New Phytologist

Research

Adaptation to high zinc depends on distinct mechanisms in metallicolous populations of *Arabidopsis halleri*

M. Sol Schvartzman[1], Massimiliano Corso[2], Nazeer Fataftah[1], Maxime Scheepers[1], Cécile Nouet[1], Bernard Bosman[3], Monique Carnol[3], Patrick Motte[1], Nathalie Verbruggen[2] and Marc Hanikenne[1]

[1]InBioS-PhytoSystems, Functional Genomics and Plant Molecular Imaging, University of Liège, Liège B-4000, Belgium; [2]Physiology and Plant Molecular Genetics, Free University of Brussels, Brussels 1050, Belgium; [3]Laboratory of Plant and Microbial Ecology, Department of Biology, Ecology and Evolution, University of Liège, Liège B-4000, Belgium

Author for correspondence:
Marc Hanikenne
Tel: +32 43663844
Email: marc.hanikenne@uliege.be

Received: *8 September 2017*
Accepted: *14 November 2017*

New Phytologist (2018)
doi: 10.1111/nph.14949

Key words: *Arabidopsis halleri*, divergent evolution, hyperaccumulation, intraspecific variation, iron deficiency response, reference transcriptome, zinc.

Summary

• Zinc (Zn) hyperaccumulation and hypertolerance are highly variable traits in *Arabidopsis halleri*. Metallicolous populations have evolved from nearby nonmetallicolous populations in multiple independent adaptation events. To determine whether these events resulted in similar or divergent adaptive strategies to high soil Zn concentrations, we compared two *A. halleri* metallicolous populations from distant genetic units in Europe (Poland (PL22) and Italy (I16)).

• The ionomic (Inductively Coupled Plasma Atomic Emission Spectroscopy (ICP-AES)) and transcriptomic (RNA sequencing (RNA-Seq)) responses to growth at 5 and 150 μM Zn were analyzed in root and shoot tissues to examine the contribution of the geographic origin and treatment to variation among populations. These analyses were enabled by the generation of a reference *A. halleri* transcriptome assembly.

• The genetic unit accounted for the largest variation in the gene expression profile, whereas the two populations had contrasting Zn accumulation phenotypes and shared little common response to the Zn treatment. The PL22 population displayed an iron deficiency response at high Zn in roots and shoots, which may account for higher Zn accumulation. By contrast, I16, originating from a highly Zn-contaminated soil, strongly responded to control conditions.

• Our data suggest that distinct mechanisms support adaptation to high Zn in soils among *A. halleri* metallicolous populations.

Introduction

Plants have developed different strategies to cope with high metal concentrations, usually considered phytotoxic, in soils. Shoot metal hyperaccumulation is a naturally selected extreme trait that represents one of these strategies (Bert *et al.*, 2002; Verbruggen *et al.*, 2009; Krämer, 2010; Hanikenne & Nouet, 2011; Merlot *et al.*, 2018). Metal hyperaccumulator plants are valuable models in which to examine the evolution of a complex trait and reveal key players in metal distribution and tolerance in plant tissues.

Arabidopsis halleri is a metal-hyperaccumulating species with an extraordinary potential to take up and accumulate up to *c.* 5.4% zinc (Zn) and/or *c.* 0.36% cadmium (Cd) as a percentage of dry biomass in shoots (Bert *et al.*, 2000; Dahmani-Muller *et al.*, 2000; Krämer, 2010; Corso *et al.*, 2017; Stein *et al.*, 2017). Quantitative genetic and transcriptomic studies allowed identification of molecular actors underlying hyperaccumulation and hypertolerance in *A. halleri* (Verbruggen *et al.*, 2009; Krämer, 2010; Hanikenne & Nouet, 2011; Merlot *et al.*, 2018). A number of major candidate genes, involved in metal transport and metal chelator synthesis and all constitutively highly expressed in *A. halleri* compared with its nonaccumulator relatives *Arabidopsis thaliana* and *Arabidopsis lyrata*, were further characterized functionally (e.g. Dräger *et al.*, 2004; Hanikenne *et al.*, 2008; Shahzad *et al.*, 2010; Deinlein *et al.*, 2012; Baliardini *et al.*, 2015; Charlier *et al.*, 2015). Briefly, several Zinc-Regulated Transporter Iron-Regulated Transporter-Like Protein (ZIP) transporters presumably contribute to enhanced rates of root metal uptake, enhanced Zn radial transport towards the root xylem and Zn distribution in shoots (Talke *et al.*, 2006; Krämer *et al.*, 2007; Lin *et al.*, 2009). High expression of *Nicotianamine Synthase 2* (*NAS2*) results in elevated concentrations of the metal chelator nicotianamine (NA) in roots of *A. halleri* compared with *A. thaliana* and is suggested to favor Zn symplastic mobility towards the xylem (Weber *et al.*, 2004; Deinlein *et al.*, 2012; Cornu *et al.*, 2015). In *A. halleri*, Zn (and Cd) loading into the xylem is driven by the Heavy Metal ATPase 4 (HMA4) protein (Talke *et al.*, 2006; Courbot *et al.*, 2007; Hanikenne *et al.*, 2008; Nouet *et al.*, 2015), which is a plasma membrane P-type ATPase pump (Hussain *et al.*, 2004; Hanikenne & Baurain, 2014). High expression of *HMA4* is required for both hyperaccumulation and hypertolerance of Zn and Cd (Talke *et al.*, 2006; Courbot *et al.*, 2007; Hanikenne *et al.*, 2008; Frérot *et al.*, 2010; Willems *et al.*, 2010). Increased gene dosage of *HMA4* was selected during evolution of *A. halleri* (Hanikenne *et al.*, 2008, 2013). HMA4 possibly contributes to metal tolerance by enabling metal storage in

shoot tissues. Metal Tolerance Protein 1 (MTP1) is probably responsible for Zn storage in shoot vacuoles, providing Zn hypertolerance (Dräger *et al.*, 2004; Talke *et al.*, 2006; Willems *et al.*, 2007; Shahzad *et al.*, 2010).

Arabidopsis halleri is distributed in Europe and in eastern Asia where it notably establishes populations in (former) industrial sites with metal-polluted soils (metallicolous (M) populations). Even though Zn hyperaccumulation is suggested to be a species-wide trait in *A. halleri*, substantial intraspecific variation is observed among M populations (Bert *et al.*, 2000, 2002; Meyer *et al.*, 2015; Stein *et al.*, 2017). The European populations of *A. halleri* are distributed in two allopatric genetic units (GUs), named the North-West and South-East zones, and a hybrid zone (HZ) linking the two GUs (Pauwels *et al.*, 2012). The colonization of anthropogenic metal-contaminated sites from nearby nonpolluted sites (nonmetallicolous (NM) populations) is thought to have taken place fairly recently and independently within each GU. Hypertolerance and hyperaccumulation properties of metallicolous populations thus potentially evolved using distinct genetic mechanisms (Meyer *et al.*, 2010, 2011, 2015; Pauwels *et al.*, 2012). It is currently unknown whether these recent events of adaptation relied on fine-tuning (e.g. allelic variation) of mechanisms that contribute to the species-wide hyperaccumulation and hypertolerance traits (e.g. HMA4 and ZIP) or whether they required additional genes. A recent study suggests that, while HMA4 contributes to Zn tolerance in both M and NM populations of *A. halleri* (Hanikenne *et al.*, 2013), the function of MTP1 in Zn tolerance may have evolved later in M populations (Meyer *et al.*, 2016).

To examine the mechanisms underlying intraspecific phenotypic variation in *A. halleri*, two geographically distant M populations from the HZ (southern Poland) and South-East (northern Italy) GUs were compared at the ionomic and transcriptomic levels. This analysis was enabled by building a reference transcriptome assembly for *A. halleri*. Our data reveal that the two M populations display similar Zn tolerance but distinct Zn hyperaccumulation properties. They also evolved distinct strategies to adjust Zn and iron (Fe) homeostasis when exposed to high Zn concentrations.

Materials and Methods

Plant material and experimental design

Arabidopsis halleri (L.) O'Kane & Al-Shehbaz seeds were collected in the field (summer 2015) in Poland (PL22) and Italy (I16) (Pauwels *et al.*, 2012; Corso *et al.*, 2017).

PL22 and I16 seeds were sown on vermiculite in a controlled growth chamber (16 h light d^{-1}; 100 µmol photons m^{-2} s^{-1} irradiance; 20°C : 18°C, day : night; 70% humidity) (Corso *et al.*, 2017). Five week-old plants were transferred to 4-l hydroponic trays containing modified Hoagland medium (Talke *et al.*, 2006) with 5 µM $ZnSO_4$ (control conditions) and were grown in a glasshouse (with supplemented light at 100 µmol photons m^{-2} s^{-1} irradiance in a 16 h 21°C : 8 h 17°C, light : dark regime). Plants of each population were distributed randomly in trays, which were randomly moved in the glasshouse every week. After 6 wk of growth, the Zn concentration in half of the trays (20–24 plants per population) was increased to 150 µM $ZnSO_4$ (treatment conditions) for 2 wk, whereas the other half was maintained in control conditions. The nutrient solution was exchanged for fresh medium every 3 d, and 3 d before harvest. Relative chlorophyll content (in arbitrary units) was measured every 3 d using a CCM-200plus device (Opti-Sciences, Hudson, NH, USA). Each measurement was obtained from 10 randomly selected plants for each population/condition and data are the average of three leaves of different ages from each plant.

After the 2 wk of treatments, three biological replicates (pools of six to eight individuals per population/growth condition) were harvested. Root tissues were separated from shoots and blotted dry, and each tissue was quickly weighed. For each sample, half of the material was immediately frozen in liquid nitrogen and stored at −80°C. The other half of the material was desorbed and used for mineral analysis by Inductively Coupled Plasma Atomic Emission Spectroscopy (ICP-AES) (Nouet *et al.*, 2015). Samples for citrate determination were prepared as follows: 100 mg of frozen tissues was mixed with 1 ml of distilled water. The pH was adjusted to 7–8 with 1 M KOH, 100 µl of perchloric acid was added and citrate was measured using a citric acid assay kit (BioSentec, Auzeville-Tolosane, France).

RNA extraction, library preparation and Illumina sequencing

Total RNAs were extracted from tissues with the Maxwell® 16 LEV Plant RNA Kit (Promega, Leiden, the Netherlands). Libraries for RNA sequencing (RNA-Seq) were prepared from 1 µg of total RNAs with the TruSeq Stranded mRNA Library Prep Kit (Illumina, San Diego, CA, USA), multiplexed and sequenced with an Illumina NextSeq500 device (high throughput mode; 75-bp paired-end reads) yielding on average *c.* 35 million reads per sample.

Reference transcriptome *de novo* assembly, functional annotation and quality assessment

Read quality was assessed using FastQC (v.0.10.1, http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Quality trimming and removal of adapters were performed using Trimmomatic (v.0.32; Bolger *et al.*, 2014). Using Trinity with default settings (v.20140717; Grabherr *et al.*, 2011), a *de novo* assembly of the *A. halleri* transcriptome (Supporting Information Fig. S1) was carried out including 283 million reads corresponding to one biological replicate of root and shoot samples (six to eight plants each) from control, Zn-exposed (this study) and Cd-exposed (Corso *et al.*, 2017) PL22 plants (*c.* 90 individuals). This initial assembly was subjected to filtering using the EvidentialGene tr2aacds pipeline (v.2013.03.11; http://arthropods.eugenes.org/EvidentialGene/about/EvidentialGene_trassembly_pipe.html) with default parameters.

The filtered contigs were annotated using reciprocal best hit alignments to the *Arabidopsis thaliana* Columbia (Col-0)

transcriptome (The Arabidopsis Information Resource, TAIR10) with the USEARCH software (v.7; Edgar, 2010). Contigs with no reciprocal best hit against *A. thaliana* were then iteratively aligned with the same strategy against the *Arabidopsis lyrata* ssp. *lyrata* (L.) O'Kane & Al-Shehbaz transcriptome (v.1.0; PHYTOZOME) and the UniRef and SwissProt databases, considering only the transcripts annotated by a Viridiplantae sequence (i.e. land plants and green algae). A number of non-Viridiplantae contigs were additionally considered when highly expressed (total count of *c.* 1000 FPKM (fragments per kilobase of exon per million reads mapped)), with counts in at least 80% of the samples. This strategy produced an assembly consisting of 34 818 transcripts, corresponding to 29 076 genes.

TRANSRATE (v.1.0.3; Smith-Unna *et al.*, 2016) was used to assess the assembly quality based on the alignment of filtered reads to annotated contigs and to compare assembly summary statistics. BUSCO (v.1.22; Simão *et al.*, 2015) was used for comparative analysis of the *A. halleri*, *A. thaliana* and *A. lyrata* transcriptomes. ORTHOFINDER (v.1.1.2; Emms & Kelly, 2015) was used to assess matching gene families between the three species with an alignment expected value threshold of 10^{-5} and a Markov cluster algorithm inflation value of 1.5. CATEGORIZER (v.3.218, Zhi-Liang *et al.*, 2008) was used to calculate gene ontology (GO) term frequencies with the Plant_GOslim as GO classification and the consolidated single occurrence method option.

The RNA-Seq reads and transcriptome assembly have been deposited in the National Center for Biotechnology Information (NCBI) Transcriptome Shotgun Assembly Sequence Database (TSA) with BioProject identification number PRJNA388549.

Statistical analysis of RNA-Seq data

Quality filtered reads were mapped to the PL22 reference transcriptome. The TRINITY pipeline was used to estimate transcript abundance using BOWTIE2 (Langmead & Salzberg, 2012) and RSEM (Li & Dewey, 2011). Principal component analysis (PCA) plots were created with the *PlotPCA* function from R using RLOG transformed data (Beginner's guide; DESEQ2 package; 13 May 2014, http://www.bioconductor.org/packages/2.14/bioc/vignettes/DESeq2/inst/doc/beginner.pdf). Differentially expressed genes (DEGs) were identified by pairwise comparisons with the DESEQ2 package (v.1.12.3; Love *et al.*, 2014). Genes were retained as differentially expressed when the $\log_2$ fold-change (FC) was > 0.5 or < -0.5, with a false discovery rate (FDR; Benjamini–Hochberg) adjusted *P*-value of < 0.05. Pairwise comparisons were carried out considering: GU as the main factor (PL22 vs I16; six replicates per population: three controls and three treatments); the treatment conditions as the main factor (treatment vs control; six replicates per condition: three PL22 and three I16); the effect of treatment in each population.

Gene ontology enrichment analysis and MAPMAN visualization

Sets of differentially expressed genes were analyzed for GO term enrichment using the BINGO application (v3.0.3) (Maere *et al.*, 2005) of CYTOSCAPE (v.3.4.0; Shannon *et al.*, 2003) using the following criteria: hypergeometric statistical test; Benjamini and Hochberg FDR correction of < 0.05. The DEGs were analyzed to visualize overrepresented categories after correction using the whole annotation as reference testing for the ontology terms biological process, molecular function and cellular components. The ontology file used was the go-basic.obo (7 July 2016 version) downloaded from http://geneontology.org/ontology/go-basic.obo and the ontology annotation files for each category were customized for the *A. halleri* reference transcriptome by assigning to each transcript the GO term corresponding to its annotation. Pathway enrichment results were obtained with the THALEMINE tool from ARAPORT11 (Krishnakumar *et al.*, 2015).

The MAPMAN (v.3.6.0; Usadel *et al.*, 2009) tool was used to classify genes in Cell Function bins. Data were mapped to the 'Ath_AFFY_ATH1_TAIR10_Aug2012.m02' annotation file.

qRT-PCR analysis

Total DNase-treated RNAs were extracted as described in the RNA extraction section above. One microgram of total RNA was used for cDNA synthesis with Oligo dT and the RevertAid H Minus Kit (Thermo Fisher Scientific, Merelbeke, Belgium). Quantitative polymerase chain reactions (qPCRs) were performed with a QuantStudio 5 machine (Applied Biosystems, Foster City, CA, USA) using PowerUp™ SYBR® green Master Mix (Applied Biosystems) (Nouet *et al.*, 2015). Based on RNA-Seq data, primers were designed in the region of the cDNAs that displayed no sequence variants in I16 and PL22 to ensure proper amplification in both populations (Table S1). A total of three technical replicates were run for each combination of cDNA and primer pair in the three biological replicates as described previously (Nouet *et al.*, 2015). Reaction efficiencies were determined using the LINREGPCR software (v.2013; Ruijter *et al.*, 2009). Mean reaction efficiencies were then determined for each primer pair from all reactions (> 36 reactions; Table S1). Relative gene expression levels were calculated by the $2^{-\Delta\Delta Ct}$ method (Pfaffl, 2001) using *Elongation Factor Alpha 1* (*EF1α*) as reference (Talke *et al.*, 2006).

Protein extraction and Western blot analysis

To extract total proteins, 100 mg of frozen ground root tissues was homogenized with 150 µl of extraction buffer as described previously (Seguela *et al.*, 2008). Western blots were performed using standard protocols. The IRT1 polyclonal antibody was raised in rabbits against the peptide H- CPANDVTLPIKEDDSS-NH2 (Eurogentec, Liège, Belgium) (Seguela *et al.*, 2008) and used at a 1/2500 dilution. After incubation with a goat anti-rabbit antibody coupled to horseradish peroxidase (HRP; Ab6721; AbCam, Cambridge, UK) at a 1/5000 dilution and chemiluminescence detection (BM Kit; Roche, Basel, Switzerland), the blot was imaged with the GENESNAP software of a Syngene G: Box (Cambridge, UK). Root tissues from two biological replicates were analyzed in two technical replicates.

Results

Building an *A. halleri* reference transcriptome

To establish an adequate and specific resource for downstream RNA-Seq analysis (this study and Corso *et al.*, 2017), a *de novo* transcriptome assembly for *A. halleri* (Fig. S1) was carried out using as reference the PL22 population, a polish M population (Corso *et al.*, 2017). Among the two populations examined in this study, the HZ population PL22 was selected as reference, as it bridges the North-West and South-East GUs. The initial *de novo* assembly obtained with TRINITY produced 156 529 genes consisting of 246 633 isoforms, which was reduced to 71 973 genes consisting of 79 480 isoforms after filtering with the EVIDENTIALGENE tool to remove redundancy. Of those, 34 818 contigs were annotated as follows and used in downstream analyses. First, the 79k contigs were matched to the *A. thaliana* transcriptome, selected as the best annotated reference transcriptome available within the *Arabidopsis* genus, and 31 211 contigs were annotated. Second, the remaining 3607 contigs were matched to *A. lyrata*, resulting in an additional 886 annotated contigs. Third, 2721 contigs were annotated through a match to other Viridiplantae (2485 contigs) or retained based on consistent expression across samples (236 contigs) (Table S2). Summary statistics for the annotated transcriptome are presented in Table S3. Overall, 99.28% of transcripts had an annotation based on a match to a Viridiplantae species and the remaining 0.72% were annotated with species belonging to diverse taxonomic groups (Fig. S1; Table S2).

The distribution of GO terms showed that, overall, the *A. halleri* transcriptome did not display any striking differences compared with the functional annotation of the *A. thaliana* and *A. lyrata* transcriptomes (Fig. S2).

Quality statistics for the *A. halleri* transcriptome assembly were computed using the TRANSRATE tool upon mapping of the quality filtered reads to the 34 818 annotated transcripts. The analysis confirmed that 90% of the annotated transcripts were correctly assembled, with a global assembly score of 0.35. In addition, a BUSCO (Benchmarking Universal Single-Copy Orthologs) analysis, which measures expected gene content based on single-copy orthologs selected from an orthology database (OrthoDB; v.9.1, http://www.orthodb.org/) and provides a measurement of transcriptome assembly completeness, also indicated that 90% of transcripts were complete (in agreement with the TRANSRATE metric), 23% duplicated, 1.8% fragmented, and 7.4% missing. These statistics are fully in line with those of the *A. thaliana* and *A. lyrata* transcriptomes (Table S4).

Orthologous gene groups in whole-transcriptome comparisons of the three *Arabidopsis* species were accurately inferred using ORTHOFINDER. The analysis revealed that 17 945 orthogroups (i.e. gene families; Table S5) were shared among *A. thaliana, A. lyrata* and *A. halleri* transcriptomes, whereas as little as 12, 13 and 24 were species-specific orthogroups, respectively (Fig. 1). The 24 *A. halleri*-specific orthogroups consisted of 92 genes (and 119 isoforms) and a GO term overrepresentation analysis showed that these genes were enriched for mitochondrial energy production (Table S6), and more specifically for the succinate dehydrogenase (SDH) complex which has been shown to play an important role in reactive oxygen species (ROS) production in plant mitochondria and to regulate plant development and stress responses (Jardim-Messeder *et al.*, 2015).

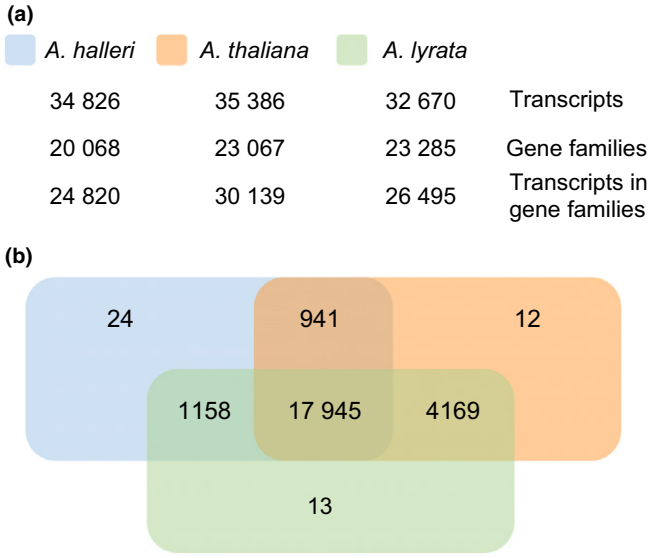

Fig. 1 Orthology relationship in the transcriptomes of *Arabidopsis lyrata*, *A. thaliana* and *A. halleri*. (a) Statistics of transcripts, gene families (or orthogroups) and number of transcripts belonging to gene families in each species. (b) Distribution of shared and unique orthogroups among the three species. A list of the corresponding genes is provided in Supporting Information Table S5.

Distinct Zn accumulation phenotypes in M populations

To examine the physiological and molecular mechanisms underlying intraspecific variation of the Zn accumulation and tolerance traits among distinct M populations, PL22 and I16 plants were grown in hydroponic solution in control (5 μM $ZnSO_4$) or treatment (150 μM $ZnSO_4$) conditions for 2 wk. The treatment had no significant effect either on the shoot chlorophyll content or on the root and shoot biomass of the two populations (Fig. 2), suggesting similar tolerance to the Zn treatment in PL22 and I16. By contrast, even though the two populations shared similar Zn concentrations in root and shoot tissues in the control conditions, PL22 displayed markedly higher Zn accumulation in both roots and shoots (Fig. 3), as observed in field-collected samples (Corso *et al.*, 2017). I16 maintained a shoot : root Zn ratio of *c.* 1 in the two growth conditions and did not reach the shoot hyperaccumulation threshold (0.3% of dry weight as defined by Krämer, 2010) (Fig. 3c). PL22 displayed higher shoot accumulation upon treatment, reaching the hyperaccumulation level. Moreover, when exposed to high Zn, PL22 accumulated more Zn in roots than in shoots, which resulted in a shoot : root Zn ratio of 0.4 (Fig. 3c). These observations highlighted the contrasting physiological behaviors of the two populations. I16 restricted root Zn uptake at high Zn exposure and maintained

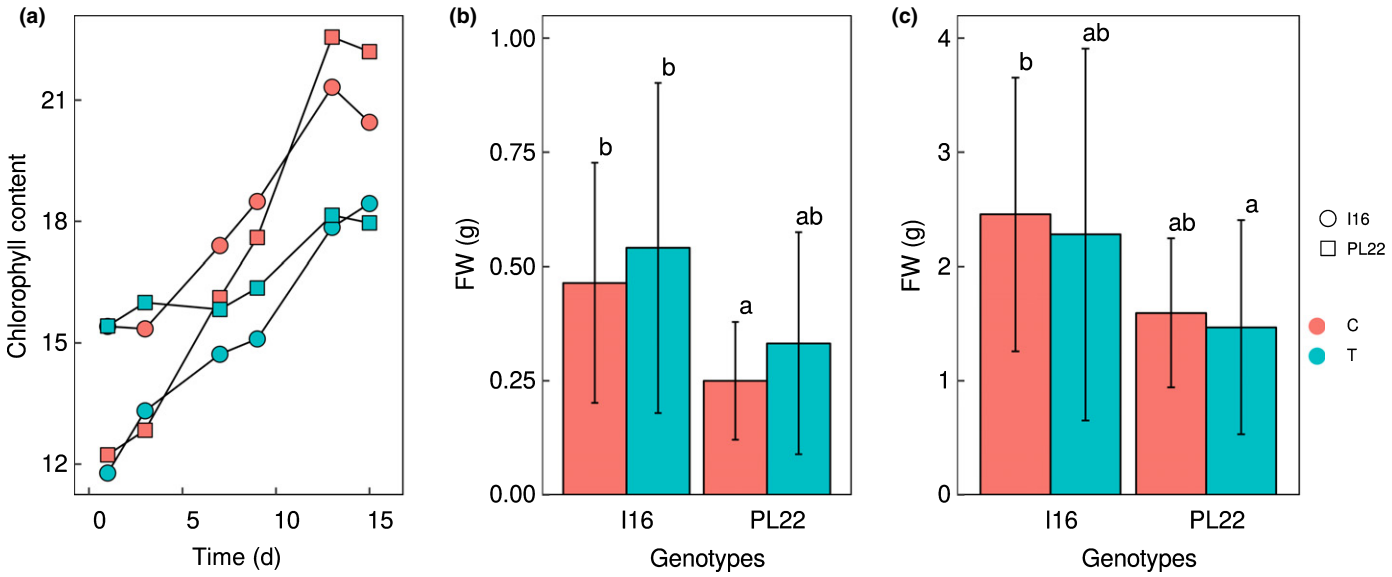

Fig. 2 (a) Chlorophyll content in the PL22 and I16 metallicolous *Arabidopsis halleri* populations during 15 d of exposure to 5 μM $ZnSO_4$ (control, C) and 150 μM $ZnSO_4$ (treatment, T). (b) Root fresh weight (FW) biomass and (c) shoot FW biomass, both recorded at the time of harvest (day 15) in control and treatment conditions. Error bars represent ± SD of three biological replicates. Labels indicate the significance of an ANOVA test calculated with the Tukey post hoc test (*P*-value < 0.05).

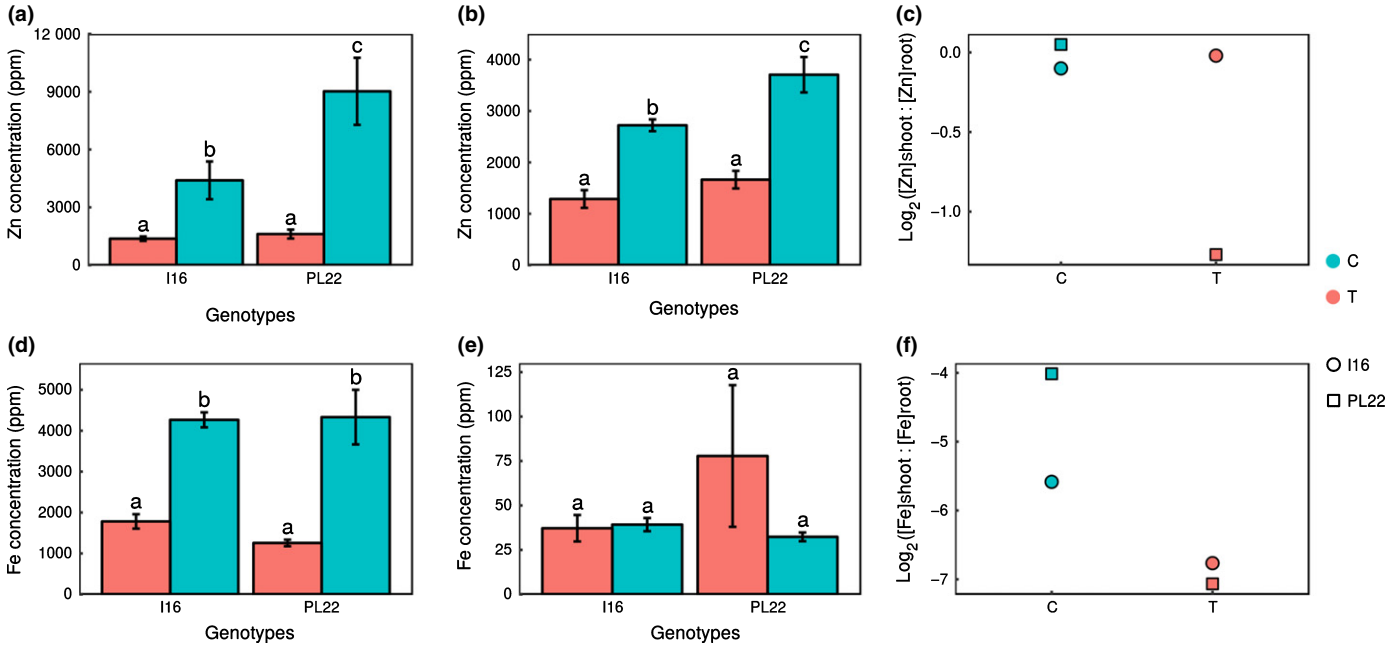

Fig. 3 (a, b) Zinc (Zn) and (d, e) iron (Fe) concentrations in (a, d) roots and (b, e) shoots of the PL22 and I16 metallicolous *Arabidopsis halleri* populations upon 15 d of exposure to 5 μM $ZnSO_4$ (control, C) and 150 μM $ZnSO_4$ (treatment, T). Error bars represent ± SD of three biological replicates. Labels indicate the significance of an ANOVA test calculated with the Tukey post hoc test (*P*-value < 0.05). (c, f) Log_2 of the root-to-shoot ratio of average (b) Zn and (f) Fe concentrations in PL22 and I16 at 5 or 150 μM $ZnSO_4$.

moderate Zn translocation to the shoot. PL22, by contrast, displayed higher Zn uptake, retention in roots, and shoot accumulation.

In addition, the Zn treatment resulted in increased accumulation of Fe in the roots in both populations and the mean Fe concentration was higher in PL22 shoots but not significantly as a consequence of the high variance in control conditions (Fig. 3d, e). Consequently, the Fe shoot : root ratio decreased after Zn treatment to a greater extent in PL22 than in I16 (Fig. 3f). Moreover, calcium (Ca) concentrations were also higher in shoots of both populations, with a significant shoot : root ratio difference between I16 and PL22 upon Zn treatment (Figs S3, S4).

Geographic origin as a major source of transcriptomic variation

The molecular mechanisms underlying the phenotypic differences observed between PL22 and I16 were further examined with a comparative transcriptomic experiment using RNA-Seq on the samples described in the previous section. After read mapping against the *A. halleri* reference transcriptome and estimation of transcript abundance for each sample, pairwise comparisons were used to estimate the contribution of the GU (PL22 vs I16) and the treatment (T; control vs treatment) to differential gene expression ($\log_2$ fold change >0.5 or <-0.5; FDR <0.05; Table S7). Principal component analysis revealed that the GU was the most significant source of variation, accounting for 83% and 88% of the gene expression variance in roots and shoots, respectively, whereas the treatment only explained 12% and 9%, respectively, of this variance (Fig. 4). In agreement with this, the GU factor had the strongest effect on the number of DEGs: the PL22/I16 comparison yielded 1505 DEGs in roots and 700 DEGs in shoots, whereas the control/treatment comparison resulted in 226 DEGs in roots and no DEGs in shoots (Fig. 5; Table S7). There was no common DEG between GU and treatment comparisons. The Zn treatment thus triggered a more profound response in roots than in shoots (Fig. 5). Strikingly, PL22 and I16 shared no common response to Zn in shoots, whereas the root response to Zn involved mostly genes with higher expression in control conditions than in the treatment (Fig. 5; Table S7).

Genes with higher expression in PL22 than in I16 were enriched for defense response, DNA metabolic process and DNA integration GO terms. Overexpressed genes in I16 compared with PL22 were related to amino acid homeostasis (transmembrane transport) and carboxylic acid transmembrane transport.

Population-specific responses to Zn

Comparisons of gene expression in control and treatment conditions within each genotype revealed population-specific responses to Zn. In roots, the Zn treatment affected the expression of 73 and 524 genes in PL22 and I16, respectively, while 72 DEGs were in common between the two populations (Fig. 5b). Most of these genes showed higher expression in control conditions than in treatment conditions, accounting for 49 DEGs in PL22, 507 in I16, and all the 72 common DEGs. This indicates that Zn exposure resulted mostly in gene downregulation in both populations, but to a different extent (Table S7). In shoots, only four genes were differentially expressed upon Zn treatment in PL22 and none in I16, confirming the limited effect of the treatment on the shoot transcriptome.

Impact of Zn on Fe homeostasis in PL22

Detailed analysis of the population-specific responses to Zn suggested that PL22 displayed distinct Fe homeostasis compared with I16 (Fig. 6). The *Iron-Regulated-Transporter 1* (*IRT1*) gene, encoding the main Fe uptake protein in *A. thaliana* (Vert *et al.*, 2002), showed constitutively higher expression in PL22 roots compared with I16 and approximately four-fold higher expression in treatment than in control conditions in PL22 roots (not significant; *P*-value >0.05). qRT-PCR and western blot analysis (Fig. S5a,b) confirmed the induction of *IRT1*, at both transcript and protein levels, in PL22 roots upon Zn exposure and the overall higher expression of *IRT1* in PL22 compared with I16. Moreover, among the 24 genes significantly induced by Zn in PL22 roots, 13 were genes whose expression is induced by Fe deficiency in *A. thaliana* (Table 1; Fig. 6). For instance, these included *Ferric Reduction Oxidase 2* (*FRO2*) (Robinson *et al.*, 1999), which contributes to Fe uptake together with *IRT1*, *IRT2* (Vert *et al.*, 2009), *Feruloyl CoA ortho-hydroxylase 1* (*F6′H1*), involved in coumarin biosynthesis, which contributes to Fe uptake (Schmid *et al.*, 2014), *ZIP8* (Milner *et al.*, 2013), *Copper Transporter 2* (*COPT2*) (Perea-García *et al.*, 2013) and *Zinc-Induced Facilitator 1* (*ZIF1*) (Haydon *et al.*, 2012). Although 11 of those genes are targets of Fe-Deficiency Induced Transcription Factor (FIT; Table 1; Colangelo & Guerinot, 2004), this transcription factor was not differentially regulated in PL22 either in response to Zn or compared with I16 (Fig. 6).

Interestingly, several transcription factors involved in the Fe deficiency response were upregulated in response to Zn in PL22: *basic Helix-Loop-Helix 038* (*bHLH038*), *bHLH039* and *bHLH100* (Wang *et al.*, 2007), and *MYeloBlastosis 10 (MYB10)* and *MYB72* (Palmer *et al.*, 2013). Note that the three bHLH

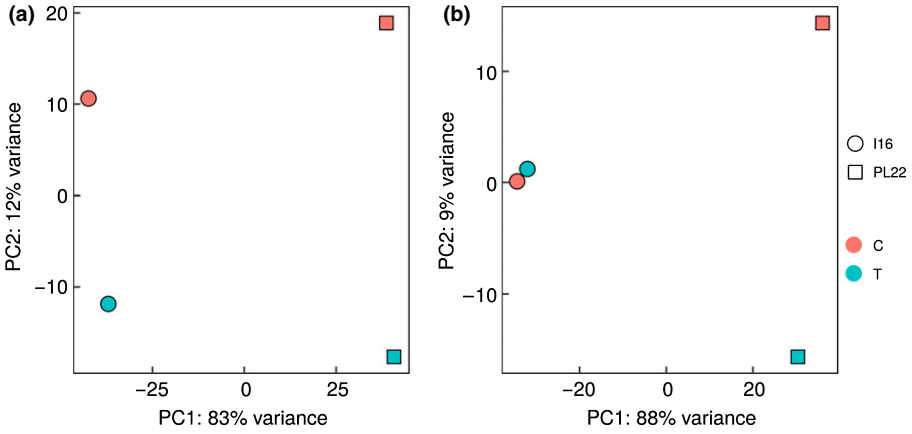

Fig. 4 Principal component analysis of RNA sequencing (RNA-Seq) gene expression data of (a) root and (b) shoot samples of the PL22 and I16 metallicolous *Arabidopsis halleri* populations upon 15 d of exposure to 5 μM $ZnSO_4$ (control, C) and 150 μM $ZnSO_4$ (treatment, T).

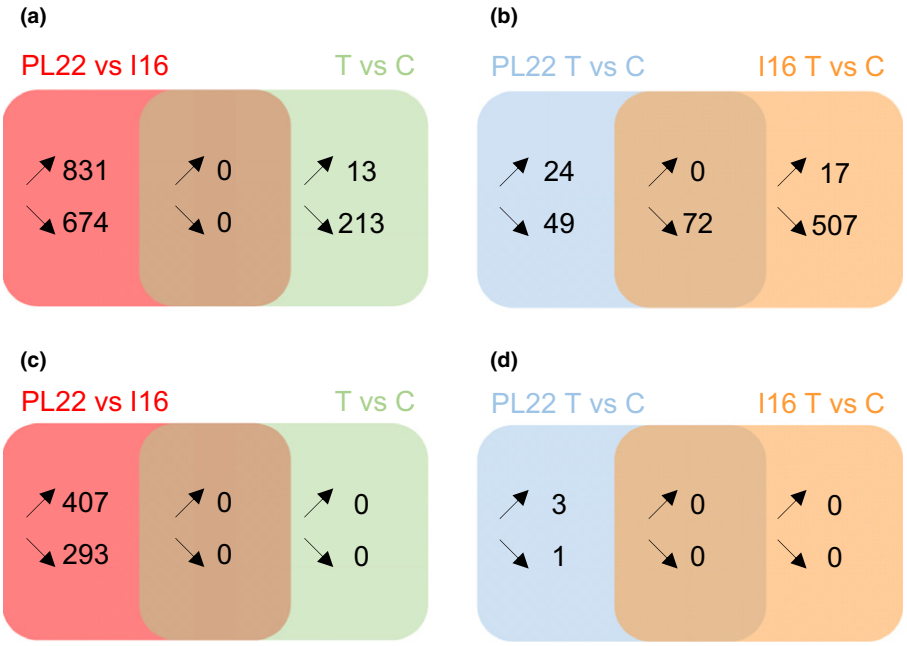

Fig. 5 Distribution of differentially expressed genes in (a, b) root and (c, d) shoot samples of the PL22 and I16 metallicolous *Arabidopsis halleri* populations upon 15 d of exposure to 5 μM $ZnSO_4$ (control, C) and 150 μM $ZnSO_4$ (treatment, T). (a, c) Comparisons using genetic unit (PL22 vs I16) and treatment conditions (T vs C) as main factors. (b, d) Comparisons of the response to the treatment within each population. In each panel, upper and lower numbers represent genes with higher or lower expression in each comparison, respectively. Lists of the corresponding genes and their annotation are provided in Supporting Information Table S7.

genes were constitutively (independent of the Zn treatment) highly expressed in I16 roots.

In addition to a strong root response, *bHLH038*, *bHLH039*, *bHLH100* and *bHLH101* were also strongly upregulated in PL22 shoots in response to Zn. They have been shown to contribute to Fe deficiency- and leaf development-related processes in shoots of *A. thaliana* (Maurer *et al.*, 2014; Brumbarova *et al.*, 2015). Finally, the *Ferritin 2* (*FER2*) gene, encoding an important protein for oxidative cell damage protection (Ravet *et al.*, 2009a,b), was constitutively more highly expressed in shoots of PL22 compared with I16 (Table 1).

Moderate effect of Zn on Fe and Zn homeostasis in I16

In contrast to PL22, the Zn treatment had little effect on the regulation of metal homeostasis-related genes in I16. However, four Fe and Zn homeostasis genes were constitutively more highly expressed in I16 (Table 1; Fig. 6). *HMA2*, encoding a Zn transporter which is required together with HMA4 for Zn root-to-shoot translocation (Hussain *et al.*, 2004; Wong & Cobbett, 2009), was more highly expressed in roots and shoots of I16. Three more genes (*ZIP3*, *NAS4* and *FRD3*) were all more highly expressed in I16 roots and were previously listed as candidate genes for a role in Zn hyperaccumulation/hypertolerance, as they were constitutively highly expressed in *A. halleri* compared with *A. thaliana* (Talke *et al.*, 2006; Krämer *et al.*, 2007). ZIP3 is a putative Zn transporter (Milner *et al.*, 2013). NAS4 contributes to NA synthesis (Curie *et al.*, 2009) and is required under Fe deficiency and for tolerance to Zn excess (Palmer *et al.*, 2013). FRD3 is a citrate transporter, loading citrate into the xylem in roots, which is required for proper Fe homeostasis and Zn tolerance (Rogers & Guerinot, 2002; Green & Rogers, 2004; Durrett *et al.*, 2007; Pineau *et al.*, 2012; Charlier *et al.*, 2015).

Massive gene overexpression in I16 in control conditions

Remarkably, 579 genes were more highly expressed in I16 roots in control conditions compared with the Zn treatment (Fig. 5b). This expression pattern might result from lower gene expression upon Zn treatment, or higher expression in control conditions. Out of the 579 genes, 72 genes were highly expressed also in PL22 (Fig. 5b), but the vast majority, 507 genes, generally showed low expression in both control and treatment conditions in PL22, confirming a surprisingly massive gene overexpression in control conditions in I16 (Fig. 7). These 507 genes were functionally enriched for ribosome, citrate and carbon metabolism pathways. They also included a group of 35 genes previously identified as responsive to Cd stress in *A. thaliana* (Table S8) (Sarry *et al.*, 2006).

Similarly, 49 genes were specifically overexpressed in control conditions in PL22 and were functionally enriched for ribosome and protein synthesis. However, the number of highly expressed ribosome pathway genes was significantly higher in I16 than in PL22, suggesting a massive overexpression in control conditions in I16 which was not observed in PL22.

Discussion

Arabidopsis halleri reference transcriptome for RNA-Seq studies

Although *A. thaliana* is still widely used as a reference for comparative genomics and transcriptomics among Brassicaceae (Koornneef & Meinke, 2010; Mondragon *et al.*, 2017) and it shares fairly high identity within coding sequences (*c.* 95%) with *A. halleri* (Talke *et al.*, 2006), a *de novo* reference transcriptome of *A. halleri* was assembled in this study to allow accurate

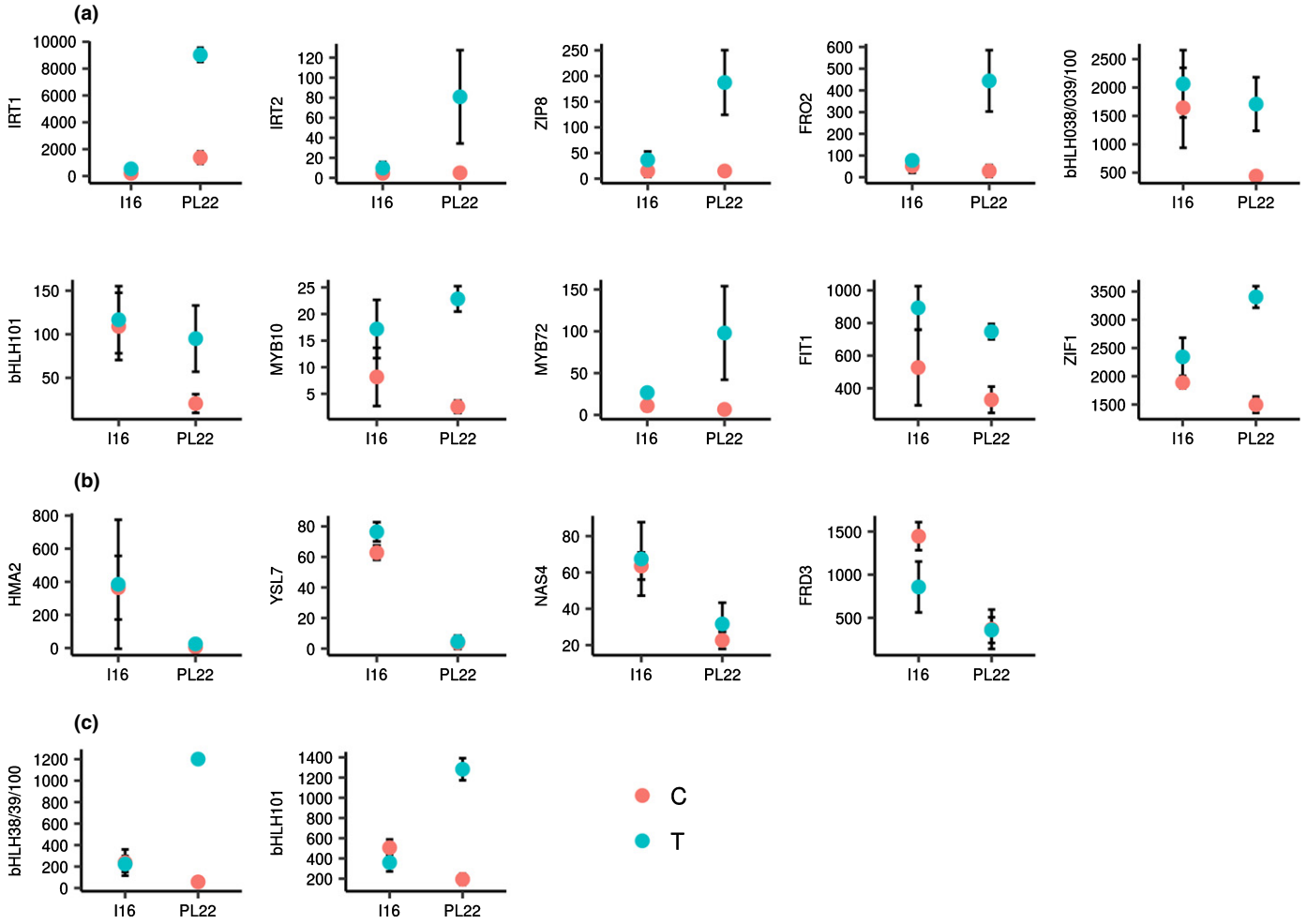

Fig. 6 RNA sequencing (RNA-Seq) normalized expression levels of genes (a) involved in iron (Fe) deficiency response and induced by zinc (Zn) in *Arabidopsis halleri* PL22 roots, (b) with constitutively higher expression in I16 roots compared with PL22 and (c) differentially expressed in shoots of PL22 upon Zn treatment. Values are the average of normalized read counts of three biological replicates and error bars represent ± SD. Red and blue dots indicate control (C; 5 μM $ZnSO_4$) and treated (T; 150 μM $ZnSO_4$) samples, respectively. Read counts are normalized by size factors (the median of the ratios of observed counts).

estimation of *A. halleri* gene expression levels. A combination of *de novo* RNA-Seq read assembly, redundancy reduction and stringent annotation using Viridiplantae as references resulted in a transcriptome including 34 818 transcripts, representing 29 076 genes expressed in root and/or shoot tissues in a set of metal treatment conditions (Corso *et al.*, 2017; this work). A detailed discussion of the quality of this assembly is presented in Notes S1 and Fig. S6. Despite some limitations, we present the first release of an *A. halleri* transcriptome assembly, which was built from a Polish population belonging to the HZ GU of the species in Europe. It constitutes a useful reference, suitable for estimating gene expression in populations of the two main GUs of the species, as illustrated with the analysis of an Italian population. Moreover, these limitations appeared to have a minor impact on gene expression estimation. Indeed, the RNA-Seq data were validated by quantitative RT-PCR for a set of 17 genes that emerged as relevant in this study. RNA-Seq and qRT-PCR transcript level estimations were highly correlated, with an R value of 0.9017 ($R^2 = 0.813$) and a P-value < 2.2e-16 between the $\log_2$ FC in RNA-Seq counts and the $\log_2$ FC in qRT-PCR transcript levels (Fig. S7), in agreement with the validation data reported in Corso *et al.* (2017).

Distinct metal accumulation and gene expression features among *A. halleri* metallicolous populations

Here and in an accompanying study (Corso *et al.*, 2017), we used this new resource to examine the mechanisms underlying intraspecific variation in two geographically distant M populations of *A. halleri*. Phylogeography and population genetics suggested that colonization of anthropogenic metal-contaminated sites was recent and occurred independently from nearby nonmetalliferous sites within each GU of the distribution of *A. halleri* in Europe (Meyer *et al.*, 2009, 2010; Pauwels *et al.*, 2012). Ionomic and transcriptomic analyses revealed that the geographic origin of the populations represented a major source of variation both for Zn (Fig. 3) and Cd (Corso *et al.*, 2017) accumulation properties and for gene expression profiles (Fig. 4). Indeed, in a PCA of RNA-Seq data, the highest factor of variation in gene expression among samples was the GU, with an outstanding 83–88% of explained

Table 1 Zinc (Zn) and iron (Fe) homeostasis-related genes differentially regulated in *Arabidopsis halleri* PL22 and I16, either constitutively or upon Zn treatment

Comparison	Gene ID[1]	Symbol[3]	Log_2 FC	Adj *P*-value
	Impact of Zn on Fe homeostasis genes in PL22			
	Root			
PL22 vs I16	AT4G19690	*IRT1*[2]	2.38	0.006
PL22 T vs C	AT4G19680	*IRT2*[2]	2.61	0.016
PL22 T vs C	AT5G45105	*ZIP8*	2.84	0.0002
PL22 T vs C	AT1G01580	*FRO2*[2]	2.91	0.0002
PL22 T vs C	AT3G56970	*bHLH38*[2]	1.81	0.001
PL22 T vs C	AT3G56980	*bHLH39*[2]	1.81	0.001
PL22 T vs C	AT2G41240	*bHLH100*[2]	1.81	0.001
PL22 T vs C	AT5G13740	*ZIF1*	1.17	5.50E-05
PL22 T vs C	AT3G46900	*COPT2*[2]	2.4	0.004
PL22 T vs C	AT3G12820	*MYB10*[2]	2.29	0.04
PL22 T vs C	AT1G56160	*MYB72*[2]	2.98	7.78E-05
PL22 T vs C	AT3G50740	*UGT72E1*[2]	2.06	1.85E-05
PL22 T vs C	AT1G34760	*GRF11*[2]	2.25	0.0098
PL22 T vs C	AT3G13610	*F6'H1*[2]	3.16	2.38E-07
	AT1G55290	*F6'H2*		
	Shoot			
PL22 T vs C	AT3G56970	*bHLH38*[2]	3.02	4.90E-08
PL22 T vs C	AT3G56980	*bHLH39*[2]	3.02	4.90E-08
PL22 T vs C	AT2G41240	*bHLH100*[2]	3.02	4.90E-08
PL22 T vs C	AT5G04150	*bHLH101*[2]	2.38	4.90E-08
PL22 vs I16	AT3G11050	*FER2*	1.92	0.003
	Moderate effect of Zn on Fe and Zn homeostasis genes in I16			
	Root			
PL22 vs I16	AT4G3110	*HMA2*	−3.8	0.02
PL22 vs I16	AT1G56430	*YSL7*	−3.69	7.46E-20
PL22 vs I16	AT3G08040	*FRD3*	−1.51	0.02
PL22 vs I16	AT4G3110	*ZIP3*	−3.17	1.42E-07
	Shoot			
PL22 vs I16	AT2G32270	*HMA2*	−1.94	2.55E-05

Log_2 fold changes (Log_2 FC) and false discovery rate adjusted *P*-values (FDR adj *P*-value < 0.05) from the Benjamini and Hochberg multiple comparison correction are shown. The log_2 FC represents the log_2 of the fold change of the first element of the comparison over the second.

[1]Corresponding Arabidopsis Gene Identifier (AGI) number in *Arabidopsis thaliana*.

[2]Genes under Fe-Deficiency Induced Transcription Factor (FIT) regulation (see the Results section).

[3]*bHLH, basic Helix-Loop-Helix; COPT2, Copper transporter 2; F6'H, Feruloyl CoA ortho-hydroxylase; FER2, Ferritin 2; FRD3, Ferric Chelate Reductase Defective 3; FRO2, Ferric Reduction Oxidase 2; GRF11, General Regulatory Factor 11; HMA2, Heavy Metal ATPase 2; IRT, Iron-Regulated-Transporter; MYB, MYeloBlastosis; UGT72E1, UDP-Glucosyl Transferase 72E1; YSL7, Yellow Stripe-Like 7; ZIF1, Zinc-Induced Facilitator 1; ZIP, Zinc-regulated transporter Iron-regulated transporter-like Protein.*

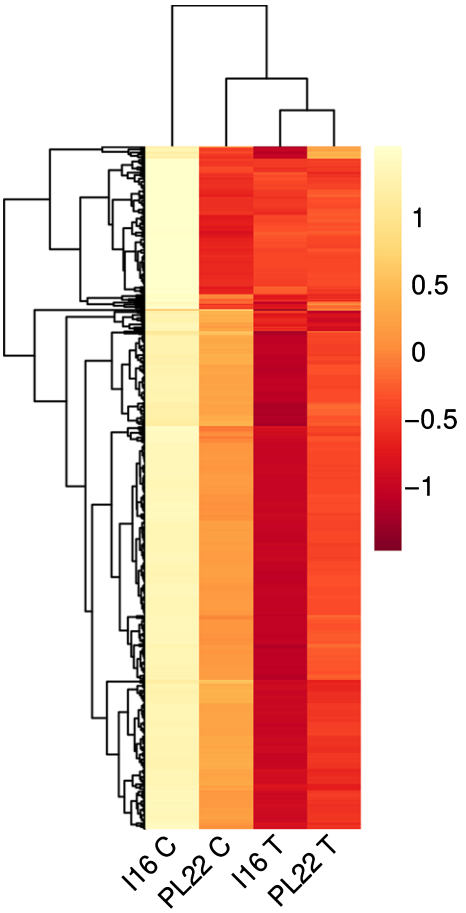

Fig. 7 Heat map depicting the expression profile of 579 genes overexpressed in roots of the I16 *Arabidopsis halleri* population in control conditions (C; 5 μM $ZnSO_4$) in both PL22 and I16 populations exposed to control and treatment (T; 150 μM $ZnSO_4$).

variance. Although a Cd treatment had a stronger effect on gene expression, GU remained by far the largest source of variation in the accompanying study (Corso *et al.*, 2017). Our data suggest that the Polish and Italian *A. halleri* populations not only constitute distinct GUs (i.e. units built based on neutral chloroplastic and nuclear genetic markers) (Pauwels *et al.*, 2012) but can also be considered as distinct transcriptomic units (i.e. with distinct physiology and possibly adaptation; see below in this section). A major contribution of geographic origin to intraspecific variation was previously reported for the pseudometallophyte *Commelina communis* (Li *et al.*, 2016).

By contrast, the Zn treatment had a much lower impact on gene expression, with seven (roots) to 200 (shoots) times fewer DEGs identified for the treatment than for the GU factor. The Zn treatment effect on gene expression in each population (Fig. 5b) was significantly higher than the global treatment effect (Fig. 5a), indicating that responses to high Zn are not shared between the two populations/GUs.

Although the Zn treatment resulted in contrasting Zn shoot accumulation levels in I16 and PL22, it strikingly had little (four DEGs in PL22) or no effect (zero DEGs in I16) on gene expression in shoots. This suggests that the two populations can accommodate a 0.27–0.37% Zn shoot accumulation level without toxicity or major Zn homeostasis reprogramming (i.e. they possess constitutive Zn tolerance), and the shoot accumulation level is mostly determined by root processes in the two populations. The first point is in agreement with previous findings in a German *A. halleri* M population (North-West GU) which suggested that high constitutive expression of metal homeostasis genes in shoots, compared with *A. thaliana*, provides tolerance to higher basal levels of Zn accumulation, but possibly also to sudden increases in Zn influx into shoot cells (Becher *et al.*, 2004; Talke

et al., 2006). The second point is in agreement with the modeling of Zn radial transport in roots (Claus *et al.*, 2013), with the key contribution of high *HMA4* expression in roots to hyperaccumulation in *A. halleri* (Hanikenne *et al.*, 2008), and with the conclusions of grafting experiments between *N. caerulescens* and the non-Zn-accumulating *Microthlaspi perfoliatum* (Guimarães *et al.*, 2009).

Compared with Cd (Corso *et al.*, 2017), the Zn treatment had also a more limited impact on (micro)nutrient homeostasis (Figs S3, S4) and metabolite profiles (data not shown). However, it affected differently Zn and Fe homeostasis in PL22 and I16 (see next section).

Furthermore, it is striking that very few (e.g. *FRD3*; see next section) of the *c.* 30 primary candidate genes for a role in Zn/Cd hyperaccumulation and hypertolerance in *A. halleri* identified previously from cross-species comparisons (Becher *et al.*, 2004; Weber *et al.*, 2004; Talke *et al.*, 2006; Krämer *et al.*, 2007; Krämer, 2010; Hanikenne & Nouet, 2011) were differentially expressed between the two populations. Indeed, none of the *HMA4* (Hanikenne *et al.*, 2008), *MTP1* (Dräger *et al.*, 2004; Shahzad *et al.*, 2010), *NAS2* (Deinlein *et al.*, 2012) or *ZIP* candidates (e.g. *IRT3*, *ZIP4*, *ZIP9* and *ZIP6*) (Talke *et al.*, 2006; Krämer *et al.*, 2007) were differentially expressed between the two populations or upon high Zn treatment. This suggests that these genes mostly contribute to constitutive high levels of Zn tolerance and hyperaccumulation in *A. halleri* and that recent adaptation to metal-contaminated sites mostly evolved through alteration of additional processes. Alternatively, evolution in M populations might have been fully convergent. In this context, a recent report based on genetic analysis and quantitative trait locus (QTL) mapping suggested that the high expression of multiple *MTP1* copies (providing increased Zn vacuolar storage in shoots; Dräger *et al.*, 2004; Krämer, 2005; Shahzad *et al.*, 2010) may have evolved in M populations only and be absent in NM populations (Meyer *et al.*, 2016). *MTP1* was highly expressed in both I16 and PL22, as previously reported in two North-West M populations (Dräger *et al.*, 2004; Talke *et al.*, 2006), which suggests either convergent evolution in the different GUs with parallel acquisition of high expression of multiple *MTP1* copies or that the lack of *MTP1* may be a specific feature of the Slovakian NM population described in Meyer *et al.* (2016).

Finally, the intraspecific variation observed between the two M *A. halleri* populations was evident from the high number of DEGs between GUs but also from other key features such as Zn accumulation, Zn and Fe homeostasis and response to control conditions.

Interplay of Zn accumulation and Fe homeostasis

In *A. thaliana*, it is well documented that Zn excess results in a secondary Fe deficiency response (Wang *et al.*, 2007; Shanmugam *et al.*, 2011, 2013). Comparative studies of *A. halleri* ssp. *gemmifera* (Shanmugam *et al.*, 2011) or *Noccaea caerulescens* (van de Mortel *et al.*, 2006) suggested that Zn hyperaccumulators evolved adapted Fe homeostasis in response to high Zn and are not facing Fe deficiency. However, in PL22, 50% of the genes differentially expressed upon Zn treatment in roots are involved in the Fe deficiency response, including a number of genes encoding transcription factors (subgroup Ib BHLHs, MYB10 and MYB72; Brumbarova *et al.*, 2015) and some of their targets involved in Fe uptake and handling (Table 1). Citrate concentrations also increased by approximately three-fold in PL22 roots exposed to Zn compared with control conditions (Fig. S5c), which was previously described in response to both Zn excess and Fe deficiency in several species (Godbold *et al.*, 1984; Zhao *et al.*, 2000; Abadía *et al.*, 2002). Citrate contributes to Fe (and Zn) mobility in the xylem (Abadía *et al.*, 2002; Cornu *et al.*, 2015). The Fe deficiency response in PL22 may be triggered, as in *A. thaliana*, by a high Zn-induced Fe deficiency sensing and response in roots and shoots (Wang *et al.*, 2007) and is probably responsible for increased Fe in roots and possibly enables limited Fe decrease in shoots upon Zn treatment in PL22 (Fig. 3). Interestingly, the Fe deficiency response observed in PL22 was somehow limited to Fe uptake genes. By contrast, genes induced in

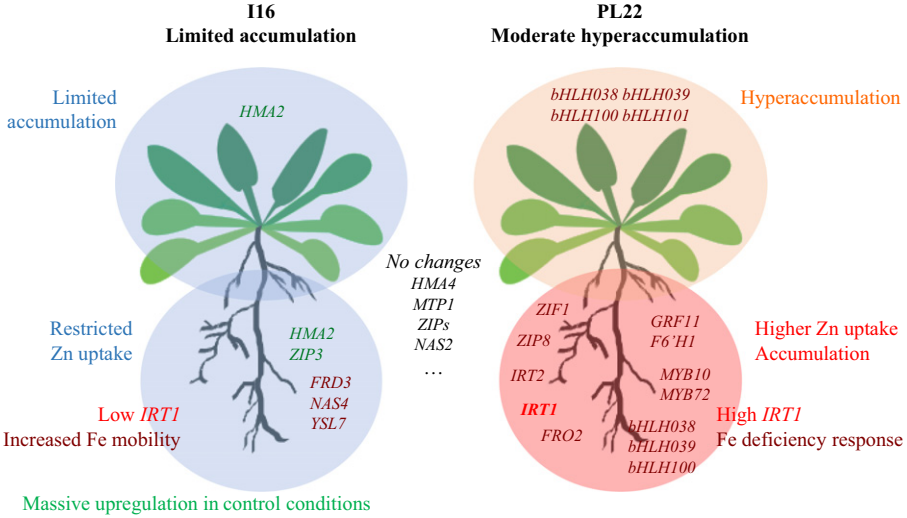

Fig. 8 Summary model of the distinct constitutive and zinc (Zn)-induced Zn accumulation properties and gene expression patterns in the PL22 and I16 *Arabidopsis halleri* metallicolous populations (see details in the Discussion section). Depicted genes for each population are either constitutively more highly expressed (bold) or induced by Zn (italic) in the corresponding population. Fe homeostasis genes are in red and Zn homeostasis genes in green.

A. thaliana by FIT as part of the Fe deficiency response to deal with the IRT1-mediated uptake of nonspecific divalent metal cations (Zn, Mn, Ni…) were not upregulated by Zn in PL22, including for instance the *MTP3*, *MTP8* and *FPN2* genes (Arrivault *et al.*, 2006; Schaaf *et al.*, 2006; Eroglu *et al.*, 2016). This observation may reflect the specialized functions of subgroup Ib BHLHs whose regulation possibly depends on a FIT-independent Zn excess/Fe deficiency signaling pathway (Wang *et al.*, 2007; Maurer *et al.*, 2014; Brumbarova *et al.*, 2015).

Although it also displayed increased Fe root concentrations upon Zn treatment, this massive transcriptional Fe deficiency response was not observed in I16, which was therefore more similar to *A. halleri* ssp. *gemmifera* (Shanmugam *et al.*, 2011) than to PL22 in that respect. This also implies that I16 bypasses the canonical Fe acquisition pathway and may use an alternative pathway to maintain adequate Fe supply. Alternatively, the very small (and not significant) induction of *IRT1*/*FRO2* by Zn observed in I16 (Figs 6, S5) might be sufficient to ensure adequate Fe uptake (Robinson *et al.*, 1999; Vert *et al.*, 2002). In addition, among the very few metal homeostasis genes that are more highly expressed in I16 compared with PL22 (Table 1), the constitutive higher expression of several genes involved in Fe mobility within the plant – *FRD3*, *NAS4* and *Yellow Stripe-Like 7* (*YSL7*) – could contribute to higher availability and mobilization of Fe within I16.

Altogether, our observations suggest that PL22 is more sensitive to changes in external Zn concentration and has a response to high Zn reminiscent of that inherited from the ancestor of the *Arabidopsis* lineage, whereas I16 and other *A. halleri* populations seem to have lost this feature and display more advanced adaptation of their Fe homeostasis network accommodating Zn hyperaccumulation. In this context, we further observed in I16 a disconnection between the expression of Fe homeostasis transcription factors and their described target genes in *A. thaliana* (Fig. 6a,b). For instance, low *MYB72* expression was combined with high expression of its target *NAS4* (Palmer *et al.*, 2013), and high constitutive expression of subgroup Ib *BHLH* genes (*bHLH38*, *39* and *100*) was combined with low expression of *IRT1*, *IRT2* and *FRO2* (Wang *et al.*, 2007).

A putative contribution of IRT1 to intraspecific variation in Zn accumulation

The I16 population colonizes soils with extremely high Zn contamination levels, twice as high as in the PL22 habitat (Corso *et al.*, 2017). In our experimental conditions, it had reduced root and shoot Zn accumulation when exposed to high Zn compared with PL22, and maintained a constant Zn shoot : root ratio independent of the Zn concentration in the growth medium (Fig. 3). The latter was in contrast with both PL22 and the German M population used in many studies (LAN, Talke *et al.*, 2006; Hanikenne *et al.*, 2008). I16 thus displayed an excluder phenotype, essentially achieved through controlled Zn uptake by roots (Fig. 8). By contrast, PL22 behaved as a Zn hyperaccumulator and showed twice as much Zn in the roots upon Zn treatment compared with I16, suggesting less restricted root uptake but controlled root-to-shoot translocation (Fig. 3). The distinct accumulation properties of the two populations do not appear to result from major changes in expression of *ZIP* genes or other genes known to be involved in Zn transport and to be differentially regulated between *A. halleri* and *A. thaliana* (Talke *et al.*, 2006; Krämer *et al.*, 2007). Indeed, these genes were not differentially regulated between I16 and PL22 or were not downregulated by Zn in I16. For instance, among the large number of genes differentially regulated upon Zn treatment in I16 roots, none have a putative function in Zn uptake (Table S7).

Among the DEGs in I16 and PL22, *IRT1* encodes a protein transporting Fe but also Zn (Vert *et al.*, 2002) and its higher expression in PL22, in particular upon Zn treatment, may represent a major contribution to higher Zn uptake in PL22 roots. By contrast, its low expression, stemming from a disconnection with its bHLH regulators (see above), in I16 may reduce Zn uptake. These features possibly represent a hallmark of an advanced adaptation of I16 to the heavily contaminated soils it colonizes, with a more specialized regulation of Fe- and Zn-regulated genes. Furthermore, this advanced adaptation may come at a cost, which is suggested by the observation of a massive overexpression of genes involved in protein and central metabolism in control conditions. This may constitute a stress response reprogramming metabolism (Houot *et al.*, 2007). Further testing of this hypothesis would require transplant experiments *in situ*.

A summary model

In this study, we examined whether geographically distant M *A. halleri* populations have developed convergent or divergent adaptations to local polluted soil conditions. Our data suggest that PL22 and I16 have developed different strategies to adapt to increasing Zn content in the soil through Fe–Zn crosstalk modifications (Fig. 8). I16 seems to control Zn uptake by downregulating genes that play a pivotal role in Fe (and Zn) uptake and mobilization; and I16 exhibits a massive overexpression of genes when exposed to sufficient Zn as opposed to high Zn, which may represent a cost of adaptation imposed on fitness in nonpolluted conditions. Taken together, the findings of this study highlight the contrasting mechanisms of adaptation encountered among *A. halleri* M populations, possibly related to their physiological backgrounds, and reflects the high phenotypic plasticity of the species.

Acknowledgements

We thank P. Salis, C. Baliardini, C-L. Meyer and A. Degueldre for support, the GIGA Genomics platform (University of Liège) and the *durandal* cluster (InBioS-Phytosystems, University of Liège) for computational resources. Funding was provided by the Fonds de la Recherche Scientifique–FNRS (FRFC-2.4583.08 and PDR-T.0206.13 to M.H. and N.V.), the University of Liège (SFRD-12/03 to M.H.) and the Belgian Program on Interuniversity Attraction Poles (IAP no. P7/44). M.H. is Research Associate of the FNRS. M.S. thanks the FRIA for his PhD fellowship.

Author contributions

M.S.S., M. Corso, N.V. and M.H. designed the research. M.H. and N.V. directed the research. M.S.S., M. Corso, N.F., C.N., M.S., B.B. and M. Carnol performed the experiments. M.S.S., M.H., P.M., N.F., N.V. and M. Corso analyzed and interpreted the data. M.S.S. and M.H. generated all figures and Supporting Information material. M.S.S. and M.H. wrote the paper. All authors commented on the manuscript.

References

Abadía J, López-Millán A-F, Rombolà A, Abadía A. 2002. Organic acids and Fe deficiency: a review. *Plant and Soil* **241**: 75–86.

Arrivault S, Senger T, Krämer U. 2006. The Arabidopsis metal tolerance protein AtMTP3 maintains metal homeostasis by mediating Zn exclusion from the shoot under Fe deficiency and Zn oversupply. *Plant Journal* **46**: 861–879.

Baliardini C, Meyer C-L, Salis P, Saumitou-Laprade P, Verbruggen N. 2015. CAX1 co-segregates with Cd tolerance in the metal hyperaccumulator *Arabidopsis halleri* and plays a role in limiting oxidative stress in Arabidopsis. *Plant Physiology* **169**: 549–559.

Becher M, Talke IN, Krall L, Krämer U. 2004. Cross-species microarray transcript profiling reveals high constitutive expression of metal homeostasis genes in shoots of the zinc hyperaccumulator *Arabidopsis halleri*. *Plant Journal* **37**: 251–268.

Bert V, Bonnin I, Saumitou-Laprade P, De Laguérie P, Petit D. 2002. Do *Arabidopsis halleri* from non metallicolous populations accumulate zinc and cadmium more effectively than those from metallicolous populations? *New Phytologist* **155**: 47–57.

Bert V, MacNair MR, De Laguérie P, Saumitou-Laprade P, Petit D. 2000. Zinc tolerance and accumualtion in metallicolous and non metallicolous populations of *Arabidopsis halleri* (Brassicaceae). *New Phytologist* **146**: 225–233.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Brumbarova T, Bauer P, Ivanov R. 2015. Molecular mechanisms governing Arabidopsis iron uptake. *Trends in Plant Science* **20**: 124–133.

Charlier JB, Polese C, Nouet C, Carnol M, Bosman B, Krämer U, Motte P, Hanikenne M. 2015. Zinc triggers a complex transcriptional and post-transcriptional regulation of the metal homeostasis gene *FRD3* in *Arabidopsis* relatives. *Journal of Experimental Botany* **66**: 3865–3878.

Claus J, Bohmann A, Chavarría-Krauser A. 2013. Zinc uptake and radial transport in roots of *Arabidopsis thaliana*: a modelling approach to understand accumulation. *Annals of Botany* **112**: 369–380.

Colangelo EP, Guerinot ML. 2004. The essential basic helix-loop-helix protein FIT1 is required for the iron deficiency response. *Plant Cell* **16**: 3400–3412.

Cornu J, Deinlein U, Horeth S, Braun M, Schmidt H, Weber M, Persson DP, Husted S, Schjoerring JK, Clemens S. 2015. Contrasting effects of nicotianamine synthase knockdown on zinc and nickel tolerance and accumulation in the zinc/cadmium hyperaccumulator *Arabidopsis halleri*. *New Phytologist* **206**: 738–750.

Corso M, Schvartzman MS, Guzzo F, Souard F, Malkowski E, Hanikenne M, Verbruggen N. 2017. Contrasting cadmium resistance strategies in two metallicolous populations of *Arabidopsis halleri*. *New Phytologist*. doi: 10.1111/nph.14948.

Courbot M, Willems G, Motte P, Arvidsson S, Roosens N, Saumitou-Laprade P, Verbruggen N. 2007. A major QTL for Cd tolerance in *Arabidopsis halleri* Co-localizes with *HMA4*, a gene encoding a heavy metal ATPase. *Plant Physiology* **144**: 1052–1065.

Curie C, Cassin G, Couch D, Divol F, Higuchi K, Le Jean M, Misson J, Schikora A, Czernic P, Mari S. 2009. Metal movement within the plant: contribution of nicotianamine and yellow stripe 1-like transporters. *Annals of Botany* **103**: 1–11.

Dahmani-Muller H, van Oort F, Gelie B, Balabane M. 2000. Strategies of heavy metal uptake by three plant species growing near a metal smelter. *Environmental Pollution* **109**: 231–238.

Deinlein U, Weber M, Schmidt H, Rensch S, Trampczynska A, Hansen TH, Husted S, Schjoerring JK, Talke IN, Krämer U *et al.* 2012. Elevated nicotianamine levels in *Arabidopsis halleri* roots play a key role in zinc hyperaccumulation. *Plant Cell* **24**: 708–723.

Dräger DB, Desbrosses-Fonrouge AG, Krach C, Chardonnens AN, Meyer RC, Saumitou-Laprade P, Krämer U. 2004. Two genes encoding *Arabidopsis halleri* MTP1 metal transport proteins co-segregate with zinc tolerance and account for high MTP1 transcript levels. *Plant Journal* **39**: 425–439.

Durrett TP, Gassmann W, Rogers EE. 2007. The FRD3-mediated efflux of citrate into the root vasculature is necessary for efficient iron translocation. *Plant Physiology* **144**: 197–205.

Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–2461.

Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology* **16**: 157.

Eroglu S, Meier B, von Wirén N, Peiter E. 2016. The vacuolar manganese transporter MTP8 determines tolerance to iron deficiency-induced chlorosis in Arabidopsis. *Plant Physiology* **170**: 1030–1045.

Frérot H, Faucon MP, Willems G, Gode C, Courseaux A, Darracq A, Verbruggen N, Saumitou-Laprade P. 2010. Genetic architecture of zinc hyperaccumulation in *Arabidopsis halleri*: the essential role of QTL × environment interactions. *New Phytologist* **187**: 355–367.

Godbold DL, Horst WJ, Collins JC, Thurman DA, Marschner H. 1984. Accumulation of zinc and organic acids in roots of zinc tolerant and non-tolerant ecotypes of *Deschampsia caespitosa*. *Journal of Plant Physiology* **116**: 59–69.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q *et al.* 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* **29**: 644–652.

Green LS, Rogers EE. 2004. FRD3 controls iron localization in Arabidopsis. *Plant Physiology* **136**: 2523–2531.

Guimarães MD, Gustin JL, Salt DE. 2009. Reciprocal grafting separates the roles of the root and shoot in zinc hyperaccumulation in *Thlaspi caerulescens*. *New Phytologist* **184**: 323–329.

Hanikenne M, Baurain D. 2014. Origin and evolution of metal p-Type ATPases in Plantae (Archaeplastida). *Frontiers in Plant Science* **4**: 544.

Hanikenne M, Kroymann J, Trampczynska A, Bernal M, Motte P, Clemens S, Krämer U. 2013. Hard selective sweep and ectopic gene conversion in a gene cluster affording environmental adaptation. *PLoS Genetics* **9**: e1003707.

Hanikenne M, Nouet C. 2011. Metal hyperaccumulation and hypertolerance: a model for plant evolutionary genomics. *Current Opinion in Plant Biology* **14**: 252–259.

Hanikenne M, Talke IN, Haydon MJ, Lanz C, Nolte A, Motte P, Kroymann J, Weigel D, Krämer U. 2008. Evolution of metal hyperaccumulation required *cis*-regulatory changes and triplication of *HMA4*. *Nature* **453**: 391–395.

Haydon MJ, Kawachi M, Wirtz M, Hillmer S, Hell R, Krämer U. 2012. Vacuolar nicotianamine has critical and distinct roles under iron deficiency and for zinc sequestration in arabidopsis. *Plant Cell* **24**: 724–737.

Houot L, Floutier M, Marteyn B, Michaut M, Picciocchi A, Legrain P, Aude J-C, Cassier-Chauvat C, Chauvat F. 2007. Cadmium triggers an integrated reprogramming of the metabolism of *Synechocystis* PCC6803, under the control of the Slr1738 regulator. *BMC Genomics* **8**: 350.

Hussain D, Haydon MJ, Wang Y, Wong E, Sherson SM, Young J, Camakaris J, Harper JF, Cobbett CS. 2004. P-type ATPase heavy metal transporters with roles in essential zinc homeostasis in Arabidopsis. *Plant Cell* **16**: 1327–1339.

Jardim-Messeder D, Caverzan A, Rauber R, de Souza Ferreira E, Margis-Pinheiro M, Galina A. 2015. Succinate dehydrogenase (mitochondrial complex II) is a source of reactive oxygen species in plants and regulates development and stress responses. *New Phytologist* **208**: 776–789.

Koornneef M, Meinke D. 2010. The development of Arabidopsis as a model plant. *Plant Journal* **61**: 909–921.

Krämer U. 2005. MTP1 mops up excess zinc in Arabidopsis cells. *Trends in Plant Science* **10**: 313–315.

Krämer U. 2010. Metal hyperaccumulation in plants. *Annual Review in Plant Biology* **61**: 517–534.

Krämer U, Talke IN, Hanikenne M. 2007. Transition metal transport. *FEBS Letters* **581**: 2263–2272.

Krishnakumar V, Hanlon MR, Contrino S, Ferlanti ES, Karamycheva S, Kim M, Rosen BD, Cheng C-Y, Moreira W, Mock SA *et al.* 2015. Araport: the Arabidopsis Information Portal. *Nucleic Acids Research* **43**: D1003–D1009.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**: 357–359.

Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**: 323.

Li J, Xu H, Song Y, Tang L, Gong Y, Yu R, Shen L, Wu X, Liu Y, Zeng W. 2016. Geography plays a more important role than soil composition on structuring genetic variation of pseudometallophyte *Commelina communis*. *Frontiers in Plant Science* **7**: 1085.

Lin YF, Liang HM, Yang SY, Boch A, Clemens S, Chen CC, Wu JF, Huang JL, Yeh KC. 2009. Arabidopsis IRT3 is a zinc-regulated and plasma membrane localized zinc/iron transporter. *New Phytologist* **182**: 392–404.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**: 550.

Maere S, Heymans K, Kuiper M. 2005. BiNGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks. *Bioinformatics* **21**: 3448–3449.

Maurer F, Naranjo Arcos MA, Bauer P. 2014. Responses of a triple mutant defective in three iron deficiency-induced *BASIC HELIX-LOOP-HELIX* genes of the subgroup Ib(2) to iron deficiency and salicylic acid. *PLoS ONE* **9**: e99234.

Merlot S, de Sanchez Garcia la Torre V, Hanikenne M. 2018. Physiology and molecular biology of trace element hyperaccumulation. In: Echevarria G, van der Ent A, Morel JL, Baker AJM, eds. *Agromining: farming for metals. Extracting unconventional resources using plants*. Cham, Switzerland: Springer, 93–116.

Meyer C-L, Juraniec M, Huguet S, Chaves-Rodriguez E, Salis P, Isaure M-P, Goormaghtigh E, Verbruggen N. 2015. Intraspecific variability of cadmium tolerance and accumulation, and cadmium-induced cell wall modifications in the metal hyperaccumulator *Arabidopsis halleri*. *Journal of Experimental Botany* **66**: 3215–3227.

Meyer C-L, Kostecka AA, Saumitou-Laprade P, Créach A, Castric V, Pauwels M, Frérot H. 2010. Variability of zinc tolerance among and within populations of the pseudometallophyte species *Arabidopsis halleri* and possible role of directional selection. *New Phytologist* **185**: 130–142.

Meyer C-L, Pauwels M, Briset L, Godé C, Salis P, Bourceaux A, Souleman D, Frérot H, Verbruggen N. 2016. Potential preadaptation to anthropogenic pollution: evidence from a common quantitative trait locus for zinc and cadmium tolerance in metallicolous and nonmetallicolous accessions of *Arabidopsis halleri*. *New Phytologist* **212**: 934–943.

Meyer C-L, Peisker D, Courbot M, Craciun AR, Cazale AC, Desgain D, Schat H, Clemens S, Verbruggen N. 2011. Isolation and characterization of *Arabidopsis halleri* and *Thlaspi caerulescens* phytochelatin synthases. *Planta* **234**: 83–95.

Meyer C-L, Vitalis R, Saumitou-Laprade P, Castric V. 2009. Genomic pattern of adaptive divergence in *Arabidopsis halleri*, a model species for tolerance to heavy metal. *Molecular Ecology* **18**: 2050–2062.

Milner MJ, Seamon J, Craft E, Kochian LV. 2013. Transport properties of members of the ZIP family in plants and their role in Zn and Mn homeostasis. *Journal of Experimental Botany* **64**: 369–381.

Mondragon M, John-Arputharaj A, Pallmann M, Dresselhaus T. 2017. Similarities between reproductive and immune pistil transcriptomes of Arabidopsis species. *Plant Physiology* **174**: 1559–1575.

van de Mortel JE, Almar Villanueva L, Schat H, Kwekkeboom J, Coughlan S, Moerland PD, Loren Ver, van Themaat E, Koornneef M, Aarts MG. 2006. Large expression differences in genes for iron and zinc homeostasis, stress response and lignin biosynthesis distinguish *Arabidopsis thaliana* and the related metal hyperaccumulator *Thlaspi caerulescens*. *Plant Physiology* **142**: 1127–1147.

Nouet C, Charlier JB, Carnol M, Bosman B, Farnir F, Motte P, Hanikenne M. 2015. Functional analysis of the three *HMA4* copies of the metal hyperaccumulator *Arabidopsis halleri*. *Journal of Experimental Botany* **66**: 5783–5795.

Palmer CM, Hindt MN, Schmidt H, Clemens S, Guerinot ML. 2013. *MYB10* and *MYB72* are required for growth under iron-limiting conditions. *PLoS Genetics* **9**: e1003953.

Pauwels M, Vekemans X, Godé C, Frérot H, Castric V, Saumitou-Laprade P. 2012. Nuclear and chloroplast DNA phylogeography reveals vicariance among European populations of the model species for the study of metal tolerance, *Arabidopsis halleri* (Brassicaceae). *New Phytologist* **193**: 916–928.

Perea-García A, Garcia-Molina A, Andrés-Colás N, Vera-Sirera F, Pérez-Amador MA, Puig S, Peñarrubia L. 2013. Arabidopsis copper transport protein COPT2 participates in the cross talk between iron deficiency responses and low-phosphate signaling. *Plant Physiology* **162**: 180–194.

Pfaffl MW. 2001. A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Research* **29**: e45.

Pineau C, Loubet S, Lefoulon C, Chalies C, Fizames C, Lacombe B, Ferrand M, Loudet O, Berthomieu P, Richard O. 2012. Natural variation at the *FRD3* MATE transporter locus reveals cross-talk between Fe homeostasis and Zn tolerance in *Arabidopsis thaliana*. *PLoS Genetics* **8**: e1003120.

Ravet K, Touraine B, Boucherez J, Briat JF, Gaymard F, Cellier F. 2009a. Ferritins control interaction between iron homeostasis and oxidative stress in Arabidopsis. *Plant Journal* **57**: 400–412.

Ravet K, Touraine B, Kim SA, Cellier F, Thomine S, Guerinot ML, Briat JF, Gaymard F. 2009b. Post-translational regulation of AtFER2 ferritin in response to intracellular iron trafficking during fruit development in Arabidopsis. *Molecular Plant* **2**: 1095–1106.

Robinson NJ, Procter CM, Connolly EL, Guerinot ML. 1999. A ferric-chelate reductase for iron uptake from soils. *Nature* **397**: 694–697.

Rogers EE, Guerinot ML. 2002. FRD3, a member of the multidrug and toxin efflux family, controls iron deficiency responses in Arabidopsis. *Plant Cell* **14**: 1787–1799.

Ruijter JM, Ramakers C, Hoogaars WMH, Karlen Y, Bakker O, van den Hoff MJB, Moorman AFM. 2009. Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Research* **37**: e45.

Sarry J-E, Kuhn L, Ducruix C, Lafaye A, Junot C, Hugouvieux V, Jourdain A, Bastien O, Fievet JB, Vailhen D *et al.* 2006. The early responses of *Arabidopsis thaliana* cells to cadmium exposure explored by protein and metabolite profiling analyses. *Proteomics* **6**: 2180–2198.

Schaaf G, Honsbein A, Meda AR, Kirchner S, Wipf D, von Wirén N. 2006. AtIREG2 encodes a tonoplast transport protein involved in iron-dependent nickel detoxification in *Arabidopsis thaliana* roots. *Journal of Biological Chemistry* **281**: 25532–25540.

Schmid NB, Giehl RFH, Döll S, Mock H-P, Strehmel N, Scheel D, Kong X, Hider RC, von Wirén N. 2014. Feruloyl-CoA 6′-hydroxylase1-dependent coumarins mediate iron acquisition from alkaline substrates in Arabidopsis. *Plant Physiology* **164**: 160–172.

Seguela M, Briat JF, Vert G, Curie C. 2008. Cytokinins negatively regulate the root iron uptake machinery in Arabidopsis through a growth-dependent pathway. *Plant Journal* **55**: 289–300.

Shahzad Z, Gosti F, Frérot H, Lacombe E, Roosens N, Saumitou-Laprade P, Berthomieu P. 2010. The five AhMTP1 zinc transporters undergo different evolutionary fates towards adaptive evolution to zinc tolerance in *Arabidopsis halleri*. *PLoS Genetics* **6**: e1000911.

Shanmugam V, Lo JC, Wu CL, Wang SL, Lai CC, Connolly EL, Huang JL, Yeh KC. 2011. Differential expression and regulation of iron-regulated metal transporters in *Arabidopsis halleri* and *Arabidopsis thaliana* - the role in zinc tolerance. *New Phytologist* **190**: 125–137.

Shanmugam V, Lo J-C, Yeh K-C. 2013. Control of Zn uptake in *Arabidopsis halleri*: a balance between Zn and Fe. *Frontiers in Plant Science* **4**: 281.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research* **13**: 2498–2504.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**: 3210–3212.

Smith-Unna R, Boursnell C, Patro R, Hibberd JM, Kelly S. 2016. TransRate: reference-free quality assessment of *de novo* transcriptome assemblies. *Genome Research* **26**: 1134–1144.

Stein RJ, Höreth S, de Melo JRF, Syllwasschy L, Lee G, Garbin ML, Clemens S, Krämer U. 2017. Relationships between soil and leaf mineral composition are element-specific, environment-dependent and geographically structured in the emerging model *Arabidopsis halleri*. *New Phytologist* **213**: 1274–1286.

Talke IN, Hanikenne M, Krämer U. 2006. Zinc-dependent global transcriptional control, transcriptional deregulation, and higher gene copy number for genes in metal homeostasis of the hyperaccumulator *Arabidopsis halleri*. *Plant Physiology* **142**: 148–167.

Usadel B, Poree F, Nagel A, Lohse M, Czedik-Eysenberg A, Stitt M. 2009. A guide to using MapMan to visualize and compare Omics data in plants: a case study in the crop species, Maize. *Plant, Cell & Environment* **32**: 1211–1229.

Verbruggen N, Hermans C, Schat H. 2009. Molecular mechanisms of metal hyperaccumulation in plants. *New Phytologist* **181**: 759–776.

Vert G, Barberon M, Zelazny E, Seguela M, Briat JF, Curie C. 2009. Arabidopsis IRT2 cooperates with the high-affinity iron uptake system to maintain iron homeostasis in root epidermal cells. *Planta* **229**: 1171–1179.

Vert G, Grotz N, Dédaldéchamp F, Gaymard F, Guerinot ML, Briat J-F, Curie C. 2002. IRT1, an Arabidopsis transporter essential for iron uptake from the soil and for plant growth. *Plant Cell* **14**: 1223–1233.

Wang HY, Klatte M, Jakoby M, Baumlein H, Weisshaar B, Bauer P. 2007. Iron deficiency-mediated stress regulation of four subgroup Ib BHLH genes in *Arabidopsis thaliana*. *Planta* **226**: 897–908.

Weber M, Harada E, Vess C, Roepenack-Lahaye EV, Clemens S. 2004. Comparative microarray analysis of *Arabidopsis thaliana* and *Arabidopsis halleri* roots identifies nicotianamine synthase, a ZIP transporter and other genes as potential metal hyperaccumulation factors. *Plant Journal* **37**: 269–281.

Willems G, Dräger DB, Courbot M, Gode C, Verbruggen N, Saumitou-Laprade P. 2007. The genetic basis of zinc tolerance in the metallophyte *Arabidopsis halleri* ssp. *halleri* (Brassicaceae): an analysis of quantitative trait loci. *Genetics* **176**: 659–674.

Willems G, Frérot H, Gennen J, Salis P, Saumitou-Laprade P, Verbruggen N. 2010. Quantitative trait loci analysis of mineral element concentrations in an *Arabidopsis halleri* × *Arabidopsis lyrata petraea* F_2 progeny grown on cadmium-contaminated soil. *New Phytologist* **187**: 368–379.

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.

Zhao FJ, Lombi E, Breedon T, McGrath SP. 2000. Zinc hyperaccumulation and cellular distribution in *Arabidopsis halleri*. *Plant, Cell & Environment* **23**: 507–514.

Zhi-Liang H, Bao J, Reecy J. 2008. CateGOrizer: a web-based program to batch analyze gene ontology classification categories. *Online Journal of Bioinformatics* **9**: 108–112.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information tab for this article:

Fig. S1 *Arabidopsis halleri de novo* transcriptome assembly and annotation pipeline.

Fig. S2 Gene ontology term frequencies in transcriptome annotations of *Arabidopsis halleri*, *A. thaliana* and *A. lyrata*.

Fig. S3 Root-to-shoot ratio of average metal nutrient concentrations.

Fig. S4 Metal nutrient concentrations in roots and shoots of PL22 and I16 *Arabidopsis halleri* metallicolous populations.

Fig. S5 *IRT1* relative transcript level, IRT1 protein level, and citrate concentration in root samples of PL22 and I16 *Arabidopsis halleri* metallicolous populations.

Fig. S6 *HMA2*, *bHLH038*, *bHLH039* and *bHLH100* relative transcript levels in root samples of PL22 and I16 *Arabidopsis halleri* metallicolous populations.

Fig. S7 Correlation between RNA-Seq and quantitative RT-PCR validation data.

Table S1 Quantitative RT-PCR primer sequences and primer efficiencies

Table S2 Annotation of 34 818 transcripts of *Arabidopsis halleri* (Polish accession PL22) reference transcriptome

Table S3 Comparative assembly metrics of the *Arabidopsis halleri* (this study), *A. lyrata* and *A. thaliana* transcriptomes using TRANSRATE

Table S4 Comparative assembly completeness assessment of the *Arabidopsis halleri* (this study), *A. lyrata* and *A. thaliana* transcriptomes using BUSCO

Table S5 Groups of orthologous genes (orthogroups) in *Arabidopsis lyrata*, *A. thaliana* and *A. halleri*

Table S6 Gene ontology term enrichment in 24 *Arabidopsis halleri*-specific orthogroups

Table S7 Lists of differentially expressed genes from all pairwise comparisons described in Fig. 5

Table S8 List of 35 genes corresponding to the 'response to cadmium ion' GO term among 507 overexpressed genes in control conditions in I16

Notes S1 Discussion of the characteristics of the *Arabidopsis halleri de novo* transcriptome assembly.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.