiotropic drug resistance 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDR7</td>
<td>isogroup12430</td>
<td>Thhalv10012050mlPACid:20892263</td>
<td>2.05E-11</td>
<td>AT1G15210</td>
<td>pleiotropic drug resistance 7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Accession</th>
<th>Bra026157PACid:22695455</th>
<th>4.48E-126</th>
<th>Bra026157</th>
<th>pleiotropic drug resistance 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDR9/</td>
<td>isogroup14785</td>
<td>Bra003137PACid:22699106</td>
<td>1.69E-99</td>
<td>AT3G53480</td>
<td>pleiotropic drug resistance 9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Accession</th>
<th>AT1G14870.1PACid:19649994</th>
<th>1.51E-88</th>
<th>AT1G14870</th>
<th>Plant Cadmium Resistance 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR2</td>
<td>isogroup15934</td>
<td>Bra026796PACid:22713063</td>
<td>1.43E-111</td>
<td>AT1G14870</td>
<td>Plant Cadmium Resistance 2</td>
</tr>
<tr>
<td>Gene</td>
<td>Isoform</td>
<td>Species</td>
<td>Accession Number</td>
<td>Expression</td>
<td>Annotation</td>
</tr>
<tr>
<td>------</td>
<td>---------</td>
<td>---------</td>
<td>-----------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>CAX1</td>
<td>isogroup01205</td>
<td>Thhalv10016611m</td>
<td>PACid:20180411</td>
<td>0</td>
<td>AT2G38170 cation exchanger 1</td>
</tr>
<tr>
<td>CAX2</td>
<td>isogroup16996</td>
<td>Thhalv10020744m</td>
<td>PACid:20182084</td>
<td>3.06E-164</td>
<td>AT3G13320 cation exchanger 2</td>
</tr>
<tr>
<td>CAX3</td>
<td>isogroup04417</td>
<td>Thhalv10020744m</td>
<td>PACid:20182084</td>
<td>9.22E-94</td>
<td>AT3G13320 cation exchanger 2</td>
</tr>
<tr>
<td>CAX4</td>
<td>isogroup07282</td>
<td>Bra009640m</td>
<td>PACid:22694292</td>
<td>5.20E-82</td>
<td>AT5G01490 cation exchanger 4</td>
</tr>
<tr>
<td>CAX5</td>
<td>isogroup04418</td>
<td>Thhalv10011487m</td>
<td>PACid:20184135</td>
<td>1.59E-79</td>
<td>AT1G55730 cation exchanger 5</td>
</tr>
<tr>
<td>CAX6</td>
<td>isogroup15879</td>
<td>Thhalv10009154m</td>
<td>PACid:20889697</td>
<td>3.20E-93</td>
<td>AT1G55730 cation exchanger 5</td>
</tr>
<tr>
<td>CAX7</td>
<td>isogroup06161</td>
<td>897667m</td>
<td>PACid:16066029</td>
<td>0</td>
<td>AT3G14070 cation exchanger 9</td>
</tr>
<tr>
<td>PCS1</td>
<td>isogroup03039</td>
<td>1670</td>
<td>Thhalv10003194m</td>
<td>PACid:20208031</td>
<td>0</td>
</tr>
<tr>
<td>PCS2</td>
<td>isogroup03040</td>
<td>1095</td>
<td>Thhalv10007644m</td>
<td>PACid:20187384</td>
<td>0</td>
</tr>
<tr>
<td>MT2a</td>
<td>isogroup01808</td>
<td>Bra029765</td>
<td>PACid:22685892</td>
<td>6.28E-24</td>
<td>Bra029765 metallothionein 2A</td>
</tr>
<tr>
<td>MT3</td>
<td>isogroup17171</td>
<td>Thhalv10015167m</td>
<td>PACid:20205299</td>
<td>1.01E-16</td>
<td>AT3G09390 metallothionein 2A</td>
</tr>
<tr>
<td>PDF1.2</td>
<td>isogroup16374</td>
<td>Thhalv10019366m</td>
<td>PACid:20192145</td>
<td>6.62E-42</td>
<td>AT5G44420 plant defensin 1.2</td>
</tr>
</tbody>
</table>

1. These isitogs were both grouped into isogroup00190
2. These isitogs were all grouped into isogroup00027
3. These isitogs were both grouped into isogroup001669
4. These isitogs were both grouped into isogroup00399
5. These isitogs were both grouped into isogroup00980
6. This isogroup was found to be most similar to AthMA3
SUPPLEMENTAL TABLES AND FIGURES

Table S1. *N. caerulescens* contigs and isotigs, with corresponding isogroups, with their BlastX best hit when compared to the Brassicaceae proteome dataset. Only 16 contigs could not be assembled into an isotig and are listed at the top. The sequence length (in bp) as well as the number of reads constituting the contig or isotig is indicated.

Table S2. *N. caerulescens* isotigs with similarity to plant sequences found in the non-redundant NCBI GenBank protein database not represented in the Brassicaceae proteome dataset

Table S3. *N. caerulescens* isotigs with similarity to non-plant proteins sequences.

Table S4. *N. caerulescens* isotigs with no similarity to any sequence in the NCBI nr database.

Table S5. Gene Ontology (GO) annotation of *N. caerulescens* isogroups

Table S6. GOSlim classification of annotated *N. caerulescens* GO terms.

Table S7. PlantCyc listing of biological pathways in which *N. caerulescens* isogroups are involved in.

Table S8. Genes involved in glucosinolate biosynthesis.

* The supplementary tables can be downloaded from the website, https://drive.google.com/?tab=mo&authuser=0#folders/0B9G8Qh7ykHV1czhBZHdIQ2Y3Rk0
User Name: yafenphdthesis@gmail.com
Password:20140324
Chapter 6

General Discussion
General aspects of metal toxicity and metal hyperaccumulation

Metal micronutrients such as Zn, Ni, Fe, Cu and Mn play important roles in plant metabolism, growth and development, but they are toxic at higher concentrations. Furthermore, non-essential elements such as Cd, Pb and Hg are also taken up by plants and are toxic even at low concentrations. Metals must therefore be kept within an appropriate concentration range to avoid mineral deficiency on one hand and damage to plant cells on the other. The appropriate range differs for each type of metal. A dose-response curve can therefore be used to find the optimal range for metals that are available in particular environments (Fig. 1 in Chapter 2). In this thesis, I report on different metal exposure experiments using both an essential nutrient (Zn) and a non-essential element (Cd). Plants exposed to high metal concentrations may demonstrate either sensitivity or tolerance (Fig. 2 in Chapter 2). Sensitive plants cannot cope with metals accumulating to high levels in the roots and shoots, and suffer effects such as short roots, yellow leaves, limited seed production and eventually death. In contrast, metal-tolerant plants can survive when exposed to high concentrations of metals, e.g. by excluding metals from the roots (heavy metal excluders), storing metals in root vacuoles (heavy metal tolerant non-hyperaccumulators), or by transporting metals for sequestration in shoot vacuoles or storage tissues, such as trichomes or older leaves (heavy metal hypertolerant hyperaccumulators).

The threshold used to define metal hyperaccumulation was recently set to 3000 μg/g Zn and 100 μg/g Cd as a component of dried foliage (Ent et al., 2013). Based on these criteria, the known 300,000 vascular plant species include 15 Zn hyperaccumulators and 5 Zn/Cd hyperaccumulators (Krämer, 2010). However, as further plants are tested for metal tolerance, the number of hyperaccumulators and their taxonomic nomenclature will change. Recently Gomphrena clausenii was described as the first South American metallophyte, and was defined as a hypertolerant hyperaccumulating bioindicator of Zn and Cd (Carvalho et al., 2013). The sequences of the ribosomal DNA internal transcribed spacer and the plastid trnL-trnF region have recently been used as evidence to reassign plants from the genus Thlaspi, e.g. T. caerulescens has been renamed Noccaea caerulescens, T. montanum has been renamed N. fendleri, T. praecox has been renamed N. praecox and T. perfoliatum has been renamed Microthlaspi perfoliatum (Koch&Al-Shehbaz, 2004, Koch&German, 2013, Meyer, 2003, 2001). A comprehensive list of the amended taxonomy of the Noccaea genus was recently presented by
Although our research focused on different *N. caerulescens* accessions, the phylogenetic context is important because it allows metal tolerance and accumulation in *N. caerulescens* to be compared with other species from the same genus to investigate the evolutionary aspects of metal homeostasis in these species. A well-characterized phylogeny allows the selection of useful comparators.

**N. caerulescens as model for metal hyperaccumulation**

*N. caerulescens* and *A. halleri* are ideal model hyperaccumulators because of their successful adaptation to extremophilic environments and their ability to hyperaccumulate multiple metals including Zn, Cd, Ni and Pb. *N. caerulescens* is more amenable than *A. halleri* because it is self-fertilizing which makes genetic analysis more straightforward, and thus facilitates the comparison of different accessions (Assunção et al., 2003a, Milner&Kochian, 2008).

Different *N. caerulescens* populations from serpentine, calamine and non-metallicolous soils provide germplasm that can be used to study the genetic basis of ecological and evolutionary behavior, based on their diverse phenotypes: For example Ganges is Zn/Cd hypertolerant hyperaccumulators, La Calamine is Zn/Cd tolerant Zn hyperaccumulator, Monte Prinzer is a Ni tolerant Zn/Ni/Cd hyperaccumulator, and Lellingen is a Zn/Ni/Cd intolerant, Zn/Ni/Cd hyperaccumulator (Assunção et al., 2003b). La Calamine is mainly a Zn hyperaccumulator (not Cd or Ni), which is useful to study Zn-related transporters such as NcZNT1 (Chapter 3). In contrast, Ganges is a Zn/Cd hypertolerant hyperaccumulator which can be used to investigate the hypertolerance/hyperaccumulation phenotype by next-generation sequencing (Chapter 5). The genetic variation among different *N. caerulescens* populations provides insight into the evolutionary adaptation of plants to suit metalliferous environments. In Chapter 4, the different nasl allele frequencies in three *N. caerulescens* populations suggested that these three populations experience either different selection pressures or represent different stages of a selective sweep towards the fixation of the nasl mutant allele. The absence of the nasl allele in accessions collected from six other populations suggests that the selective advantage of nasl only exists in the La Calamine, Plombières and Prayon accessions, or that the transposon-induced nasl mutation is a recent event that is still comparatively rare in the species germplasm.
In addition to the multiple and sometimes distinct metal hyperaccumulation properties in different *N. caerulescens* ecotypes, the species is closely related (88.5% identity at the transcriptome level) to the non-accumulator *Arabidopsis thaliana*. This allows researchers to benefit from the established methods such as *in planta* transformation, and the abundant resources available for this well-characterized model. For example, *A. thaliana* sequences have been used to clone homologous *N. caerulescens* genes like *NcZNT1*, which was cloned based on the conserved sequences of *A. thaliana* ZIP genes (Chapter 3) (Assunção et al., 2001), and *NcNAS* genes (Chapter 4). Although *N. caerulescens* can be transformed by flower dipping or tissue culture methods, the gene transfer frequency and stability remain low (Guan et al., 2008, Peer et al., 2003). An efficient hairy root transformation system based on *Agrobacterium rhizogenes* was therefore used to introduce GFP and GUS reporter genes driven by the *NcZNT1* promoter into *N. caerulescens* roots, showing that the *NcZNT1* gene is expressed in the cortex, endodermis, pericycle and vascular tissues, suggesting its involvement in transferring Zn to the stele for long-distance transport (Chapter 3). This method is only applicable in root tissues, so the overexpression of *N. caerulescens* metal-homeostasis genes in shoots was carried out in *A. thaliana*, which has a well-established transformation system, in order to investigate the function of these genes *in planta*. The constitutive overexpression of *NcZNT1* in *A. thaliana* suggested that its function is to increase Zn/Cd tolerance and accumulation.

*N. caerulescens* is a member of the family Brassicaceae, which encompasses ~25% of the known metal hyperaccumulators (Hörger et al., 2013). For example, among 15 known Zn hyperaccumulators in nine taxa, there are three Brassicaceae species (*A. halleri*, *N. caerulescens* and *Arabis paniculata*) representing at least three independent origins of Zn hyperaccumulation (Krämer, 2010). The qualitative and quantitative comparison of transcript sequences among these species (Haudry et al., 2013), and among non-hyperaccumulator Brassicaceae species such as *A. thaliana*, *A. lyrata*, *Brassica rapa*, *Capsella rubella* and *Eutrema halophila*, will help to identify candidate genes that may regulate heavy metal homeostasis and provide insight into the underlying molecular mechanisms of metal tolerance and accumulation (Chapter 5).
Ecological aspects of metal hyperaccumulation

Hyperaccumulators show extreme tolerance to high concentrations of metals and extraordinary metal storing ability, attracting the interest of researchers looking for new hyperaccumulator species, exploring the molecular and physiological mechanisms of metal hyperaccumulation and hypertolerance, and examining the application of these species for phytoremediation (Boyd, 2004). There is also interest in ecological studies that aim to understand the relationship between metal tolerance/accumulation traits and plant fitness or fecundity, especially where hyperaccumulation is considered to be an adaptive trait. When plants hyperaccumulate metals, this can have an impact on other species, including herbivores and pathogens, in their native habitats (Boyd, 2010, Lürling&Scheffer, 2007). There is mounting evidence that the elemental/inorganic defense hypothesis may explain the selective advantages of hyperaccumulation in *N. caerulescens*, in preference to the tolerance/disposal hypothesis, the allelopathy hypothesis, and the drought resistance hypothesis, which are not supported by robust empirical data (Boyd, 2007).

The elemental/inorganic defense hypothesis proposes that metal hyperaccumulation is a defensive trait that protects against pathogens and/or herbivores (Hörger et al., 2013). Although the role of Zn in elemental defense is a matter of debate (Boyd, 2012), our investigation provided evidence that high Zn levels can protect *N. caerulescens* from the herbivore *Pieris rapae*. The higher Zn concentrations in *nas1* mutants compared to WT plants caused a stronger negative impact on caterpillars (Chapter 4). *N. caerulescens* can also resist several other natural herbivores and pathogens such as the grasshopper *Schistocerca gregaria* (Behmer et al., 2005), mildew (*Erysiphe* spp.) (Fones et al., 2010), and a slug (*Deroceras caruanae* Pollonera) (Pollard&Baker, 1997), but no significant resistance has been observed against thrips (*Frankliniella occidentalis* Pergande) (Jiang et al., 2005) and snails (*Helix aspersa*) (Noret et al., 2005) under high Zn conditions. Zn hyperaccumulation appears to preferentially inhibit natural enemies that are dependent on *N. caerulescens*. However, all demonstrations of defense against herbivores were observed under laboratory conditions and there are no corresponding field data. The only field test to investigate the Zn defense hypothesis concluded that *N. caerulescens* was subject to lower herbivory pressure from slugs and snails in metalliferous soils than normal soils around Belgium/Luxembourg and southern France (Noret et al., 2007). However, this
study did not invoke the Zn defense hypothesis in the protection of *N. caerulescens*, but proposed that glucosinolates played a role in the organic defense of metal hyperaccumulators against herbivores. Comparing the results from laboratory and field tests, high Zn levels did not protect *N. caerulescens* against snails in either environment (Noret *et al.*, 2005, Noret *et al.*, 2007), but protected against slugs under laboratory conditions (Pollard&Baker, 1997), although not in natural environments (Noret *et al.*, 2007). We also considered the Zn defense hypothesis as an explanation for the prevalence of the *nas1* mutant along with WT plants at a natural calamine site in Plombières, Belgium. However, the observed metal levels in this *N. caerulescens* population were highly variable between plants and could not be correlated with a particular genotype, probably because they reflect a combination of variable metal concentrations in the soil and genetic diversity among the natural population. Such variability makes it difficult to test the hypothesis directly in natural environments and means that controlled laboratory tests are required even though these do not accurately represent natural habitats and populations.

The *nas1* case provides an example of ecological studies related to exceptionally high metal concentrations (Chapter 4). The elevated metal levels in the natural *N. caerulescens* accession containing the *nas1* transposon insertion allele conferred a selective advantage by protecting against herbivores (Chapter 4), but even so the frequency of the allele differed significantly among three natural *N. caerulescens* populations: Plombières (0.59), La Calamine (0.55) and Prayon (0.91). This suggests that the *nas1* allele may be experiencing different degrees of selection pressure in the three populations, or that it may represent different stages in a selective sweep towards fixation (Babbitt *et al.*, 2010, Tian *et al.*, 2002). We scanned for polymorphisms in and around the *NAS1* locus to identify signatures of selection by homozygous WT and *nas1* plants from Plombières (Nielsen, 2005). Few polymorphisms were identified (data not shown), which suggests the allele is more likely to be undergoing a selective sweep towards positive directional selection.

The selective advantage of the *nas1* mutant (Chapter 4) suggests a trade-off between metal-accumulating traits and plant survival in metalliferous environments (Maestri *et al.*, 2010). As soon as a mutation occurs, the plant must compensate for the genetic change by reallocating its resources to maintain fitness. The selective advantage of the *nas1* allele is based on its ability to accumulate more metal, which involves committing energy resources in the form of ATP to
produce and activate metal transporters, thus enhancing metal translocation and sequestration. This diverts energy from plant growth and metal tolerance is reduced, resulting in selective disadvantages such as lower biomass accumulation, delayed development and greater metal sensitivity. Therefore, the plants must strike a balance between metal hyperaccumulation and physiological performance to maintain fitness in nature.

**Mechanisms of metal hyperaccumulation**

The molecular basis of heavy metal accumulation in plants can be subdivided into three major processes: 1) mobilization, uptake and sequestration of heavy metals in roots; 2) xylem loading/unloading and root-to-shoot translocation; and 3) tissue distribution and the sequestration of heavy metals in shoots and leaves (Krämer et al., 2007). Many genes encoding metal transporters are overexpressed in hyperaccumulators such as *N. caerulescens* and *A. halleri*, compared to the non-accumulator *A. thaliana* (Becher et al., 2004, Chiang et al., 2006, van de Mortel et al., 2006, van de Mortel et al., 2008, Weber et al., 2004). There are many ways to increase the transcription of these genes, including a higher copy number and more efficient transactivation. In *N. caerulescens*, hypertolerance or hyperaccumulation reflects at least three types of modifications: copy number expansion, enhanced promoter activity and more efficient transcription factors. These may have led to the evolution of metal hyperaccumulators and the resulting selective advantages of this trait in *N. caerulescens*.

Copy-number expansion has been reported for several metal homeostasis genes in *A. halleri* (*HMA4, ZIP3, ZIP9 and MTP1*) (Dräger et al., 2004, Hanikenne et al., 2008, Mirouze et al., 2006, Shahzad et al., 2010) and *N. caerulescens* (*HMA4, IRT1 and HMA3*) (Ó Lochlainn et al., 2011, Plaza et al., 2007, Ueno et al., 2011) compared to orthologs in *A. thaliana*. This increase in copy number is proposed to be essential for Zn hyperaccumulation in *A. halleri*, which contains three copies of *HMA4* that boost expression compared to *A. thaliana* (Hanikenne et al., 2008). The copy number of the Zn/Cd transporter *NcHMA4* is not fixed at the species level. Higher copy numbers were found in *N. caerulescens* accessions from calamine sites in Saint Laurent Le Minier, France (four copies), Saint-Félix-de-Pallières, France (four copies), Prayon, Belgium (three copies) and La Calamine, Belgium (three copies) compared to accessions from serpentine soils such as Puente Basadre, Spain (two copies) (Craciun et al., 2012, Ó Lochlainn et al., 2011).
When comparing accessions from serpentine and calamine sites, the lower *NcHMA4* expression levels observed in serpentine accessions was probably at least partially due to the lower copy number, which may also explain the lower Cd translocation capacity of these plants (Craciun *et al.*, 2012). The transcriptome analysis of *N. caerulescens* (Chapter 5) did not reveal much on copy number variation, as the sequence difference between copies appear to be too little to be distinguished by the assembly software that was used. Only for the *MT3* gene, an additional copy was detected, while all other examined Brassicaceae species carry one *MT3* gene copy. Copy number expansion of *MT3* corresponds to higher *MT3* gene expression in the southern accessions (Roosens *et al.*, 2004, Roosens *et al.*, 2005).

The transcription of metal homeostasis genes may also be enhanced by optimizing the *cis*-acting elements in the promoter. For example, placing the *AtHMA4* gene downstream of *NcHMA4* promoters from different *N. caerulescens* ecotypes resulted in changes in the capacity of the transgenic plants to accumulate Zn (Iqbal *et al.*, 2013). We found likewise that the *NcZNT1* gene was expressed more strongly and more widely in *N. caerulescens* than its ortholog in *A. thaliana* (van de Mortel *et al.*, 2006). This was determined by the promoter, as shown by the comparative results of expressing the constructs *proNcZNT1::GUS* and *proAtZIP4::GUS* in *N. caerulescens* roots (Chapter 3). The enhanced activity of the promoter in *N. caerulescens* results in a more efficient response to high Zn and Cd levels and contributes to the hypertolerant/hyperaccumulator phenotype. Furthermore, the *NcZNT1* promoter behaved differently in *N. caerulescens* and *A. thaliana*, i.e. *proZNT1::GUS* activity was only induced under conditions of Zn deficiency in *A. thaliana* but was constitutive in *N. caerulescens*. This suggests that differences in transactivation also contribute to the expression profiles, as discussed in more detail below.

The promoter region of *AtZIP4* contains two conserved *cis*-acting elements (RTGTCGACAY) known as zinc deficiency response elements (ZDREs) (Assunção *et al.*, 2010). One or two ZDREs are also found in the promoter regions of several other Zn transporter genes such as *ZIP1*, *ZIP3*, *ZIP4*, *ZIP5*, *ZIP9*, *ZIP12* and *IRT3*. These transporter genes are regulated by ZDRE binding transcription factors bZIP19 and bZIP23 (Assunção *et al.*, 2010), which therefore act as executive switches in the control of Zn homeostasis in *A. thaliana*. These transcription factors are important for the adaptation of *A. thaliana* under Zn deficiency conditions by promoting Zn uptake and transport when Zn is limiting.
The same ZDREs are also found in the *NcZNT1* promoter (*N. caerulescens*) and orthologous promoters in other species including *Cochleria pyrenaica*, *A. halleri* and *A. lyrata* (Talukdar, 2007). The function of the *NcbZIP19* and *NcbZIP23* transcription factors in *N. caerulescens* has not been studied in detail. Preliminary data suggest that the *N. caerulescens* *NcbZIP19* gene is repressed in roots but induced in shoots when plants are exposed to high Zn and Cd levels (data not shown). This suggests that *NcbZIP19* plays a role in metal homeostasis. When *N. caerulescens* is exposed to high metal concentrations, the inhibition of *NcbZIP19* in the roots prevents the uptake of excess metals. However, enhanced *NcbZIP19* expression in the shoots promotes the mobilization of metals that are already in the roots and translocates them to aboveground tissues. This matches the expression of *ZNT1*, which is regulated by the same bZIP factors and plays a role in the long-distance translocation of metals.

Additional functional experiments (such as knocking down *NcbZIP19* in *N. caerulescens* or attempting functional complementation of the *A. thaliana bzip19bzip23* double mutant with *NcbZIP19*) are required to clarify the role of *NcbZIP19* in hyperaccumulation and also to determine whether *NcbZIP19* is differentially regulated in *N. caerulescens* and *A. thaliana* as observed for *ZNT1/ZIP4*. Preliminary data indicate that *NcbZIP23* can be amplified by PCR from most of the *N. caerulescens* accessions we collected, but not from La Calamine, St. Félix de Pallières or Durfort (data not shown). This suggests that the *NcbZIP23* sequence may be poorly-conserved in these populations or that the gene may have been lost. Although *NcbZIP23* can be amplified from the GA accession, it was not isolated from our transcriptome set, suggesting it is expressed at minimal levels in this accession. If *NcZIP23* is confirmed to be absent or silent in *N. caerulescens*, it may imply its normal role in hyperaccumulation has been subsumed by *NcbZIP19*. The functional analysis of *NcbZIP19* in accessions with and without *NcbZIP23* will help to clarify the role of *NcbZIP23* in metal hyperaccumulation and homeostasis.

Although metal hypertolerance and hyperaccumulation are usually studied at the genetic level, epigenetic processes such DNA methylation, chromatin remodeling and regulation by small RNAs may also contribute to the regulation of genes required for plants to adapt to high levels of metal in the environment (Grativol *et al.*, 2012, Hanikenne&Nouet, 2011). However, the epigenetic control of metal homeostasis has not been investigated in detail. The microRNA miRNA398 was shown to be inhibited under excess Cd and Cu conditions and also
induced ROS detoxification and antioxidant accumulation, which can reduce the damage caused by high levels of Cu and Cd (Fig. 3 in Chapter 2). A comprehensive comparison of different \textit{N. caerulescens} accessions and related non-accumulator plants using techniques such as whole-genome bisulfite sequencing and small-RNAseq analysis will help to determine the potential role of these processes in hyperaccumulation.

\section*{The genomic analysis of metal hyperaccumulation}

The molecular analysis of metal hyperaccumulation in \textit{N. caerulescens} has relied heavily on sequence data and other resources from the laboratory model \textit{A. thaliana}, with which it shares 88.5\% sequence identity at the transcriptome level (Peer \textit{et al.}, 2003, Rigola \textit{et al.}, 2006). Although microarray-based transcript profiling studies revealed broad differences in gene expression that affect heavy metal homeostasis in \textit{N. caerulescens} (van de Mortel \textit{et al.}, 2006, van de Mortel \textit{et al.}, 2008), the limited number of probes on the array, the lack of sequence information, and the sequence divergence between \textit{N. caerulescens} and \textit{A. thaliana} prevented a detailed analysis of metal-regulated gene expression in \textit{N. caerulescens}. The first attempt to generate substantial transcriptome sequence data from \textit{N. caerulescens} (Rigola \textit{et al.}, 2006) involved the production of an expressed sequence tag (EST) database containing just over 3700 transcript sequences, but this resulted in the identification of only 35 metal-related genes. We therefore used a more prolific high-throughput next-generation sequencing technique to acquire a comprehensive set of transcript sequences (23836 isotigs) representing a 5.6-fold improvement over the previous research, including 198 known metal homeostasis genes (Chapter 5). This cDNA collection provides a firm foundation for a well-annotated \textit{N. caerulescens} genome sequence and will contribute to the functional analysis of genes related to metal hypertolerance and hyperaccumulation. Although an extensive set of \textit{N. caerulescens} expressed sequences was isolated from our transcriptome data, many known metal homeostasis genes were surprisingly absent, including ZIP3, ZIP5, ZIP7, ZIP8, ZIP12 and IRT2 from the ZIP gene family, and MTP2, MTP3, MTP4, MTP9 and MTP10 from the MTP gene family. However, these genes may not be expressed in the tissues or stages we sampled, or under the environmental conditions we chose for cultivation.

Although these \textit{N. caerulescens} transcript sequences will facilitate further studies of hypertolerance and hyperaccumulation, transcriptome data do not
provide intron sequences or intragenic sequences, which usually feature greater sequence diversity than transcripts. Such diversity would provide useful data for the analysis of genetic variation, gene regulatory regions, and metal adaptation, and this can only be provided by whole genome sequencing. The Zn/Cd tolerant, Zn/Cd hyperaccumulator accession Ganges has been selected for this purpose. The sequencing, assembly and annotation is nearly complete, but the advanced molecular evolution analysis and the annotation of repeats and putative microRNA sequences is still in progress.

The *N. caerulescens* Ganges genome is approximately 319 Mb in size, with 29,712 predicted genes representing 15,874 gene families. A comparison with other completed plant genome sequences (*A. thaliana*, *Ricinus communis*, *Theobroma cacao*, *Populus trichocarpa*, *Fragaria vesca*, *Carica papaya*, *Oryza sativa*, *Vitis vinifera* and *Thellungiella (Schrenkiella) parvula*) allows a preliminary phylogenetic tree to be constructed (Fig. 1). From these species, *S. parvula* is the most closely related one to *N. caerulescens*, followed by *A. thaliana*, the only other Brassicaceae. Whole genome duplication analysis using the 4DTv method to determine distance-transversion rates at four-fold degenerate sites of paralogous gene pairs (Huang et al., 2009) showed that the speciation event between *N. caerulescens* and *S. parvula* occurred after the genome duplication. The availability of the *N. caerulescens* whole genome sequence now allows comparisons with related non-accumulators that also have complete genome sequences, such as *Microthlaspi perfoliatum*, *A. thaliana* and *A. lyrata* and thus promotes the identification of metal hypertolerance/hyperaccumulation candidate genes for subsequent functional analysis. The genome will also contribute to the analysis of other metal hyperaccumulators from the Brassicaceae family, such as *A. halleri*. This is more difficult to sequence directly because it is self-incompatible and therefore a predominantly outcrossing species, making it difficult to acquire an inbred line for sequencing. For the similar challenge in potato and soybean, it was necessary to use a doubled haploid clone and a palaeopolyploid, respectively (Schmutz et al., 2010, Xu et al., 2011).

Closely-related hyperaccumulators such as the Zn/Cd hyperaccumulator *N. praecox* (Tolra et al., 2006) and the Ni hyperaccumulator *N. goesingense* (Salt, 2001) are also interesting species for comparative genomics. Based on the *N. caerulescens* Ganges whole genome sequence, the re-sequencing of other *N. caerulescens* populations differing in metal tolerance and accumulation capacity
will allow the genetic variation within and between populations to be assessed and correlated with phenotypic differences, allowing their evolutionary histories to be reconstructed. The *N. caerulescens* whole genome sequence will therefore contribute to the investigation of genetic variation, ecological adaptation and metal hypertolerance/hyperaccumulation in many plant species.

**Fig 1. Phylogenetic tree showing sequenced plant genomes, estimating the divergence times and substitution rates.** The blue numbers on the nodes represent the divergence time from present day (million years ago, Mya). The calibration time is the *A. thaliana* – *C. papaya* divergence (54–90 Mya) (red dot) as previously reported (Huang et al., 2009). The full names of the species are: *Noccaea caerulescens*, *Arabidopsis thaliana*, *Ricinus communis*, *Theobroma cacao*, *Populus trichocarpa*, *Fragaria vesca*, *Carica papaya*, *Oryza sativa*, *Vitis vinifera* and *Thellungiella* (Schrenkiella) *parvula*. 

179
References


References


Inoue H, Higuchi K, Takahashi M, Nakanishi H, Mori S, Nishizawa NK. 2003. Three rice nicotianamine synthase genes, OsNAS1, OsNAS2, and OsNAS3 are expressed in cells involved in long-distance transport of iron and differentially regulated by iron. The Plant Journal 36: 366-381.


References


Koch MA, German D. 2013. Taxonomy and systematics are key to biological information: Arabidopsis, Eutrema (Thellungiella), Noccaea and Schrenkiella (Brassicaceae) as examples. *Frontiers in Plant Science* **4**:267.


Morina F, Jovanovic L, Mojovic M, Vidovic M, Pankovic D, Veljovic Jovanovic S. 2010. Zinc-induced oxidative stress in Verbacum thapsus is caused by an
accumulation of reactive oxygen species and quinhydrone in the cell wall. Physiologia Plantarum 140: 209-224.


References


Sarret G., Harada E., Choi Y.-E., Isaure M.-P., Geoffroy N., Fakra S., Marcus M.A., Birschwilks, M., Clemens, S., Manceau, A., 2006. Trichomes of Tobacco excrete...


References


Shanmugam V, Tsednee M, Yeh K-C. 2012. ZINC TOLERANCE INDUCED BY IRON 1 reveals the importance of glutathione in the cross-homeostasis between zinc and iron in Arabidopsis thaliana. The Plant Journal 69: 1006-1017.


Steffens B, Sauter M. 2009. Epidermal cell death in rice is confined to cells with a distinct molecular identity and is mediated by ethylene and H$_2$O$_2$ through an autoamplified signal pathway. The Plant Cell 21: 184-196.


202
References


Van De Mortel JE, Schat H, Moerland PD, Van Themaat EVERL, Van Der Ent S, Blankestijn H, Ghandilyan A, Tsiatsiani S, Aarts MGM. 2008. Expression differences for genes involved in lignin, glutathione and sulphate metabolism in


van de Mortel JE. 2007. Heavy metal tolerance and accumulation in Thlaspi caerulescens, a heavy metal hyperaccumulating plant species.


Wong CKE, Cobbett CS. 2009. HMA P-type ATPases are the major mechanism for root-to-shoot Cd translocation in Arabidopsis thaliana. *New Phytologist* **181**: 71-78.


Zhang ZC, Chen BX, Qiu BS. 2010. Phytochelatin synthesis plays a similar role in shoots of the cadmium hyperaccumulator *Sedum alfredii* as in non-resistant plants. *Plant, Cell & Environment* 33: 1248-1255.


Zимери AM, Dhankher OP, McCaig B, Meagher RB. 2005. The plant MT1 metallothioneins are stabilized by binding cadmiums and are required for cadmium tolerance and accumulation. *Plant Molecular Biology* 58: 839-855.
Summary

Some plants can tolerate and accumulate unusually high levels of toxic metals, and the analysis of such plants can provide insight into the ecology of environments that are polluted with heavy metals due to human industrial activities. The study of heavy metal hyperaccumulators such as *Noccaea caerulescens* can show how plants cope with excess metals and increase their fitness when growing in metalliferous environments. In this thesis, I compared the molecular mechanisms of zinc (Zn) and cadmium (Cd) homeostasis in the hyperaccumulator species *Noccaea caerulescens* and its non-accumulator relative *Arabidopsis thaliana* by investigating the regulation of the *ZNT1/ZIP4* gene that promote Zn uptake, Zn/Cd tolerance, and perhaps also metal transport. I also studied the ecological advantages of metal hyperaccumulators in nature and screened an extensive set of transcripts related to metal homeostasis to find candidate genes that control metal hyperaccumulation and provide an evolutionary perspective to the emergence of this trait.

The functional characterization of the *N. caerulescens* and *A. thaliana* Zn-transporter genes *NcZNT1* and *AtZIP4* (and their promoters) showed how their differential expression pattern contributed to their role in metal tolerance and accumulation. The *NcZNT1* gene is induced by Zn deficiency and the NcZNT1 protein is localized in the plasma membrane. Transgenic *N. caerulescens* roots with the gene for green fluorescent protein (GFP) driven by the *NcZNT1* promoter revealed GFP fluorescence localized to pericycle and vascular tissues, suggests that *NcZNT1* facilitates metal loading into xylem and long-distance metal transport. The overexpression of *NcZNT1* in *A. thaliana* increased Zn and Cd tolerance and the capacity to accumulate these metals compared to wild-type plants. These results suggest that NcZNT1 plays an important role in Zn and Cd hypertolerance and hyperaccumulation in *N. caerulescens*, where it is expressed in both Zn-sufficient and Zn-excess conditions. The differential behaviour of the *NcZNT1* and *AtZIP4* promoters in *N. caerulescens* and *A. thaliana* implies that different *cis*-regulatory elements and *trans*-regulatory factors are found in both species.

The ecological advantage of hyperaccumulators in metal-contaminated soils was investigated by studying the *nas1* mutant, in which the *NAS1* gene is
disrupted by a transposon insertion. This allele was found in three natural *N. caerulescens* populations, which were compared to wild-type plants in terms of phenotype and adaptive advantage. Although the transposon disrupted the gene, the loss of activity was compensated by enhanced expression of *NAS3* and *NAS4*, resulting in the induction of nicotianamine (NA) production and therefore enhanced the Zn and Cd accumulation in the *nas1* mutants. This increased their metal sensitivity compared to wild-type plants, but also made them more toxic towards *Pieris rapae* caterpillars, which developed more slowly and gained less weight when fed on mutant plants exposed to excess Zn or Cd. Therefore, the possible selective advantage of the *nas1* mutant in nature is high metal accumulation and the protection of plants from herbivores. Differences in *nas1* allele frequency among the three natural populations suggests that *nas1* alleles experience different degrees of natural selection or may be at different stages on the route to fixation.

Molecular evolutionary studies involve the identification of candidate genes that play a role in adaptation. Therefore, a comprehensive set of transcript sequences was obtained from *N. caerulescens* accession Ganges (GA) by 454 pyrosequencing. In total, the collected 23,836 isotigs (putative transcripts) were grouped into 20,378 isogroups and 93.2% of them could be matched to a Brassicaceae protein sequences, which allowed preliminary functional annotation. A total of 87 isogroups was annotated as metal homeostasis related genes, including metal transporter families, metal chelator biosynthesis families, and metal tolerance gene families, which are candidate genes for the molecular analysis of heavy metal homeostasis mechanisms. A group of genes required for the synthesis of glucosinolates, which are important secondary metabolites that protect plants against herbivores, were also shown to be expressed in *N. caerulescens*. Only transcripts similar to *METALLOTHIONEIN3 (MT3)* were present as multiple copies in the *N. caerulescens* genome. These transcripts will provide an important tool to annotate the sequences that become available following whole-genome sequencing of the *N. caerulescens* Ganges accession. This genome sequence will also be the basis of genome comparisons between the different *N. caerulescens* accessions that have different levels of metal accumulation and metal tolerance and which may therefore differ in terms of gene expression levels of modes of actions that affect metal homeostasis.
Om de ecologie van gebieden te begrijpen die vooral door menselijk ingrijpen verontreinigd zijn met zware metalen, is er een toenemende interesse om te weten te komen op welke wijze sommige planten tolerant zijn tegen hoge concentraties van dergelijke giftige metalen en hoe ze deze metalen soms ook in hoge mate kunnen ophopen.

De studie van zware metaal-accumuleerders zoals *Noccaea caerulescens* (zinkboerenkers) kan helpen bij het begrijpen van de manier waarop planten omgaan met een overmaat aan metalen in de grond en hoe de planten hun ‘fitness’ kunnen vergroten wanneer ze op metaalhoudende gronden groeien.

In dit proefschrift heb ik de verschillende moleculaire mechanismen van zowel Zn- als Cd- homeostase van de metaal hyperaccumuleerder *Noccaea caerulescens* vergeleken met die van de niet accumulator *Arabidopsis thaliana* (zandraket) door de regulatie te bestuderen van het *ZNT1/ZIP4* gen, dat de Zn opname en de tolerantie voor Zn en Cd en mogelijk ook hun transport bevordert. Daarnaast werd het ecologische voordeel van deze metaal hyperaccumuleerders in de natuur onderzocht. Vervolgens werd een grote verzameling van transcripten die mogelijk iets te maken hebben met metaalhomeostase geanalyseerd met de bedoeling kandidaatgenen te vinden die de hyperaccumulatie van metalen bepalen. Hierbij werden ook de evolutionaire aspecten van de eigenschap betrokken.

De functionele karakterisering en promoter vergelijking van de voor een Zn transporter coderende genen *NcZNT1* en *AtZIP4* werden gedaan in *N. caerulescens* en *A. thaliana* met de bedoeling hun verschillende expressepatroon te begrijpen en daarmee ook hun rol in metaal tolerantie en -accumulatie. Expressie van het *NcZNT1* gen wordt geïnduceerd door Zn deficiëntie en het eiwit is gelokaliseerd in het plasmamembraan. Transgene wortels van *N. caerulescens* met het *NcZNT1 promoter::GFP* construct bevatten, lieten GFP localisatie zien in de pericycel en in het vaatbundelweefsel. Dit suggereert dat NcZNT1 betrokken is bij het opnemen van metalen in het xyleem waardoor ze over grotere afstanden in de plant getransporteerd kunnen worden. De overexpressie van *NcZNT1* in *A. thaliana* verhoogde de Zn- en Cd-tolerantie, evenals het vermogen om Zn en Cd in de planten op te hopen, in vergelijking met niet getransformeerde wild type planten. Deze resultaten suggereren dat NcZNT1 een belangrijke rol speelt bij Zn- en Cd-
Samenvatting

hypertolerantie en bij het ophopen van metalen in *N. caerulescens*, waar dit gen tot expressie komt wanneer de planten opgroeien onder normale en onder overmaat Zn condities. De verschillen in promoter activiteit van *NcZNT1* en *AtZIP4* tussen *N. caerulescens* en *A. thaliana* kunnen waarschijnlijk verklaard worden door verschillen in *cis-* en in *trans*-regulatiemechanismen.

Om het ecologische voordeel van metaal-hyperaccumuleerders te begrijpen wanneer deze op sterk met metalen verontreinigde gronden groeien, werd een *nicotineamine synthase (nas1)* mutant geïdentificeerd in drie natuurlijke *N. caerulescens* populaties. Deze mutant, waarbij de mutatie veroorzaakt werd door een transposon-insertie in het gen, werd vergeleken met wildtype *NAS1* planten (WT). Hoewel het transposon in de *nas1* mutant de functie van het *NAS1* gen verstoorde, bleek dit effect te worden gecompenseerd door verhoogde expressie van de *NAS3* en *NAS4* genen, hetgeen leidde tot hogere nicotineamine (NA) productie. Dit verklaarde een grotere Zn- en Cd-ophoping, waardoor de planten gevoeliger werden voor hoge concentraties van Zn en Cd dan de WT-planten. Rupsen van *Pierus rapae* (koolwitje) die gevoed werden door de *nas1* mutant, vertoonden een verminderde groei en vertraagde ontwikkeling in vergelijking met het wildtype wanneer de planten opgekweekt werden op verhoogde Zn- en Cd-concentraties. Daarom werd geconcludeerd dat *nas1* mutanten in de natuur een voordeel kunnen hebben, omdat ze meer metalen ophopen en omdat ze de plant beschermen tegen herbivoren. De verschillen in allele frequenties tussen de drie natuurlijke populaties suggereren dat de *nas1* mutanten een selectiedruk ondergaan die verschilt tussen de populaties of dat ze in een verschillend stadium zijn van door selectie gestuurde vervanging van het wildtype allele door het mutante *nas1* allele wat een ‘selective sweep’ genoemd wordt.

Moleculaire en evolutoirne studies maken gebruik van gensequenties van die genen die een rol spelen bij de adaptatie van organismen aan hun omgeving. Daarom werd een uitgebreide verzameling van DNA sequenties van transcriptieproducten geanalyseerd. Deze transcripten werden verkregen van de *N. caerulescens* accessie Ganges (GA) met behulp van 454 pyrosequencering. Het totaal van 28.836 isotigs (waarschijnlijke transcripten) werd gegroepeerd in 20.378 isogroepen waarvan 93.2 % op een bekende eiwitsequentie van een Brassicaceae-soort leek, wat een inzicht gaf in de functies van deze *N. caerulescens* genen. 87 isogroepen konden worden beschreven als genen die te maken hebben met metaalhomeostase, waaronder metaaltransportgenen, genen voor het maken van verbindingen die metaal binden (chelators) en genen die metaaltolerantie kunnen bepalen. Deze zijn daarom kandidaatgenen voor studies van zware
Samenvatting

metaalhomeostase. Daarnaast werd een groep genen geïdentificeerd, die betrokken is bij de biosynthese van glucosinolaten. Deze verbindingen kunnen planten beschermen tegen herbivoren. Van alle genen die horen bij bovengenoemde klassen werd alleen voor METALLOTHIONEIN3 (MT3) een extra genkopie gevonden in N. caerulescens. Deze sequenties van transcripten zullen een belangrijk hulpmiddel zijn bij het volledig in kaart brengen van het genoom van de N. caerulescens accessie Ganges. Deze genoomsequentie zal ook de basis zijn van genoomvergelijkingen met verschillende N. caerulescens accessies die verschillen in de mate van metaalophoping en waarvan verwacht wordt dat ze verschillen hetzij in het niveau van expressie hetzij in het mechanisme van genen betrokken bij metaalhomeostase.
Acknowledgements

Five years ago, I came to Wageningen with nothing but Jimmy; five years later, I was fully filled with your friendship and kindness. After another five years, I may have forgotten how I have run all the experiments in this thesis, but I will never forget you all and the priceless time in Wageningen.

Dear Dr. Mark Aarts, dear supervisor, I express my deepest gratitude to you for your supervision on my PhD project and kindly supporting me to go to many courses and meetings in the past five years. This enabled me to develop useful skills and connections. I also very much appreciate your patience to give comments and to revise my poor English. In addition, I enjoyed your hospitality and very much liked those events at your sweet house, BBQ, Sinterklaas, and the international dinners, at which I had a lot of fun and a good time with your family and my colleagues.

Dear Prof. Dr. Maarten Koornneef, dear promotor, I’m very sincerely grateful to you for the kind and warm help you gave me in the past five years. I very much like the way you are getting along with us, like a friend or a kindly grandfather, which made me feel comfortable. I especially thank you to put all the efforts on helping me and encouraging me very positively at the very last and the most difficult moments, when I was writing my thesis. Without your help and support, I for sure would not have been able to accomplish my aim and submit my thesis. I would also like to extend my gratitude to your wife, Elly, for her kind help on the translation of my Dutch summary.

I warmly appreciate my external supervisor, Dr. Henk Schat, for the kind and patient supervision on metal concentration measurements and statistics analysis (my weakest part). I want to express my appreciation to Prof. Dr. Bas Zwaan, who always gave me very useful comments and helped me a lot on negotiation. I would like to acknowledge several collaborators: many thanks to Prof. Dr. Richard Immink, who helped me to design the constructs for protein expression and always sends warm greeting to us; to Dr. Cezary Smaczniak, who very patiently taught me how to perform EMSA experiments; to Prof. Dr. Joop van Loon, who very professionally and carefully performed caterpillar and butterfly experiments with me (my first experience with this insect, very interesting); to Prof. Dr. Stephan Clemens and Dr. Holger Schmidt, who helped me on the measurement of Nicotianamine; to Dr. Wilco Lighterink, Dr. Bas Dekkers, Leo Willems, and Dr. Ronny Joosen, who helped me a lot on performing qPCR and germinating Noccaea
Acknowledgements

*caeruleascens* seeds; to Prof. Dr. Joost Keurentjes, who joined our literature discussion and kindly gave us useful advice; to Peter de Gijsel and Liya Yi, who helped for the measurement of nitrogen/proteins; to Dr. Norbert de Ruijter, who helped a lot on the use of microscopy; to Dr. Elio Schijlen, who helped on the transcriptome analysis; to Dr. Arja Tervahauta and Pauliina Halimaa, who put great efforts on publishing our co-author paper; to Prof. Dr Sirpa Kärenlampi, who helped on checking the whole thesis very carefully and give many valuable comments. Without any of you, I could not have made the thesis as it is now.

I’m indebted to my colleagues at the Laboratories of Genetics and Plant Physiology, Ana Carolina Atala Lombelo Campos (a huge), Ana G. L. Assunção, Aina Prinzenberg, Anke Kristin Trilling, Ariadne Ribeiro Henriques, Bertha Koopmanschap, Charles Neris Moreira, Desalegn Etalo, Diaan Jamar, Diana Zuluaga, Edouard Severing, Eric Bastiaans, Esmer Jongedijk, Frank Becker, Hedayat Bagheri, Johanna Molenaar, Jose van de Belt, Lidiya Sergeeva, Liu Hong, Mina Tomaz Villafort Carvalho, Mohamed El Soda, Myriam Olortegui Guzman, Natalia Carreno Quintero. Nihal Erol-Oztolan, Pádraic Flood (our best man), Pingping Huang, Rik Kooke, Roxanne van Rooijen, Thais Córdova, Vanessa Stein, Yanli Wang, Zeshan Hassan. Thank you a million to help me a lot on experiments and life. Especially thanks to Corrie Hanhart, who supervise me the first experiment at lab and kindly deal with my thousands of questions; and to Wytske Nijenhuis and Rina Anthonijsz, who always act like a mother to send regards and care for me a lot.

I appreciate the efforts of many students that have ever been involved in this project and helped me a lot. I am thankful for MSc students, Nguyen Duy, Sam Van Es, and Alfred Arulandhu; BSc students, René Boesten, Niels van der Last, Annemiek Stegehuis, Diede Maas, and Allison Luger; internship student, Erika; and guest researchers, Aiyen and Chia-Kai Kang.

To my dear Taiwanese friends, Chang-Lin Chen, Chenling Huang (Jenny), Che-Yang Liao, Chia-chi Chang, Chia-Kai Kang, Chia-Yi Liu, Chia-Ying Lin, Chih-Chieh Chuang, Ellen Lin, Han-Ni Tang, Hsiao-Li Hsu, Hsi-Yu Chen, Hsuan Chen, I-Chiao Lee, Jacourr Liu, Jen-Jung Liu, Jhan-Ye Wang, Jia-Yuh Lee (Eunice), Jie-Hong Huang (Red), Jo-Chien Liao (Wendy), Jou-Fu Wang, Lilly Hsiao, Michelle Yeh, Meng-Yuan Jen, Wei-Shan Chen (MoMo), Shang-Jung Lin, Shu-Chen Liu (Phabbie), Tzeyi, Huang, Wen-Kai Feng, Yu-Fang Chen, Yun-Shan Liu, Yun-Yueh Lu, Yu-Wei Chang, I want to express my appreciation for your company at Wageningen and for those great & funny times we had. My sincere thanks to Pengying Piena-Yang for kindly taking care of us as your children and always preparing very nice food for us; to Ya-Ping Chang for your kind help on the cover design; and to Song-Bin Chang for your kind help on our application to the PhD at Wageningen University.
To my Chinese friends, Bing Bai, Bo Wang, Chenlei Hua, Chunxu Song, Feng Zhu, Guodong Wang, Hanzi He, Huchen Li, Hui Li, Jianhua Zhang, Jimeng Li, Jun He, Ke Lin, Lemeng Dong, Liping Gao, Lisha Zhang, Na Li, Ningwen Zhang, Qing Liu, Suxian Zhu, Ting Yang, Tingting Xiao, Wei Gao, Wei Liu, Wei QinWei Song, Weicong Qi, Xi Chen, Xi Cheng, Xianwen Ji; Xiaomin Tang, Xu Cheng, Yan Wang, Yanli Wang, Yanxia Zhang, Ying Zhang, Yuanyuan Zhang, Yumeng Zhang, I am thankful for all your support and kindness. I will never forget how delicious of steamed dumplings you made and how interesting of the games, Dunkey and Legends of the Three Kingdoms, we have ever played.

To my friends at the Student Chaplaincy, Rev. Josine, Fr. Wiel, Ingeborg Brouwer, Alexandre Villela, Danielle Lucas Barbosa, and Robin Liu, I’m so grateful for the time we had in Taizé, for the sing at a prison, for the great food at global dinner, for the good wishes you delivered to us, and for the peace we prayed. Thank you very much from the very bottom of my heart.

To my Malaysia friends, Ameen, Izan, Fatimah, Naim, Yani, Asyraf, Ilah, Razak, Nozie, Huda, Su, Aniey, Siew Ling Chong, and Apple Teo, it’s generous of you to make so many nice Malaysia Food in each gathering, and to be thoughtful for me and Yew-yew. Thank you for everything you have done. To my “Malaysia Match Group”, Shi Pey Wong, Loo Wee Chia, Freddy Yeo, Meng Foong Cheon, and Rui Jack Chong, saying thank you is not enough for our friendship and those wonderful time we have being together. “I Love You” guys, hope to see you soon.

I express my heartiest gratitude to my beloved parents, An-Ching Lin (老爸) and Su-Chiu Huang (老妈). Thank you for the unlimited love and for standing behind me in all that I do. To my sisters and their husbands, Mei-Hsiu Lin (小美人) & Jen-Cheng Liao, Yi-Chun Lin (小貝) & Chih-Yung Wu (火鍋), Feng-Ching Lin (一匹野馬晴); and my brother, Yen-Ming Lin (黑人), I would like to acknowledge you for the strong support at my back and to take care of papa and mama when I was not with them. I am also thankful for my aunt, Tzu-Chu Lin (二姑), who never miss any important moment in my life, no matter where it is.

At last but not the least, to my dearly beloved husband, Jimmy, and my lovely Yew-yew, I love you more than words can say. We share so much together, no matter it’s happiness or disappointed. Your loving and accompany makes the trouble disappear. Thank you from my deepest heart.

My apologies and thanks also to those I may forget to put your names in this acknowledgement.

Ya-Fen Lin

Wageningen, March 4, 2014
Ya-Fen Lin was born in Taipei, Taiwan, on July 25, 1980. She started her academic studies in the field of Agronomy in 1999. In 2002, she joined Prof. Dr. Li-Fei Liu’s research group and soon became involved in a genetically modified organism (GMO) detection project, which was supported by The Council of Agriculture, Executive Yuan, Taiwan. She acquired her B.Sc. degree in Agronomy, National Taiwan University in 2003. Soon thereafter, she continued her M.Sc. research in the same group and worked on the topic, “Functional studies of the genes related to rice (Oryza sativa) spikelets” by using a stunted lemma/palea (slp/slp) rice mutant. After obtaining her MSc major in crop science, Agronomy, National Taiwan University in 2005, she continued on her career path in research and became a research assistant in Dr. Kuo-Chen Yeh’s laboratory at the Agricultural Biotechnology Research Center (ABRC), Academia Sinica, Taipei, Taiwan. Her research focused on studying heavy metal-related transporters and chaperones in Arabidopsis thaliana. In 2007, she started working as PhD student at the Institute of Plant Biology at National Taiwan University in Taipei, where she was rotating with Dr. Hsieh, Hsu-Liang and learning how to analyse light signalling in seedlings, including growth analysis in red, far red, and blue light. In 2008, she joined the group of Dr. Mark Aarts and Prof. Dr. Maarten Koornneef at the Laboratory of Genetics, Wageningen University, The Netherlands, and worked on the PhD thesis entitled, “An evolutionary perspective on differential regulation of zinc and cadmium homeostasis genes in Arabidopsis thaliana and Noccaea caerulescens”.
Publications

Journal Papers


  * Fig 2a was selected as the cover of the journal.

- Pauliina Halimaa, **Ya-Fen Lin**, Viivi Ahonen, Daniel Blande, Stephan Clemens, Attila Gyenesei, Elna Häikiö, Sirpa Kärenlampi, Asta Laiho, Marc Aarts, Juha-Pekka Pursiheimo, Henk Schat, Holger Schmidt, Marjo Tuomainen, Arja Tervahauta (2014) Gene expression differences between *Noccaea caerulescens* ecotypes help identifying candidate genes for metal phytoremediation. (accepted by Environmental Science & Technology)


- Zeshan Hassan*, **Ya-Fen Lin***, Sangita Talukdar*, Liu Hong, Alfred Arulandhu, Henk Schat and Mark G.M. Aarts (2013) The expression of the ZNT1 zinc transporter from the metal hyperaccumulator *Noccaea caerulescens* confers
enhanced zinc and cadmium tolerance and accumulation to *Arabidopsis thaliana*. (in preparation for New Phytologist)


**Thesis**

Education Statement of the Graduate School
Experimental Plant Sciences

Issued to: Ya-Fen Lin
Date: 24 March 2014
Group: Genetics, Wageningen University & Research Centre

1) Start-up phase

- **First presentation of your project**
  An evolutionary perspective on differential transcriptional regulation of zinc homeostasis genes in *A. thaliana* and *T. carnea*. Mar 02, 2009
- **Writing or rewriting a project proposal**
- **Writing a review or book chapter**
- **MSc courses**
- **Laboratory use of isotopes**
  Safe handling of Radioactive isotopes, level 5B Jul 2010

| Subtotal Start-up Phase | 9.9 credits |

2) Scientific Exposure

- **EPS PhD student days**
  EPS PhD Student Day, Naturis Museum, Leiden, NL Feb 26, 2009
  EPS PhD Student Day, Utrecht University, NL Jun 01, 2010
  EPS PhD Student Day, Wageningen UR, NL May 20, 2011
  ExpectationS Day (EPS Career Day) Nov 18, 2011
- **EPS theme symposia**
  EPS theme 4 ‘Genome Plasticity’, Wageningen UR, NL Dec 12, 2008
  EPS theme 2 ‘Interactions between Plants and Biotic Agents’, Utrecht University, NL Jan 22, 2009
  EPS theme 3 ‘Symposium Metabolism and Adaptation’, University of Amsterdam, NL Feb 19, 2010
  EPS theme 1 ‘Developmental Biology of Plants’, Wageningen UR, NL Jan 28, 2010
  EPS theme 3 ‘Symposium Metabolism and Adaptation’, Leiden University, NL Feb 19, 2010
  EPS theme 3 ‘Symposium Metabolism and Adaptation’, Utrecht University, NL April 26, 2012
- **NWO Lunteren days and other National Platforms**
  ALW meeting ‘Experimental Plant Sciences’, Lunteren, NL Apr 06-07, 2009
  ALW meeting ‘Experimental Plant Sciences’, Lunteren, NL April 19-20, 2010
  ALW meeting ‘Experimental Plant Sciences’, Lunteren, NL April 04-05, 2011
  ALW meeting ‘Experimental Plant Sciences’, Lunteren, NL Apr 02-03, 2012
- **Seminars (series), workshops and symposia**
  Invited seminar SyS ‘Nitrogen’ Nov 27, 2008
  Invited seminars Otto Leyser, Hiro Nonogaki Aug-Sep 2009
  Plant Physiology / Genetics seminar series Sep 2009-Nov 2012
  Genetics (USS) seminar series Sep 2009-Nov 2013
  Invited seminars Nick Panopoulos, Justin Borevitz, Laurent Zimmerli, Chris Hawes, John Yoder, David Baulcombe, Kirsten Bomblies, Jose Jimenez-Gomez, Alex Peczika Jan-Nov 2010
  Ecogenomics Day 2010 Apr 21, 2010
  Invited seminar Javier Puljak Aug 25, 2011
  Invited seminars (Paul Fraser, Rainer Melzer, Inez Stam-Loodt, Dorgi Wagner, David G. Robinson) Feb-Sep 2012
- **Seminar plus**
- **International symposia and congresses**
  FESPB 2010 (Valencia, Spain) Jul 04-09, 2010
  LOCOMET meeting 2012 (Lille, France) Jul 04-06, 2012
  ASPD 2012 (Austin, Texas) Jul 20-24, 2012
- **Presentations**
  Oral, Ecogenomics Day 2010 Apr 21, 2010
  Oral, FESPB 2010, Valencia, Spain Jul 04-09, 2010
  Oral, ALW meeting ‘Experimental Plant Sciences’, Lunteren Apr 04-05, 2011
  Oral, EPS theme 3 ‘Metabolism and Adaptation’ Apr 26, 2012
  Poster, ASPI 2012, Austin, Texas Jul 28-24, 2012
- **Lab interview**
  An interview with a member of the International Advisory Board Nov 14, 2012
- **Excursions**
  Green Life Science Company Visit at Seed Valley, ENZA and INCOTEC (Eekhuiizen, NL) Jun 23, 2011

| Subtotal Scientific Exposure | 20.8 credits |

3) In-Depth Studies

- **EPS courses or other PhD courses**
  Postgraduate course ‘Confolical Microscopy’ Jan 26-30, 2009
  Summer school ‘Mineral nutrition in photosynthetic organism: molecular, physiological and ecological aspects’ Jun 17-20, 2009
  Postgraduate course, Spring School ‘RNA & The world of small RNA molecules’ Apr 14-16, 2010
  Postgraduate course ‘Transcription Factors’ May 09-11, 2011
  Postgraduate course ‘Bioinformatics: A User’s approach’ Aug 29-Sep 2, 2011

CONTINUED ON NEXT PAGE
### Education statement

<table>
<thead>
<tr>
<th>Postgraduate course 'Basics of Parameter Estimation'</th>
<th>Feb 10, 2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>Journal club</td>
<td></td>
</tr>
<tr>
<td>Literature discussion Plant Genetics group</td>
<td></td>
</tr>
<tr>
<td>Individual research training</td>
<td></td>
</tr>
</tbody>
</table>

**Subtotal In-Depth Studies**: 8.8 credits*

### 4) Personal development

<table>
<thead>
<tr>
<th>Skill training courses</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project- and Time Management</td>
<td>Mar-Apr 2010</td>
</tr>
<tr>
<td>Academic writing I</td>
<td>Sep-Jan 2011</td>
</tr>
<tr>
<td>Academic writing II</td>
<td>Mar-Jun 2011</td>
</tr>
<tr>
<td>Techniques for writing and presenting a scientific paper</td>
<td>Jan 05-08, 2011</td>
</tr>
<tr>
<td>Stress identification and Management</td>
<td>Apr 05, 2012</td>
</tr>
<tr>
<td>Organization of PhD students day, course or conference</td>
<td></td>
</tr>
<tr>
<td>Organizer of literature discussion</td>
<td>Apr 2009-Aug 2012</td>
</tr>
<tr>
<td>Membership of Board, Committee or PhD council</td>
<td></td>
</tr>
</tbody>
</table>

**Subtotal Personal Development**: 9.3 credits*

**TOTAL NUMBER OF CREDIT POINTS**: 47.7

Hereewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS which compiles of a minimum total of 30 ECTS credits.

* 4 credit represents a normative study load of 28 hours of study.
This project was financially supported by the graduate school Experimental Plant Sciences (EPS), Wageningen University, The Netherlands

Layout and cover design: Ya-Fen Lin
Printed by Uitgeverij BOXPress || Proefschriftmaken.nl ||