the **genetics**society

ARTICLE

Genetic architecture of a plant adaptive trait: QTL mapping of intraspecific variation for tolerance to metal pollution in Arabidopsis halleri

Marie-Joe Karam[1] · Dima Souleman[1] · M. Sol Schvartzman[2] · Sophie Gallina[1] · Julien Spielmann[2] · Charles Poncet[3] · Olivier Bouchez[4] · Maxime Pauwels[1] · Marc Hanikenne[2] · Hélène Frérot[1]

Abstract
Anthropogenic activities are among the main drivers of global change and result in drastic habitat modifications, which represent strong evolutionary challenges for biological species that can either migrate, adapt, or disappear. In this context, understanding the genetics of adaptive traits is a prerequisite to enable long-term maintenance of populations under strong environmental constraints. To examine these processes, a QTL approach was developed here using the pseudometallophyte *Arabidopsis halleri*, which displays among-population adaptive divergence for tolerance to metallic pollution in soils. An F2 progeny was obtained by crossing individuals from metallicolous and non-metallicolous populations from Italian Alps, where intense metallurgic activities have created strong landscape heterogeneity. Then, we combined genome de novo assembly and genome resequencing of parental genotypes to obtain single-nucleotide polymorphism markers and achieve high-throughput genotyping of the progeny. QTL analysis was performed using growth parameters and photosynthetic yield to assess zinc tolerance levels. One major QTL was identified for photosynthetic yield. It explained about 27% of the phenotypic variance. Functional annotation of the QTL and gene expression analyses highlighted putative candidate genes. Our study represents a successful approach combining evolutionary genetics and advanced molecular tools, helping to better understand how a species can face new selective pressures of anthropogenic origin.

These authors contributed equally: Marie-Joe Karam and Dima Souleman

These authors contributed equally: Marc Hanikenne and Hélène Frérot

✉ Hélène Frérot
helene.frerot@univ-lille.fr

[1] Université de Lille, CNRS, UMR 8198 - Unité Evolution-Ecologie-Paléontologie, 59000 Lille, France

[2] InBioS - PhytoSystems, Functional Genomics and Plant Molecular Imaging, University of Liège, Liège, Belgium

[3] Plateforme Nationale Stratégique INRA - GENTYANE, UMR 1095 INRA/UBP GDEC - Site de Crouel, 5 chemin de Beaulieu, 63039 Clermont-Ferrand Cedex 2, France

[4] INRA, US 1426, GeT-PlaGe, Genotoul, Castanet-Tolosan, France

Introduction

Anthropogenic activities are one of the main sources of global change (Steffen et al. 2006). Locally, they may cause habitat modification, fragmentation, or even destruction, thus generating spatially heterogeneous landscapes. Consequently, populations may also be exposed to novel environmental conditions and selective pressures. Persistence of populations facing new environmental conditions relies on the capacity of individuals to migrate or rapidly adjust their phenotype (Alberti et al. 2017). Such rapid phenotypic changes may result from non-genetic and/or genetic processes. However, genetic processes are still considered as the main drivers of adaptation in the long term (Charlesworth et al. 2017).

In the context of anthropogenic modifications that occur rapidly, i.e., during a limited number of generations, it remains uncertain whether evolution through genetic adaptation can be generally expected. Examples of genetic adaptation to anthropogenic (mostly urban) habitats are accumulating (Johnson and Munshi-South 2017), but

remain sparse in plants. Some edifying examples include seed dispersal traits in response to fragmentation in the weed *Crepis sancta* (Cheptou et al. 2008), decreased chemical defenses in response to minimum winter temperatures in *Trifolium repens* (Thompson et al. 2016), or early flowering to avoid drought generated by climate fluctuation in *Brassica rapa* (Franks 2011).

Moreover, the possibility of rapid genetic adaptation may depend on the evolutionary dynamics of ecologically relevant traits (Schoener 2011). These traits can be relatively simple (e.g. tolerance to herbicides, insecticides, or pollutants) or more complex (life history traits, e.g. body size, fecundity, or mating behavior) (Reznick and Ghalambor 2001), which particularly raises the question of their genetic bases. Interestingly, several authors indicated that genetic bases of adaptive traits are still largely unknown, and also reported appropriate genomic approaches to investigate them (Bergelson and Roux 2010; Savolainen et al. 2013). High-throughput molecular tools have proved to be helpful for phenotype–genotype association studies requiring moderate (e.g. for Quantitative Trait Locus Mapping) to high (e.g. for Genome-Wide Association Mapping) density of markers along the genome.

This study aims at exploring the genetic bases of a trait involved in the adaptation to a particular case of anthropogenic impact, namely metallic pollution. Mining and industrial activities are indeed responsible for the expansion of areas highly contaminated with zinc (Zn), cadmium (Cd), and lead, also called calamine areas. Survival in calamine habitats usually suggests the adaptive acquisition of metal tolerance, a trait defined in evolutionary terms as the capacity of a plant to survive and reproduce on a metal-contaminated soil without showing any toxicity symptom (Macnair and Baker 1994). The evolution of metal tolerance is a typical adaptive response recorded for plants occurring in metal-polluted urban environments (McDonnell and Hahs 2015). In pseudometallophyte species, which are able to grow on both contaminated (metallicolous populations, M) and non-contaminated (non-metallicolous populations, NM) soils (Antonovics et al. 1971), metal tolerance usually occurs only or at highest levels in M populations, suggesting adaptive evolution. This most probable adaptive divergence among M and NM populations in pseudometallophytes makes them relevant models for the study of the genetics of adaptive traits associated with anthropogenic habitats.

Among pseudometallophytes, *Arabidopsis halleri* is a montane species distributed throughout Europe (Northern France, Silesia in Poland, Harz in Germany and Northern Italy) (Clapham and Akeroyd 1993; O'Kane and Al-Shehbaz 1997). It has been extensively studied for its Zn and Cd tolerance properties (Bert 2000; Meyer et al. 2015; Pauwels et al. 2006). According to all these studies, Zn and Cd tolerance are quantitative and species-wide properties, but tolerance levels are usually higher in M populations. It is also a close relative to the model *Arabidopsis thaliana* from which it diverged at least 5 Myr ago (Bechsgaard et al. 2006). *A. halleri* and *A. thaliana* share about 94% DNA sequence identity in coding part, despite the fact that 2n = 16 with a genome size of approximately 250 Mbp/1C in *A. halleri* (Briskine et al. 2016) and 2n = 10 with a genome size of 135 Mbp in *A. thaliana* (Hohmann et al. 2015; Novikova et al. 2016). *A. halleri* is even closer to *Arabidopsis lyrata*, a species sensitive to Cd and Zn, with a similar genome size (210 Mbp/1C) and the same number of chromosomes (Hohmann et al. 2014). *A. halleri* and *A. lyrata* may have diverged between 0.4 and 2 Myr ago, respectively (Novikova et al. 2016) and can be successfully intercrossed.

In the present study, we developed an innovative approach to investigate the genetic bases of Zn tolerance, an adaptive trait associated with metal pollution of anthropogenic origin. NM and M *A. halleri* individuals from geographically close populations were collected from the Lombardy region in Italy, where intense metallurgic activities have created landscape heterogeneity for metal pollution. An F2 intraspecific cross was obtained to perform QTL analysis. The linkage map was constructed using whole-genome sequencing and de novo assembly, genome resequencing of parental genotypes, and the mapping of single-nucleotide polymorphisms (SNPs) segregating in the F2 progeny. Four biomass-related traits and one physiological trait were used to phenotype Zn tolerance in controlled conditions and to detect underlying QTL regions. Functional annotation of the QTL regions and gene expression analyses further allowed to consider the putative molecular mechanisms involved in adaptation to metal-polluted habitats.

Materials and methods

Sampling sites

This study was carried out using accessions of *A. halleri* from two different sites in Bergamo province of Italy, southern Alps (Frérot et al. 2018). This Italian region is one of the European *A. halleri* habitats where some metallurgic activities (either mining or industrial) have intensively occurred (Fig. S1). The I35 M population is located at a mine exit in the "Val del Riso" valley, where the landscape includes old mine wastes and ruins resulting from mining activities. In I35, the mean concentration of ammonium-EDTA extractable Zn is about 11,417 kg mg^{-1} (Frérot et al. 2018). The I30 NM population is located in another valley about 40 km away from the first valley near the

"Sommaprada" locality. This site is covered by hayfields and forests and is not contaminated by metal trace elements.

Crossing scheme to produce a F2 segregating progeny

A. halleri being self-incompatible, the production of an F2 progeny requires crossing two independently obtained F1 hybrids. The F1 progenies were produced from two crosses involving one NM female and one M male parents called I30.16 and I35.12 for one cross, and I30.13 and I35.6 for the other (Fig. S2). One individual from each F1 progeny was randomly selected to be crossed. Seeds were collected on the I30.16 × I35.12 F1 individual and sowed to obtain the F2 progeny. DNA extracts were used to positively assign 175 F2 individuals to the crossing parents using a multiplex panel of eight microsatellite markers (Godé et al. 2012).

Assessment and statistical analysis of Zn tolerance

Six clones of each F2 genotype were obtained by cuttings, generating a total of 1050 plants. Cuttings were grown in the greenhouse on sand during 8 weeks to allow root development. Plants were then transferred to pots containing 1 liter of nutrient solution for 8 weeks in a growth control chamber (temperature: 20 °C day and 15 °C night; light: 14 h day and 10 h night; hygrometry: 80% for the first two weeks of acclimation and then 65%). Three clones per genotype were randomly distributed in 175 pots (three plants/pot), and the other three clones were distributed in a second lot of 175 pots. The 350 pots were then randomly placed on three quarters of a circular rotating table to homogenize culture conditions. The parental and F1 individuals died from viral infection before this tolerance experiment.

The nutrient solution contained essential micro- and macro-elements for plant growth, including 0.02 mM FeEDDHA, 0.1 µM $(NH_4)_6Mo_7O_{24}\cdot H_2O$, 0.1 µM $CuSO_4\cdot 5H_2O$, 0.025 mM H_3BO_3, 2 µM $MnSO_4\cdot H_2O$, 1 µM KCl, 0.1 µM NaCl, 0.5 mM $MgSO_4\cdot 7H_2O$, 1 mM $NH_4H_2PO_4$, 2 mM $Ca(NO_3)_2\cdot 4H_2O$, 3 mM KNO_3 and 10 µM of Zn added as $ZnSO4\cdot 7H_2O$. Metal bioavailability was insured using a buffer solution of 2 mM of MES (2-(*N*-morpholino)ethanesulfonic acid) whose pH was adjusted to 5 using a KOH solution. To avoid deficiency and root anoxia, the nutrient solution was changed each week for all pots. The 10 µM concentration of Zn was maintained for all pots for 2 weeks for plant acclimation. After 2 weeks, the Zn tolerance test started (week 0, W_0) by applying a 2000 µM Zn concentration in one lot of pots for the six remaining weeks (Meyer et al. 2010).

Zn tolerance was estimated by measuring root length and leaf width (mean of three mature leaves) at W_0, and two

(W_2) and four (W_4) weeks after W_0. Root and shoot dry biomasses were measured at W_6 after plant harvesting, cleaning using osmosis water, and oven-drying at 60 °C for 48 h. Photosystem II (PSII) yield (mean of three mature leaves) was measured at W_2 and W_4. The PSII yield is a proxy of the photosynthetic efficiency. It takes into account the loss of energy by fluorescence at the photosystem II level. In our experiment, we measured it using the PAM-2100 portable Chlorophyll fluorometer (Walz, Effeltrich, Germany). In order to standardize the lightning conditions across the plants, we did the measures in an artificially enlightened custom chamber. The PSII yield was calculated as follows: $(F'_m - F_t)/F'_m$, where F_t is the fluorescence level under the chamber light and F'_m the maximum fluorescence measured under a saturating pulse (Genty et al. 1989).

For each trait and each genotype, three values of the tolerance index ("*TI*") were obtained by dividing the phenotypic value of each clone in the polluted condition by the median phenotypic value over three clones in the non-polluted condition. The tolerance index of each genotype was then obtained by averaging over the clones (Meyer et al. 2010).

For each trait, the broad-sense heritability (H^2) was estimated with a variance component analysis using the Restricted Maximum Likelihood (REML) method implemented in the R/VCA package (version 1.3.3). For each trait, the genotype was considered as a random effect and only individuals with available phenotypic data for the three clones were considered (i.e., from 136 to 161 individuals depending on trait and condition, 153 individuals on average). After variant component estimation, the broad-sense heritability was computed by dividing the genotypic variance by the total variance (Lynch and Walsh 1998). The significance of the heritability values (H1 hypothesis: $H^2 \neq 0$) was assessed using a custom exact statistical test with 1000 permutations.

Comparisons of mean values between non-polluted and polluted conditions were performed using the non-parametric Wilcoxon statistical test for paired data, implemented in R 3.1.2. In order to visualize the correlations between the traits, a Principle Component Analysis (PCA) implemented in R 3.1.2 (R package FactoMineR) was applied on the trait values at W_4 for root length and leaf width and at W_6 for root and shoot dry biomass, in both conditions and on their tolerance indices.

A. halleri de novo genome assembly

The sequencing and assembly of an *A. halleri* reference genome is detailed in supporting information (Method S1; Table S1). For this assembly, the genomic DNA of one individual plant of a metallicolous Polish population (PL22, cited in Meyer et al. 2010) was extracted. The

Fig. 1 SNP filtration after SNP and genotype calling. Filters in the black dashed line frames (resp., black continuous line frame) were applied according to genotyping requirements (resp., genetic mapping suitability)
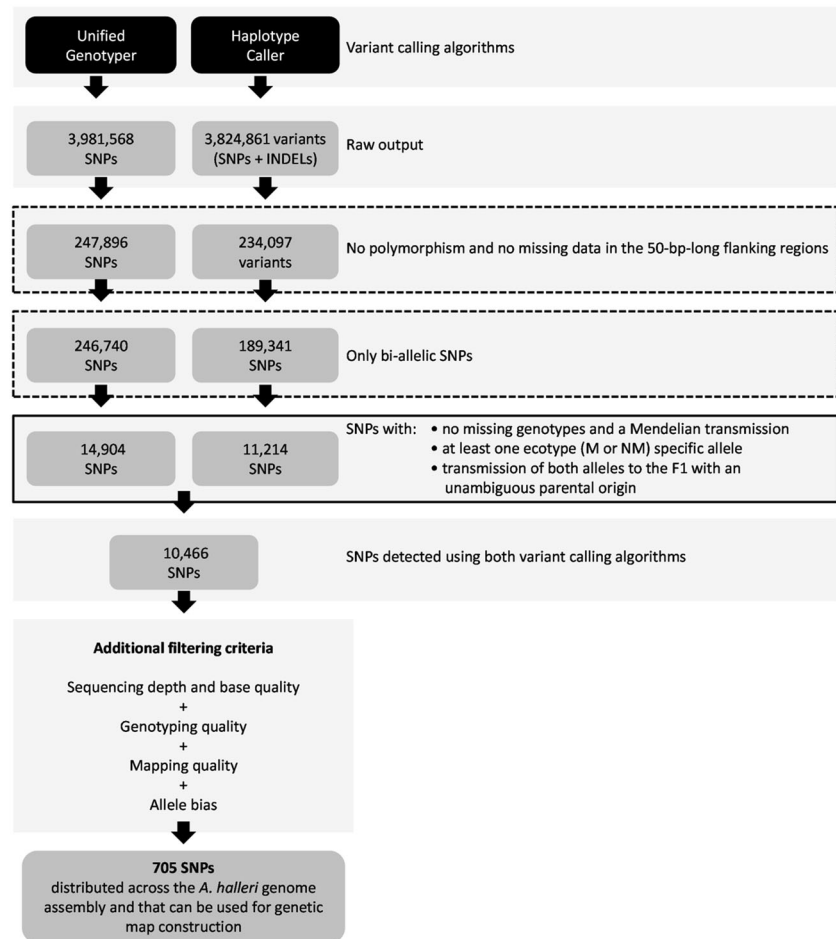
Parents and F1 high-throughput sequencing and read processing

DNA from the four parents and two F1 individuals was extracted and paired-end sequenced as detailed in Method S2. Library preparation and sequencing were done at the GET PLAGE platform (http://get.genotoul.fr/). Raw reads were checked for quality and filtered (Method S2). The reads have been deposited at the NCBI Database with BioProject identification number PRJNA427096.

SNP selection and genotyping

All the steps of SNP and genotype calling and filtering are detailed in supporting information (Method S3) and in Fig. 1. Genotyping of the F2 individuals using the KASP chemistry is described in Method S4.

reads and genome assembly have been deposited at the NCBI Database with BioProject identification number PRJNA422908.

Construction of the genetic map

The genetic map construction was done using the *R/qtl* package version 1.39–5 (Broman et al. 2003). Markers showing significant segregation distortion (Chi-square test, $\alpha_{\text{Bonf}} \simeq 14\text{e}^{-05}$), i.e. differing from the expected 1:2:1 ratio for an F2 progeny, were discarded.

For linkage group construction, the logarithm-of-odds (LOD) score minimum threshold was set to 10 and the maximum recombination fraction threshold to 0.5. Ordering markers along each linkage group was done by fixing the positions of previously ordered markers and adding one marker at a time in the position giving the minimum number of obligate crossovers. The genotyping error rate taken into consideration in the genetic map construction was estimated at 7% (see Results) according to the method described in Method S4. We used Kosambi's mapping function to infer the genetic distances from the recombination rates (Kosambi 1944). The scaffolds harboring SNPs included in the genetic map of *A. halleri* were extracted from the *A. halleri* genome assembly and grouped by linkage groups to

study the shared synteny of our genetic map with the genomes on *A. lyrata* and *A. thaliana*.

QTL mapping

QTL mapping was achieved using the tolerance indices of the traits. We started using the Interval Mapping (IM) method implemented in *R/*qtl package version 1.39-5 (Broman and Sen 2009). The LOD score representing the likelihood of a QTL was computed every centiMorgan (cM) along the linkage groups. The genome-wide significance threshold for the LOD score was set for each trait using a permutation test (1000 permutations) and corresponding to a 0.05 significance threshold (Churchill and Doerge 1994). For the detected QTLs, an approximate Bayesian credible interval was computed, with a 0.95 coverage probability. We computed the QTL additive "*a*" and dominance "*d*" effects using the *effectscan* function implemented in *R/qtl*. The QTL additive effect is calculated as the half difference between the means of the phenotype values for the homozygotes for a given trait. The QTL dominance effect is calculated as the difference between the mean phenotypic value of the heterozygotes and the midpoint between the mean phenotypic values for the homozygotes for a given trait. The degree of dominance for a QTL was computed as the absolute value of the ratio of dominance effect to the additive effect ($|d/a|$). A QTL is considered additive if $|d/a|$ is lower than 0.2, partially dominant if $|d/a|$ is between 0.2 and 0.8, dominant if $|d/a|$ is between 0.8 and 1.2, and overdominant if $|d/a|$ is greater than 1.2.

Fine mapping in the major QTL region

After QTL identification, additional SNP markers from the Illumina HiSeq data were sought in order to increase the density of markers in the QTL region and to reduce the QTL credible interval. Informative *A. halleri* scaffolds were identified, and some SNPs issued from a slightly relaxed filtering procedure were selected (Method S5).

After marker densification in the QTL region, the genetic map was reconstructed, and QTL mapping was redone. To get more statistical power and more precision in defining the QTL region, we applied the multiple QTL mapping (MQM) method implemented in *R/qtl* on the QTL linkage group. To deal with missing data, we used the "*mqmaugment*" function with the imputation strategy to model multiple individual-variants per individual associated with a probability and keep the most likely ones ("data augmentation" method implemented in *R/qtl*). The minimum probability used to consider an individual-variant was set to 0.16, corresponding to the percentage of missing data in the linkage group 4 as suggested in the *R/qtl* manual. The maximum number of individual variants to model per individual was set to $3^{6.401} \simeq 1133$ corresponding to a mean of 6.401 missing markers per individual. We did the MQM analysis with the unsupervised backward elimination method implemented in *R/qtl* to analyze cofactors. We chose a step size of 1 cM and a window size of 25 cM. Automatic co-factor selection was done using the "*mqmsetcofactors*" function with the option to set a cofactor every 5 cM.

Gene annotation in the major QTL region

The *A. lyrata* genomic annotation v2.1 (genome assembly v.1, https://phytozome.jgi.doe.gov) was used as a reference to annotate the *A. halleri* scaffolds by homology analysis (Method S6). The *A. lyrata* genes identified in the QTL region were then compared to the *A. thaliana* genes using *blastn* (*A. thaliana* genome annotation version: TAIR10). Gene description was extracted using the TAIR (The Arabidopsis Information Resource) platform (https://www.arabidopsis.org/tools/bulk/genes/index.jsp). When no hits were found, genes were compared using *blastx* to the *swissprot* database (last update 5/11/2017, ftp://ftp.ncbi.nlm.nih.gov/blast/db/).

Gene expression analysis

The plants used for gene expression analysis came from the *A. halleri* collection maintained in non-polluted conditions in greenhouse (University of Lille), so that results may reflect constitutive expression levels. Leaves from four F2 plants from each homozygous genotype at the closest SNP to the QTL were collected: $F2_{I30}$ (resp., $F2_{I35}$) had two identical alleles specific to I30 (resp., I35) at the "*fm1776_3057*" SNP maker. RNA extraction and expression analyses are detailed in Table S2 and Method S7.

Results

Phenotypic variation in the F2 progeny

The 175 F2 individuals were phenotyped for five traits over the course of a 6-week tolerance assay. For leaf width and root length measured at W_0, W_2, and W_4, the comparison of trait means and distributions between non-polluted (NP) and Zn-polluted (P) conditions suggested a tendency towards increasing values in P conditions with time (Table 1; Fig. 2). The same tendency was observed at W_6 for shoot and root dry biomass (Table 1; Fig. 2). Accordingly, for all morphological traits, tolerance indices (TIs) were elevated, with TI values frequently higher than 1 (Table 1; Fig. 3). In contrast, the distribution of PSII yield values at W_2 and W_4 was significantly lower in P than in NP conditions (Table 1;

Table 1 Summary statistics and broad-sense heritabilities of the traits in the non-polluted and Zn-polluted conditions and their tolerance indices

	Mean ± SD			Broad-sense heritability H^{2a}		
	Non-polluted condition	Zn-polluted condition	Tolerance index	Non-polluted condition	Zn-polluted condition	Tolerance index
Mean leaf width at W_0 (cm)	1.11 ± 0.41	1.04 ± 0.32	1.00 ± 0.31	0.32***	0.33***	0.32***
Mean leaf width at W_2 (cm)	1.20 ± 0.38	1.32 ± 0.32	1.21 ± 0.52	0.27***	0.37***	0.61***
Mean leaf width at W_4 (cm)	1.20 ± 0.32	1.43 ± 0.28	1.26 ± 0.42	0.23***	0.48***	0.68***
Root length at W_0 (cm)	8.19 ± 1.70	8.16 ± 1.61	1.01 ± 0.24	0.23***	0.14***	0.27***
Root length at W_2 (cm)	9.85 ± 1.98	11.96 ± 2.10	1.24 ± 0.30	0.16***	0.29***	0.46***
Root length at W_4 (cm)	10.33 ± 2.29	12.34 ± 1.76	1.28 ± 0.64	0.22***	0.26***	0.88***
Shoot dry biomass (g)	0.56 ± 0.35	0.65 ± 0.34	2.16 ± 3.03	0.27***	0.41***	0.58***
Root dry biomass (g)	0.05 ± 0.04	0.15 ± 0.07	4.76 ± 4.22	0.06	0.33***	0.60***
Mean PSII yield × 1000 at W_2	735.04 ± 31.17	692.73 ± 52.88	0.94 ± 0.08	0.08*	0.38***	0.41***
Mean PSII yield × 1000 at W_4	747.20 ± 17.07	706.62 ± 54.19	0.94 ± 0.08	0.15**	0.33***	0.37***

[a] A custom statistical exact test was applied doing 1000 permutations

*p-value ≤ 0.05, **p-value ≤ 0.01, and ***p-value ≤ 0.001

Fig. 2). TIs for PSII yield also showed the lowest mean values (Table 1; Fig. 3). However, in the PCA analysis, the morphological traits and PSII yield appeared uncorrelated in NP condition (Fig. 4a), and this was clearer in P condition (Fig. 4b) and for tolerance indices (Fig. 4c).

The broad-sense heritabilities (H^2) of the traits in P and NP conditions and of the TIs were estimated and summarized in Table 1. For all traits an increase in the heritability values was observed in the P compared to NP conditions (except root length at W_0). Moreover, for all morphological traits, higher heritability values based on TIs were observed compared to values obtained for the same traits in P conditions (except leaf width at W_0).

Genotyping of the F2 progeny

To build a high-resolution map, the genome of four parents and the two F1 individuals used in the I30 × I35 crossing scheme were sequenced to detect high-quality SNPs. Processed paired reads (841,110,582 reads, representing 90.46% of the raw reads) were mapped against the *A. halleri* genome assembly. The average number of processed paired reads across the four parents and the two F1 individuals was 140,185,097 ± 38,924,730 reads. The average number of mapped reads was 66,088,413 ± 9,071,361, with a mean depth of coverage 40.37 ± 5.54X and a genome assembly coverage 81.10 ± 0.59% across these individuals (Table S3).

Mendelian violations between the parents and F1 individuals in the detected SNPs were examined with *GATK*. 95% of the examined SNPs had a quality score equal to or above 100 (risk of false polymorphism $\leq 10^{-10}$). None of these SNPs showed Mendelian violation confirming that the parents were the true parents of the F1 individuals.

After applying highly stringent filters (Method S3; Fig. 1), 705 high-quality SNPs were selected (Table S4, column "high_quality_SNP"). Out of these, 377 SNPs were randomly chosen for genotyping. Seven additional SNPs were intentionally chosen because they were close to candidate genes for Zn tolerance (Table S4, columns "candidate_genes" and "mapping_SNP").

The 175 individuals of the F2 progeny were then genotyped using the KASP technology (384 SNPs × 175 individuals). The results showed a conversion rate (number of polymorphic SNPs over the total number of SNPs) of 93% approximately. The percentage of missing data per individual had a mean of 22.5% ± 20.5. The percentage of missing data per SNP had a mean of 22.8% ± 23.1. Moreover, a number of additional technical controls were analyzed to ensure the genotyping quality. First, plant material replicates (independent DNA extractions) gave a mean error rate at the intra-plate level of 7.5% ± 4.1. DNA replicates (a single DNA extraction) gave a mean error rate at the inter-

Fig. 2 Trait distributions in the non-polluted (NP) and Zn-polluted (P) conditions for all genotypes. Density plots are drawn behind the boxplots. Boxplots middle, lower, and upper hinges correspond to the median, first, and third quartiles, respectively. The upper (resp., lower) whisker goes from the upper (resp., lower) hinge to the maximum value no further than (resp., minimum value at most) 1.5×IQR from (resp., of) the hinge. Points beyond whiskers are plotted individually. IQR is the inter-quartile range. The p-values represented below the plots are the results of the Wilcoxon statistical test for paired data. ** indicates p-values $\leq 10^{-3}$. * indicates p-values $\leq 10^{-2}$
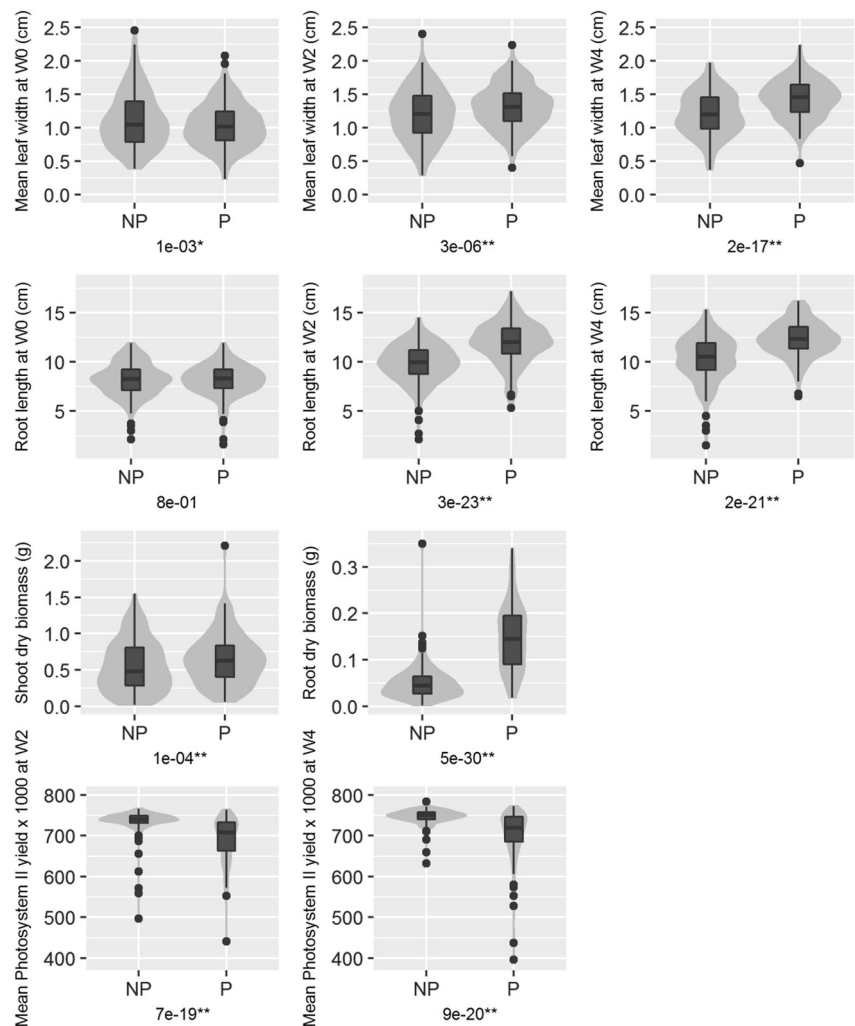

plate level of 4.7% ± 0.3. Overall, the global genotyping mean error rate to use for genetic map construction was estimated to 7%.

Genetic map construction

Six individuals and 27 SNPs with more than 25% of missing data were discarded from the analysis. As three pairs of individuals shared more than 99% of genotype identity, one individual from each of these pairs was discarded. Five markers showing significant segregation distortion were further discarded. As a result, 167 F2 individuals and 352 SNP markers were used for genetic map construction.

The genetic map construction resulted in 17 linkage groups, from which the eight largest groups carried 340 markers. Eight linkage groups corresponded to the expected number of chromosomes in A. halleri. In these linkage groups, the number of markers per linkage group varied

from 29 to 61. Linkage group sizes varied from 57.7 to 124.8 cM. The average spacing and maximum spacing between markers per linkage group varied between 1.3 and 3.0 and between 9.9 and 61.4 cM, respectively (Fig. 5).

Anchoring the genomic scaffolds on the genetic map

The mapped SNPs (i.e. used to build the genetic map) were included in 246 scaffolds of the A. halleri genome assembly described in this study (see Methods). These scaffolds had a minimum, maximum, and median size of 1003, 485,225, and 82,380 bp, respectively. They covered 25,226,224 bp (~15.8%) of the genome assembly.

As expected, the ordering of scaffolds on the genetic map revealed a strong signal of shared synteny between A. halleri and its close relative A. lyrata (Fig. S3). Each linkage group in A. halleri corresponded mostly to one chromosome in A. lyrata. A. thaliana chromosome shared

Fig. 3 Tolerance index distributions for all traits for all genotypes, reported on the same scale. Density plots are drawn behind the boxplots. Boxplots middle, lower, and upper hinges correspond to the median, first, and third quartiles, respectively. The upper (resp., lower) whisker goes from the upper (resp, lower) hinge to the maximum value no further than (resp., minimum value at most) 1.5×IQR from (resp., of) the hinge. Points beyond whiskers are plotted individually. IQR is the inter-quartile range
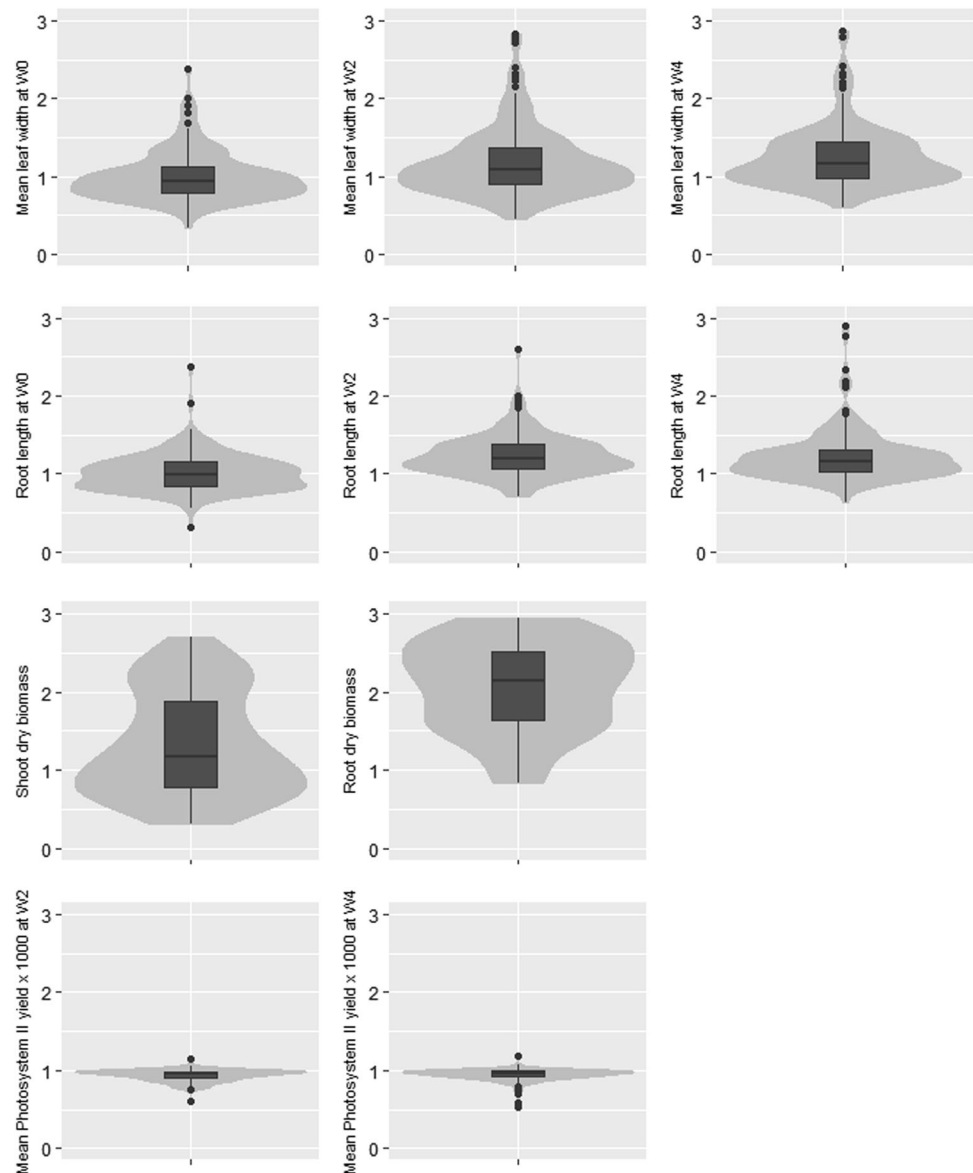
considerable syntenic blocks with at least two *A. halleri* linkage groups as previously described (Roosens et al. 2008; Willems et al. 2007).

QTL detection and effect

QTL detection through the IM approach was achieved using the tolerance indices of the traits. One significant QTL region associated with the PSII yield was detected. It is worth noting that the QTL signal was also significant for the same trait in the P but not in the NP condition, suggesting that the covered genomic region contribute to the genotype × environment interaction (data not shown). The QTL region is located on LG4. The LOD score significance threshold ($\alpha = 0.05$) was set at the genome-wide scale to 3.22 and 3.04 for PSII yield measurements made at W_2 and

W_4, respectively. The association was significant at both W_2 and W_4 and more pronounced at W_4 (Fig. 6a; Fig. S4a). At W_2 (resp., W_4), the maximum LOD score was of 5.96 (resp., 8.19), located at the position of 5 cM (resp., 4 cM) at a pseudo-marker position. The credible interval with a 0.95 coverage probability was between 0 and 10 cM for W_2 and 0 and 8 cM for W_4. The contribution of the QTL to the phenotypic variance was 18.02% for W_2 and 27.19% for W_4.

Alignment of the *A. halleri* genomic scaffolds flanking the PSII yield QTL with the *A. lyrata* chromosome 4 delimited the QTL to a region of 5,412,659 bp (i.e. between positions 606,990 and 6,019,649 of chromosome 4) in *A. lyrata*, called here the "*Alqtl*" region. The reciprocal alignment of this *A. lyrata* sequence against the *A. halleri* RepeatMasked genome assembly revealed 279 genomic scaffolds around or within the QTL region. Nine of these

Fig. 4 Principle component analyses on the trait values in the **a** non-polluted and **b** Zn-polluted condition and **c** for the trait tolerance indices. These analyses were done on the F2 individuals. The two first components and the explained variance (in %) are plotted. NP = non-polluted condition; P = Zn-polluted condition; TI = tolerance indices; LW4 (resp., RL4, PSII4) = mean leaf width (resp., root length, mean Photosystem II yield × 1000) measured after 4 weeks from the start of the Zn tolerance experiment; SB (resp., RB) = shoot dry biomass (resp., root dry biomass) measured at harvesting 6 weeks after the Zn tolerance experiment started
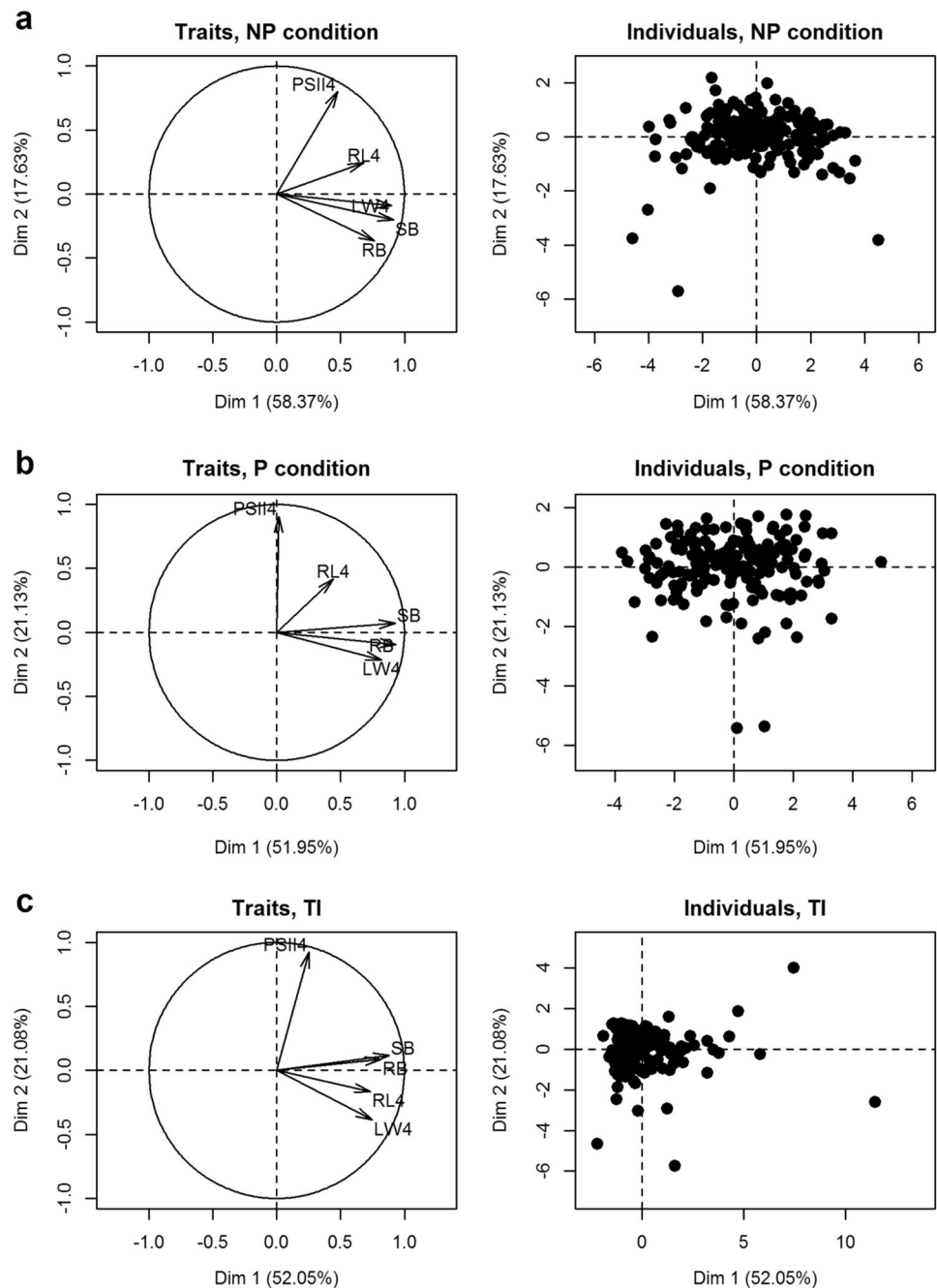
scaffolds included SNPs that passed our quality filters but had not been used for genotyping. These nine scaffolds were roughly distributed between 1 bp and 4×10^6 bps of the *Alqtl* region (Fig. S5). Seven SNPs yielded readable genotypes usable for fine mapping. After their addition to the initial dataset, the genetic map was constructed again and, as expected, the new SNPs covered the region of interest, i.e. the gap between the scaffolds "*65*" and "*1964*" in the expected order (Fig. S6). This improved QTL detection procedure conducted to a finer delimitation of the QTL region. IM on the tolerance indices of PSII yield at W_2

and W_4 detected a significant QTL signal in the expected region of LG4 (Fig. 6b; Fig. S4b). The maximum LOD score was of 10.6, located at the position of 5.28 cM at the SNP "*1776_3057*". After marker densification, the credible interval with a 0.95 coverage probability was tighter: between 4 and 6.7 cM (Fig. 6b).

To further refine the QTL location, the Multiple QTL Method was applied on the PSII yield trait at W_4. This method has a higher power in detecting QTLs, is more precise in locating the QTLs and protects against over fitting (Arends et al. 2010). With this method, the QTL peak
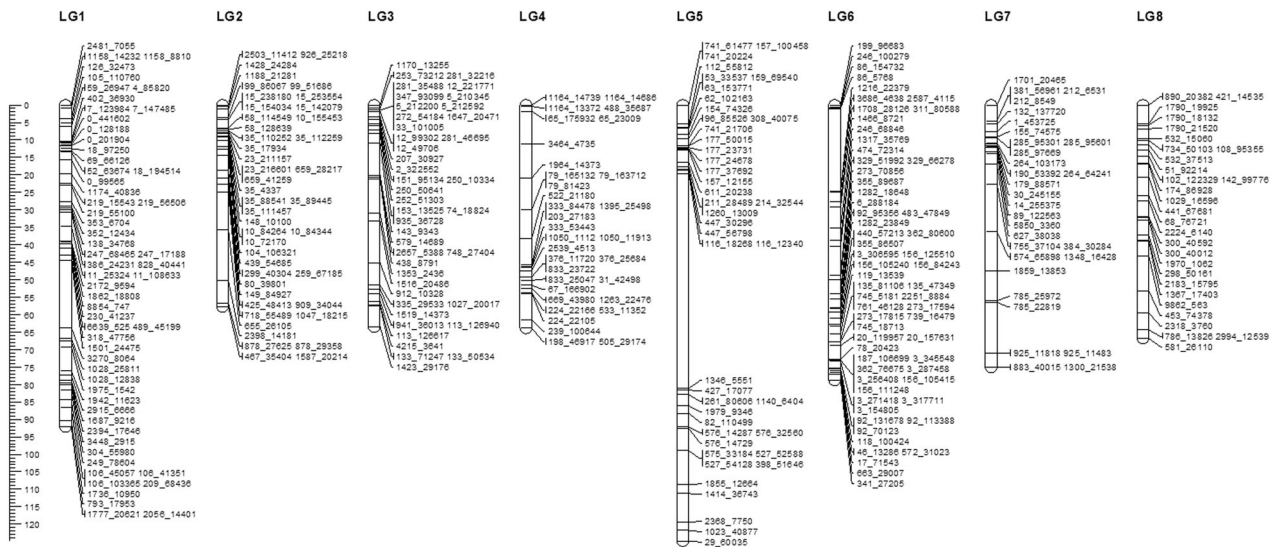
LG1 LG2 LG3 LG4 LG5 LG6 LG7 LG8

Fig. 5 Linkage map of the *A. halleri* F2 cross constructed with *R/qtl*. LG = linkage group. Distances on the genetic map are expressed in cM on the left scale. The 352 SNPs included in the map are represented by the scaffold number and the position on the scaffold separated by an underscore. The scaffolds belong to the *A. halleri* genome assembly described in this study
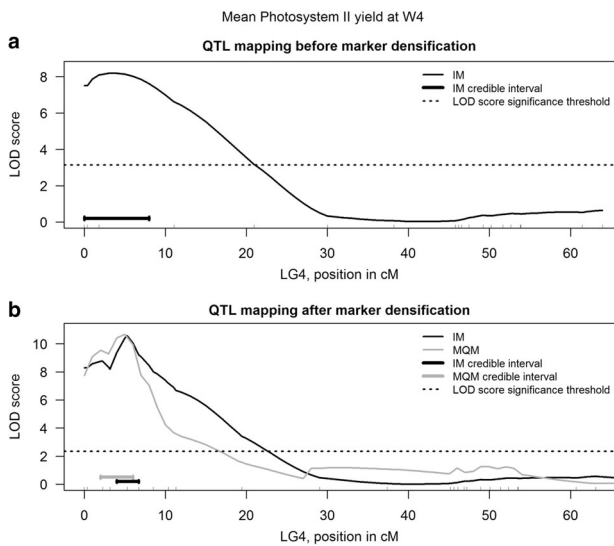
Mean Photosystem II yield at W4

a QTL mapping before marker densification

b QTL mapping after marker densification

Fig. 6 QTL mapping results for the tolerance indices of the photosystem II yield at W4, **a** before fine mapping and **b** after fine mapping. The marker positions of the linkage group 4 (LG4) are represented by the hatch marks along the *x* axes. The horizontal dashed lines represent the LOD score significance thresholds with $\alpha = 0.05$. The thick segments represent the credible intervals with a 0.95 coverage probability. For panel **a**, the LOD score significance threshold was estimated at 3.04 and the credible interval extended from 0 to 8 cM. In panel **b**, QTL mapping results were obtained by IM (in black) and multiple QTL mapping (MQM, in gray). The linkage group is truncated on its right side (between 63 and 115 cM, where there is no QTL signal). The LOD score significance threshold was estimated at 2.35 and the credible interval extended from 4 to 6.7 cM and from 2 to 6 with IM and MQM, respectively

and QTL credible interval were slightly shifted to the left (Fig. 6b). The maximum LOD score was similar (10.67 vs. 10.57 with Interval Mapping) and was located at a pseudo-marker position at 5 cM. The credible interval was located between 2 and 6 cM, vs. 4 and 6.7 cM with Interval Mapping. The markers in the credible interval of the QTL were "65_175932", "65_23009", "fm364_13413", "fm364_133 54", "fm1776_3057" (Fig. S6).

When examining the closest marker to the QTL peak (65_23009 before fine mapping and fm1776_3057 after fine mapping), the M allele increased the photosynthetic yield of 10.5% and 12.1% before and after maker densification at W4, respectively (Fig. 7). The M allele increased the photosynthetic yield of 12% at W_4 (Fig. 7). The additive effect (resp., dominance effect) of the QTL was estimated to 0.0463 ± 0.0072 (resp., 0.0458 ± 0.0104) before marker densification and 0.0542 ± 0.0072 (resp., 0.0475 ± 0.0099) after marker densification at W4. The degree of dominance of the QTL was estimated to 0.99 before marker densification and 0.88 after marker densification at W_4, so the QTL was partially dominant to dominant. This was reflected by the position of the phenotypic distribution of the heterozygotes that was closer to that of the M homozygotes, whatever the marker density (Fig. 7).

Gene annotation in the major QTL region

The fine-mapped *A. halleri* QTL credible interval (*Ahfmqtl*) included 34 genomic scaffolds (scaffold length comprised

Fig. 7 Effect plots of the QTL associated with the tolerance indices of the photosystem II yield at W4, before and after marker densification. The QTL mapping method used is interval mapping. The allele «A» is exclusively transmitted by the non-metallicolous parents (from I30) and the allele «B» by the metallicolous parents (from I35)

Mean Photosystem II yield at W4
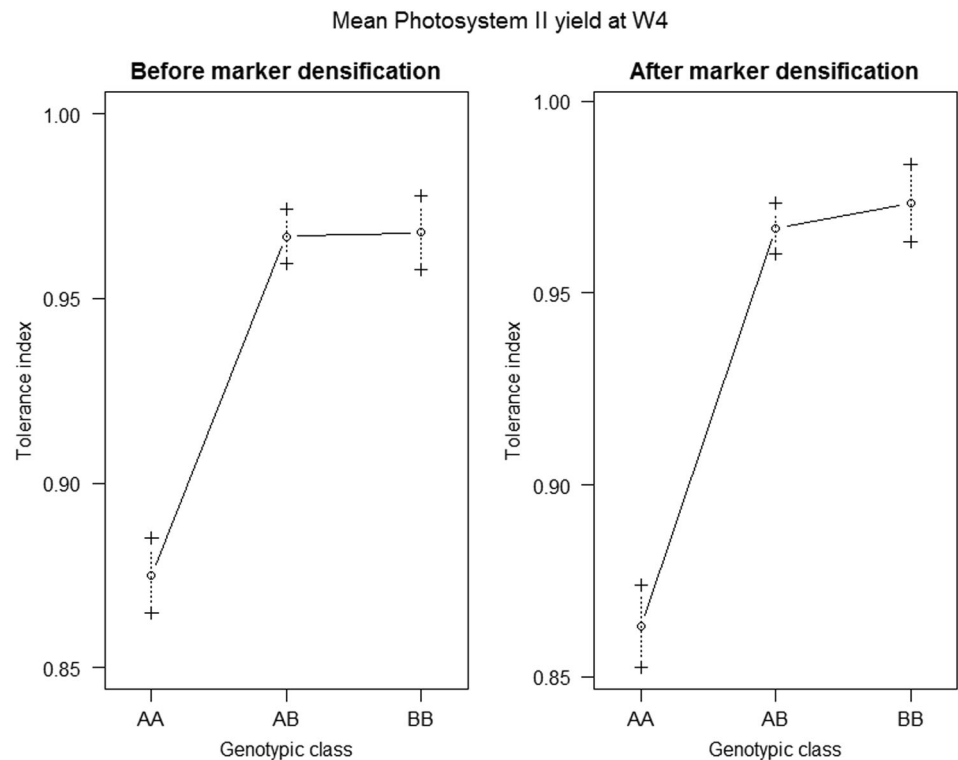
between 2132 and 203,886 bp, mean scaffold length = 22,087 bp, and scaffold total length = 993,899 bp) (Fig. S7). The homologous *A. lyrata* region (*Alfmqtl*) was 1,038,635 bp long and was located on the *A. lyrata* chromosome 4. *Alfmqtl* and *Ahfmqtl* shared 127 *A. lyrata* genes. Twenty-three *A. lyrata* genes were only found in *Alfmqtl* and 160 *A. lyrata* genes were only found in *Ahfmqtl* (Fig. S8). The homology-based functional annotation of the *A. lyrata* genes in the *Ahfmqtl* and *Alfmqtl* regions is represented in Supporting Information (Table S5).

Transcript levels of putative candidate genes

Among the genes present in the PSII yield QTL, three were metal homeostasis-related genes: *NRAMP3* (*Natural Resistance-Associated Macrophage Protein 3*), *MT4b* (*Metallothionein 4b*), and *HMA1* (*Heavy Metal ATPase 1*). Their relative expression levels were examined by quantitative RT-PCR in $F2_{I30}$ and $F2_{I35}$ progenies, respectively, representing homozygotes for the I30 and I35 alleles. Although it was not detected in the QTL interval, the *MTP1* expression level was also determined, considering its potential implication in intraspecific variation in *A. halleri* (Meyer et al. 2016). No significant difference in relative expression of *MTP1*, *HMA1*, and *MT4b* was detected between the two genotypic classes (Fig. 8b–d). In contrast, *NRAMP3* was on average 1.6-fold more highly expressed

($p < 0.05$) in the F2 individuals harboring an I35 allele in the QTL interval (Fig. 8a).

Discussion

The ability of phenotypic change in response to rapid environmental alterations is a major challenge conditioning the survival of species populations. Such phenotypic change may or may not involve genetic adaptation, yet the selection of genetic mechanisms may better promote long-term survival of populations. This study aimed at expanding our knowledge on the genetic architecture of a plant adaptive trait associated with anthropized habitats. For this purpose, a QTL mapping approach on an intraspecific F2 progeny from a cross between neighboring, although ecologically contrasted, populations of *A. halleri*, was used. This was challenging for several reasons. It required (1) that adaptive genetic changes actually occurred between the source populations, (2) appropriate phenotyping methods to reveal intraspecific variations in response to culture conditions, (3) a sufficient number of genetic markers, both scattered over the genome and polymorphic at the local geographical scale to allow genetic mapping and QTL analysis, and (4) sufficient knowledge about the gene content of detected QTL region to allow discussing potential candidate genes.
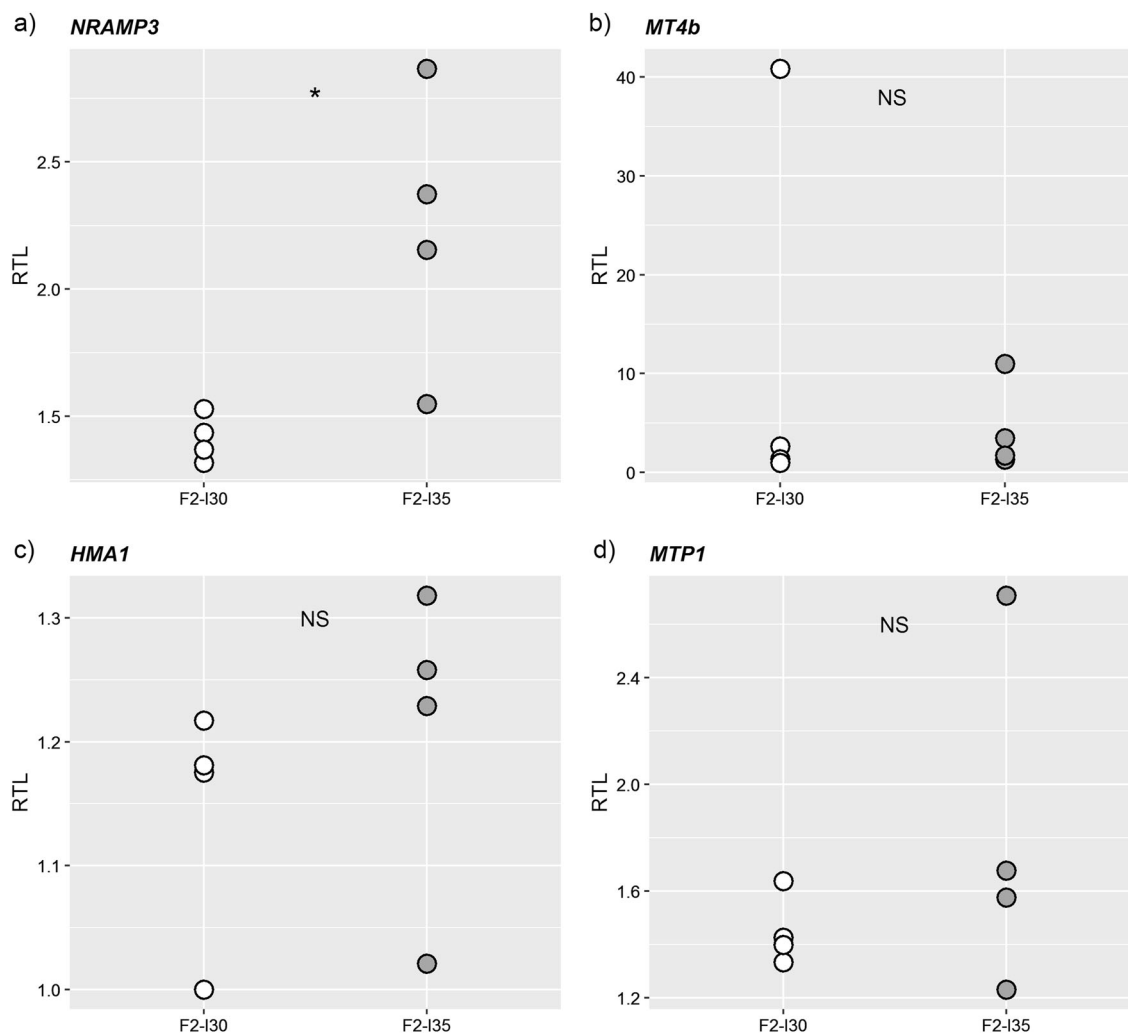
Fig. 8 Relative transcript levels (RTL) of metal homeostasis-related genes in leaf tissues of eight F2 progenies of the I30 × I35 cross collected from plants grown on soil. Four of the progenies are homozygous for the I30 allele (in white) and four for the I35 allele (in gray) at the QTL. *NRAMP3* (**a**), *MT4b* (**b**) and *HMA1* (**c**) genes were identified in the QTL region. *MTP1* (**d**) was added as a candidate gene in literature. A non-parametric Mann-Whitney test was applied to compare the genotypic classes with a significance level of 0.05. '*' $p < 0.05$; 'NS': non-significant

Generating a progeny segregating for Zn tolerance

To generate an intraspecific F2 progeny segregating for Zn tolerance, parents were selected based on previous knowledge about the environmental heterogeneity and structure of metal-related traits among a set of M and NM populations of *A. halleri* in Bergamo Alps, Italy (Frérot et al. 2018). The two isolated source populations (I30 (NM) and I35 (M)) were shown to colonize highly contrasted environments, in particular regarding soil metal pollution, and to display phenotypic divergence for some metal-related traits (Frérot et al. 2018). Here, Zn tolerance estimation was based on variation over time in several morphological and physiological traits, as previously performed on Polish and Slovakian *A. halleri* populations (Meyer et al. 2010). Heritability values suggested

strong genetic basis for the observed variation. Higher values observed in Zn-polluted (relatively to non-polluted) condition and for tolerance indices may indicate an increased genetic variance in response to Zn pollution, rather supporting the selective influence of Zn on tolerance traits. However, Zn pollution did not have the same phenotypic effect on all traits. Hence, Zn exposure caused a general decrease in the PSII yield of the F2 plants while, unexpectedly, morphological trait values increased with Zn exposure. Such increase in the biomass-related traits under Zn-polluted condition was also observed for metallicolous individuals from Germany grown in pots in an experimental garden (Kazemi-Dinan et al. 2015). Nevertheless, these findings differed from what was previously observed on Polish and Slovakian *A. halleri* populations (Meyer et al. 2010). This result

could be related to differences in genetic backgrounds, in particular Italian and Polish/Slovakian populations were assigned to distinct genetic units (Pauwels et al. 2012; Šrámková et al. 2017). Recent transcriptomic studies highlighted significant differences between M populations of these two genetic units, supporting the existence of specific mechanisms (Corso et al. 2018; Schvartzman et al. 2018).

Construction and validation of the *A. halleri* genetic map

Next Generation Sequencing technologies are offering a fast and low-cost tool for developing a large number of DNA markers that can be used in population and quantitative genetics (Metzker 2009; Nielsen et al. 2011). Here, genome sequencing of parental and F1 individuals allowed detecting millions of SNPs widely distributed across the genome. The high density of obtained SNPs allowed applying very stringent quality and mapping filters and selecting 352 SNPs covering both neutral and non-neutral loci. The large list of available SNPs also allowed swift selection of additional markers for the fine mapping.

SNPs were used to genotype the F2 progeny and build a robust *A. halleri* genetic map. As expected, it consisted in eight linkage groups corresponding to the eight *A. halleri* chromosomes. The map was larger in size (622.8 vs. 567 and 526 cM), more dense in markers (384 vs. 85 and 70 markers, 2.0 vs. 6.6 and 8.5 cM average spacing between two adjacent markers) than previously obtained (Frérot et al. 2010; Willems et al. 2010). Ideally, QTL mapping requires homogeneous distribution of molecular markers on the linkage groups. However, a non-random marker distribution was observed in the map, reflected by large gaps between marker groups. This could reflect bias of recombination rates along the genome, or even bias in the SNP selection procedure preventing detection of SNPs in highly repeated or low complexity genomic regions. The localization of some gaps was consistent with previous *A. halleri* maps, especially on LG1 and LG6 (Frérot et al. 2010; Willems et al. 2010), rather suggesting bias in recombination rates in these regions.

Detection of a major QTL for Zn tolerance

From the set of phenotypic traits, a single QTL was detected, associated with PSII yield only. The absence of QTLs for the other traits could have multiple reasons. The genetic bases of biomass traits are possibly highly complex (i.e. integrative traits) and could be governed by multiple small effect QTLs that would require a higher detection power than what was achieved here. Additionally, these traits may not be appropriate to detect quantitative variation for Zn tolerance at the intraspecific level, as suggested by the absence of significant correlation with PSII yield. The QTL for PSII yield was detected at both W_2 and W_4, from both raw phenotype data in the P condition (data not shown) and TIs. Interestingly, the QTL signal got stronger with time, confirming the robustness of the association. It appeared as a major tolerance QTL explaining up to 27% of the total phenotypic variance. With an estimated broad-sense heritability of 37%, this QTL would explain 73% approximately of the trait genetic variance. In comparison, Zn tolerance QTLs previously detected in *A. halleri* × *A. lyrata* progenies explained about 42% of the genetic variance (Meyer et al. 2016; Willems et al. 2007).

Interestingly, this study corroborates previous results that investigated the population structure of various phenotypes in Polish and Slovakian populations of *A. halleri* (Meyer et al. 2010). The latter study indeed supported an adaptive phenotypic divergence among M and NM populations for PSII yield in contrast to other traits (i.e. shoot biomass, root biomass, root length, and leaf width).

QTL specificity at the intraspecific scale

Until recently, the genetic architecture of Zn tolerance in *A. halleri* was studied at the interspecific level, supposedly more appropriate to observe segregation for metal-related traits in the progeny. QTL analysis of Zn tolerance involved either an M (Willems et al. 2007) or an NM (Meyer et al. 2016) accession of *A. halleri* crossed with *A. lyrata*. However, these studies were not relevant to uncover the genetic mechanisms involved in the adaptive divergence between M and NM populations of *A. halleri*. They only provided indirect information. Indeed, the comparison between results at the interspecific level suggested that the *AhMTP1-A* and *AhMTP1-B* genes, encoding a Zn^{2+}/H^+ antiporter responsible for Zn cytoplasmic detoxification (Fasani et al. 2017; Shahzad et al. 2010), were covered by QTL regions only detected using an M accession (from Auby, France). This indirectly suggested that they could be implicated in increased Zn tolerance observed in M populations (Meyer et al. 2016). A high expression of *MTP1* was indeed observed in unrelated Italian and Polish M populations (Schvartzman et al. 2018). In our intraspecific level study, there was no QTL co-localizing with any *MTP1* gene copy, and there was no significant differences in *MTP1* expression in F2$_{I35}$ compared to F2$_{I30}$ plants (Fig. 8). Our results therefore suggest that *AhMTP1* genes may be not specifically involved in adaptive mechanisms related to anthropogenic metal pollution in the investigated region. Such contradictory outcomes may result from the fact that

accessions used in different studies can be geographically highly distant, and likely belong to different genetic units that might have experience contrasted evolutionary histories (Pauwels et al. 2012; Šrámková et al. 2017), so that genetic resources locally implied may differ (Pauwels et al. 2008). In conclusion, the QTL detected here from an intraspecific cross did not co-localize with any other QTL regions identified so far, suggesting that specific candidate genes could be uncovered.

Molecular mechanisms of adaptation to metal pollution

Among the metal homeostasis genes identified in the Zn tolerance QTL interval, no significant difference in relative expression of *HMA1* and *MT4b* was detected in F2 progenies (Fig. 8b, c), questioning their potential involvement in enhanced Zn tolerance of I35.

In contrast, *NRAMP3* may directly contribute to increased Zn tolerance in I35 compared to I30. Indeed, NRAMP3 belongs to a family of divalent metal cation transporters. *NRAMP3* is part of a small set of metal homeostasis genes that are constitutively more highly expressed in both *A. halleri* and *N. caerulescens* compared to *A. thaliana* (Krämer et al. 2007; Oomen et al. 2009; Talke et al. 2006; van de Mortel et al. 2006), suggesting that it may play a role in Zn hyperaccumulation and/or tolerance. In *A. thaliana*, NRAMP3 and NRAMP4 are both transporting iron (Fe) and manganese (Mn) out of the vacuole (Lanquar et al. 2005, 2010; Thomine et al. 2003, 2000). Both proteins play a key role in Fe remobilization during seedling germination and in Mn supply to PSII in adult leaves. Moreover, NRAMP3 and NRAMP4 also contribute to excess Zn and Cd tolerance by mediating appropriate supply of Fe and Mn to chloroplasts from vacuolar stores, thus maintaining the photosynthetic function (Molins et al. 2013; Oomen et al. 2009). Constitutive higher expression of *NRAMP3* in F2$_{I35}$ individuals (Fig. 8a) may thus support increased Zn tolerance through higher PSII yield efficiency upon Zn excess. Altogether, although additional genes present in the QTL interval may contribute to Zn tolerance, *NRAMP3* may be a good candidate to at least partly explain the QTL effect.

These observations suggest that fine-tuning of mechanisms that contribute to the species-wide hyperaccumulation and tolerance traits, i.e. the high expression of *NRAMP3* here, played a role in adaptation to a metal-contaminated site. Transcriptomic comparisons of Italian and Polish metallicolous *A. halleri* populations recently suggested that other mechanisms additional to the species-wide features also contributed to adaptive traits associated with polluted soils (Corso et al. 2018; Schvartzman et al. 2018).

Conclusion

Metal tolerance is a typical example of adaptive response of plants to anthropogenic metal pollution. The genetic architecture of this trait, probably contributing to adaptive divergence among populations occurring in contrasted habitats, was investigated here using high-throughput molecular technologies at the intraspecific level for the first time. Our approach proved to be successful to detect QTL regions and discuss the genetic resources that may allow populations to survive in highly anthropized habitats. It should help to inspire other studies using alternative models to better understand how a species can adapt to face new selective pressures of anthropogenic origin.

Data Archiving

The reads and genome assembly have been deposited at the NCBI Database with BioProject identification number PRJNA422908 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA422908). The reads of parental and F1 individuals have been deposited at the NCBI Database with BioProject identification number PRJNA427096 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA427096).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Alberti M, Correa C, Marzluff JM, Hendry AP, Palkovacs EP, Gotanda KM et al. (2017) Global urban signatures of phenotypic change in animal and plant populations. Proc Natl Acad Sci USA 114:8951–8956

Antonovics J, Bradshaw AD, Turner RG (1971) Heavy metal tolerance in plants. Adv Ecol Res 7:1–85

Arends D, Prins P, Jansen RC, Broman KW (2010) R/qtl: high-throughput multiple QTL mapping. Bioinformatics 26:2990–2992

Bechsgaard JS, Castric V, Charlesworth D, Vekemans X, Schierup MH (2006) The transition to self-compatibility in *Arabidopsis thaliana* and evolution within S-haplotypes over 10 Myr. Mol Biol Evol 23(9):1741–1750

Bergelson J, Roux F (2010) Towards identifying genes underlying ecologically relevant traits in *Arabidopsis thaliana*. Nat Rev Genet 11:867–879

Bert V (2000) Tolérance aux métaux lourds et accumulation chez Arabidopsis halleri (Brassicaceae). Université de Lille 1, Lille

Briskine RV, Paape T, Shimizu-Inatsugi R, Nishiyama T, Akama S, Sese J et al. (2016) Genome assembly and annotation of *Arabidopsis halleri*, a model for heavy metal hyperaccumulation and evolutionary ecology. Mol Ecol Resour 17(5):1025–1036

Broman KW, Sen S (2009). A guide to QTL mapping with R/qtl. Springer, New York.

Broman KW, Wu H, Sen Å, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. Bioinformatics 19(7):889–890

Charlesworth D, Barton NH, Charlesworth B (2017) The sources of adaptive variation. Proc R Soc B Biol Sci 284:20162864

Cheptou PO, Carrue O, Rouifed S, Cantarel A (2008) Rapid evolution of seed dispersal in an urban environment in the weed Crepis sancta. Proc Natl Acad Sci USA 105:3796–3799

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963–971

Clapham AR, Akeroyd JR (1993) Cardaminopsis. In: Tutin TG, Heywood VH, Burges NA, Valentine DH, Walters SM, Webb CT (eds) Flora Europaea, Cambridge University Press, Cambridge, vol. I, p 290.

Corso M, Schvartzman MS, Guzzo F, Souard F, Malkowski E, Hanikenne M et al. (2018) Contrasting cadmium resistance strategies in two metallicolous populations of *Arabidopsis halleri*. New Phytol 218(1):283–297

Fasani E, DalCorso G, Varotto C, Li M, Visioli G, Mattarozzi M et al. (2017) The MTP1 promoters from *Arabidopsis halleri* reveal cis-regulating elements for the evolution of metal tolerance. New Phytol 214:1614–1630

Franks SJ (2011) Plasticity and evolution in drought avoidance and escape in the annual plant *Brassica rapa*. New Phytol 190:249–257

Frérot H, Faucon M-P, Willems G, Godé C, Courseaux A, Darracq A et al. (2010) Genetic architecture of zinc hyperaccumulation in *Arabidopsis halleri*: the essential role of QTL x environment interactions. New Phytol 187(2):355–367

Frérot H, Hautekèete N-C, Decombeix I, Bouchet M-H, Créach A, Saumitou-Laprade P et al. (2018) Habitat heterogeneity in the pseudometallophyte *Arabidopsis halleri* and its structuring effect on natural variation of zinc and cadmium hyperaccumulation. Plant Soil 423(1–2):157–174

Genty B, Briantais J-M, Baker NR (1989) The relationship between the quantum yield of photosynthetic electron transport and quenching of chlorophyll fluorescence. Biochim Biophys Acta 990(1):87–92

Godé C, Decombeix I, Kostecka A, Wasowicz P, Pauwels M, Courseaux A et al (2012) Nuclear microsatellite loci for Arabidopsis halleri (Brassicaceae), a model species to study plant adaptation to heavy metals Am J Bot 99:e49–e52

Hohmann N, Schmickl R, Chiang T-Y, Lučanová M, Kolář F, Marhold K et al. (2014) Taming the wild: resolving the gene pools of non-model Arabidopsis lineages. BMC Evol Biol 14(1):224

Hohmann N, Wolf EM, Lysak MA, Koch M (2015) A time-calibrated road map of Brassicaceae species radiation and evolutionary history. Plant Cell 27:2770–2784

Johnson MTJ, Munshi-South J (2017) Evolution of life in urban environments Science 358:eaam8327

Kazemi-Dinan A, Barwinski A, Stein RJ, Krämer U, Müller C (2015) Metal hyperaccumulation in Brassicaceae mediates defense against herbivores in the field and improves growth. Èntomol Exp Appl 157(1):3–10

Kosambi D (1944) The estimation of map distances from the recombination values. Ann Eugen 12:172–175

Krämer U, Talke IN, Hanikenne M (2007) Transition metal transport. FEBS Lett 581:2263–2272

Lanquar V, Lelièvre F, Bolte S, Hamès C, Alcon C, Neumann D et al. (2005) Mobilization of vacuolar iron by AtNRAMP3 and AtNRAMP4 is essential for seed germination on low iron. EMBO J 24(23):4041–4051

Lanquar V, Ramos MS, Lelièvre F, Barbier-Brygoo H, Krieger-Liszkay A, Krämer U et al. (2010) Export of vacuolar manganese by AtNRAMP3 and AtNRAMP4 is required for optimal photosynthesis and growth under manganese deficiency. Plant Physiol 152:1986–1999

Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer Associates, Inc, Sunderland, MA

Macnair MR, Baker AJM (1994). Metal-tolerant plants: an evolutionary perspective. In: Farago ME (ed) Plant and the chemical elements. Biochemistry, uptake, tolerance andtoxicity. VCH, New York p 68–83.

McDonnell MJ, Hahs AK (2015) Adaptation and adaptedness of organisms to urban environments. Annu Rev Ecol, Evol Syst 46:261–280

Metzker ML (2009) Sequencing in real time. Nat Biotechnol 27:150

Meyer C-L, Juraniec M, Huguet Sp, Chaves-Rodriguez E, Salis P, Isaure M-P et al. (2015) Intraspecific variability of cadmium tolerance and accumulation, and cadmium-induced cell wall modifications in the metal hyperaccumulator *Arabidopsis halleri*. J Exp Bot 66(11):3215–3227

Meyer C-L, Kostecka AA, Saumitou-Laprade P, Créach A, Castric V, Pauwels M et al. (2010) Variability of zinc tolerance among and within populations of the pseudometallophyte *Arabidopsis halleri* and possible role of directional selection. New Phytol 185 (1):130–142

Meyer C-L, Pauwels M, Briset L, Godé C, Salis P, Bourceaux A et al. (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 Phytol 212(4):934–943

Molins H, Michelet L, Lanquar V, Agorio A, Giraudat J, Roach T et al. (2013) Mutants impaired in vacuolar metal mobilization identify chloroplasts as a target for cadmium hypersensitivity in *Arabidopsis thaliana*. Plant Cell Environ 36(4):804–817

Nielsen R, Paul JS, Albrechtsen A, Song YS (2011) Genotype and SNP calling from next-generation sequencing data. Nat Rev Genet 12:443

Novikova PY, Hohmann N, Nizhynska V, Tsuchimatsu T, Ali J, Muir G et al. (2016) Sequencing of the genus Arabidopsis identifies a complex history of nonbifurcating speciation and abundant trans-specific polymorphism. Nat Genet 48(9):1077–1082

O'Kane SLJ, Al-Shehbaz IA (1997) A synopsis of *Arabidopsis* (Brassicaceae). Novon 7(3):323–327

Oomen RJ, Wu J, Lelièvre F, Blanchet S, Richaud P, Barbier-Brygoo H et al. (2009) Functional characterization of NRAMP3 and NRAMP4 from the metal hyperaccumulator *Thlaspi caerulescens*. New Phytol 181:637–650

Pauwels M, Frérot H, Bonnin I, Saumitou-Laprade P (2006) A broad-scale study of population differentiation for Zn-tolerance in an emerging model species for tolerance study: *Arabidopsis halleri* (Brassicaceae). J Evol Biol 19(6):1838–1850

Pauwels M, Roosens N, Frérot H, Saumitou-Laprade P (2008) When population genetics serves genomics: putting adaptation back in a spatial and historical context. Curr Opin Plant Biol 11 (2):129–134

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 Phytol 193(4):916–928

Reznick DN, Ghalambor CK (2001) The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. Genetica 112-113:183–198

Roosens N, Willems G, Godé C, Courseaux A, Saumitou-Laprade P (2008) The use of comparative genome analysis and syntenic relationships allows extrapolating the position of Zn tolerance QTL regions from *Arabidopsis halleri* into *Arabidopsis thaliana*. Plant Soil 306:105–116

Savolainen O, Lascoux M, Merilä J (2013) Ecological genomics of local adaptation. Nat Rev Genet 14:807–820

Schoener TW (2011) The newest synthesis: understanding the interplay of evolutionary and ecological dynamics. Science 331:426–429

Schvartzman MS, Corso M, Fataftah N, Scheepers M, Nouet C, Bosman B et al. (2018) Adaptation to high zinc depends on distinct mechanisms in metallicolous populations of *Arabidopsis halleri*. New Phytol 218(1):269–282

Shahzad Z, Gosti F, Frérot H, Lacombe E, Roosens N, Saumitou-Laprade P et al. (2010) The five AhMTP1 zinc transporters undergo different evolutionary fates towards adaptive evolution to zinc tolerance in *Arabidopsis halleri*. PLoS Genet 6(4): e1000911

Šrámková G, Záveská E, Kolář F, Lucanova M, Španiel S, Marhold K (2017) Range-wide genetic structure of *Arabidopsis halleri* (Brassicaceae): glacial persistence in multiple refugia and origin of the Northern Hemisphere disjunction. Bot J Linn Soc 185:321–342

Steffen W, Sanderson RA, Tyson PD, Jäger J, Matson PA, Moore III B et al. (2006) Global change and the earth system: a planet under pressure. Springer-Verlag Berlin Heidelberg New York.

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

Thomine S, Lelièvre F, Debarbieux E, Schroeder JI, Barbier-Brygoo H (2003) AtNRAMP3, a multispecific vacuolar metal transporter involved in plant responses to iron deficiency. Plant J 34 (5):685–695

Thomine Sb, Wang R, Ward JM, Crawford NM, Schroeder JI (2000) Cadmium and iron transport by members of a plant metal transporter family in Arabidopsis with homology to Nramp genes. Proc Natl Acad Sci 97(9):4991–4996

Thompson KA, Renaudin M, Johnson MTJ (2016) Urbanization drives the evolution of parallel clines in plant populations. Proc R Soc B Biol Sci 283:20162180

van de Mortel JE, Villanueva LA, Schat H, Kwekkeboom J, Coughlan S, Moerland PD et al. (2006) Large expression differences in genes for iron and zinc homeostasis, stress response, and lignin biosynthesis distinguish roots of *Arabidopsis thaliana* and the related hyperaccumulator *Thlaspi caerulescens*. Plant Physiol 142:1127–1147

Willems G, Dräger D, Courbot M, Godé 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* x *Arabidopsis lyrata petraea* F2 progeny grown on cadmium-contaminated soil. New Phytol 187(2):368–379