Plant Soil (2008) 306:105–116
DOI 10.1007/s11104-007-9383-5

REGULAR ARTICLE

The use of comparative genome analysis and syntenic relationships allows extrapolating the position of Zn tolerance QTL regions from *Arabidopsis halleri* into *Arabidopsis thaliana*

Nancy H. C. J. Roosens · Glenda Willems · Cécile Godé · Adeline Courseaux · Pierre Saumitou-Laprade

Received: 6 July 2007 / Accepted: 6 August 2007 / Published online: 14 September 2007

Abstract *Arabidopsis halleri* is a species that has undergone natural selection for zinc (Zn) tolerance. Isolation of the quantitative trait loci (QTL) associated with this trait holds great promise for the identification of the main genes responsible for this adaptation. Using a segregating progeny produced by an interspecific cross, we previously constructed a genetic linkage map of *A. halleri* × *A. lyrata petraea* and mapped the three main QTL that confer Zn tolerance in *A. halleri* (Willems et al.). The goal of the present study is to compare the genetic linkage map of *A. halleri* × *A. l. petraea* to the annotated *A. thaliana* genome sequence to generate a tool for *A. halleri* genomic approaches. To achieve this aim, we constructed a genetic linkage map with 81 markers anchored on *A. thaliana*, including 23 genes known to be involved in metal homeostasis. First, this provided an extensive overview of the chromosomal rearrangements that have occurred since the divergence between *A. thaliana* and its closest relative *A. halleri*. Second, on the basis of the syntenic relationships assessed experimentally through this work, we transferred the QTL confidence intervals for Zn tolerance to the *A. thaliana* physical map, allowing access to all the genes localized in the corresponding regions. Third, we validated from the 23 genes involved in metal homeostasis the three ones localized in the QTL regions that can be considered the best candidates for conferring Zn tolerance.

Keywords *Arabidopsis halleri* · Zinc tolerance · QTL mapping · Synteny

Nancy H. C. J. Roosens and Glenda Willems contributed equally to this paper.

Responsible Editor: Ismail Cakmak.

N. H. C. J. Roosens · G. Willems · C. Godé · A. Courseaux · P. Saumitou-Laprade (✉)
Laboratoire de Génétique et Evolution des Populations Végétales, UMR CNRS 8016, Université des Sciences et Technologies de Lille, Lille 1, Bâtiment SN2, 59655 Villeneuve d'Ascq Cedex, France
e-mail: pierre.saumitou@univ-lille1.fr

Introduction

Interest in hyperaccumulating plant species has greatly increased this last decade because of their potential use in phytoremediation, that is, the decontamination of soils with high metal concentrations (Assunçao et al. 2003). These plants are not only able to survive on extreme substrates like mine wastes and smelter sites, but also accumulate exceptionally high concentrations of heavy metals like zinc (Zn), cadmium (Cd) or nickel (Ni) in their above ground biomass (Reeves et al. 2001). The mechanisms that allow plant hyperaccumulators to tolerate very high metal concentrations are not fully understood, although they are believed to have evolved within the

metal homeostasis network (Clemens et al. 2002). To investigate the physiological and molecular bases of metal tolerance in hyperaccumulating plants, several plant models have been described, particularly in the family of the Brassicaceae (Assunçao et al. 2003; Becher et al. 2004; Peer et al. 2003; Weber et al. 2004). Among them, *Arabidopsis halleri* is one of the most promising because it is the closest metal tolerant and hyperaccumulating relative of the biological model species *A. thaliana*, from which it diverged 5 Ma ago (Bechsgaard et al. 2006). Although the species differ in chromosome number, they share 94% identity at the nucleotide level within coding regions (Becher et al. 2004). Recently, this close relationship allowed the use of *A. thaliana* GeneChips to compare gene expression levels between *A. halleri* and the non-tolerant *A. thaliana*, and, consequently, permitted the identification of genes potentially involved in metal tolerance and/or hyperaccumulation (Becher et al. 2004; Chiang et al. 2006; Filatov et al. 2006; Weber et al. 2004; Weber et al. 2006).

Recent studies on a representative range of *A. halleri* populations clearly established the constitutive nature of the high tolerance to Zn in this species (Pauwels et al. 2006). Because of this constitutive nature, the genetic bases of metal tolerance were investigated through interspecific crosses between *A. halleri* and its non-tolerant relative *A. l. petraea* (Macnair et al. 1999). This work concluded that there is a small number of genes governing Zn tolerance (Macnair et al. 1999). Very recently, a first generation backcross (BC1) progeny of an *A. halleri* × *A. l. petraea* (*Ah* × *Alp*) cross was produced for deciphering the genetic architecture of Zn tolerance (Willems et al. 2007). Three quantitative trait loci (QTL) conferring Zn tolerance were identified and their respective contributions to the trait established (Willems et al. 2007). Because of their size, each QTL region that was identified in *A. halleri,* may contain hundreds to thousands of genes; the next challenge is to identify from among them the best candidates for genes involved in Zn tolerance. However, the limited access to genomic data on wild species, such as *A. halleri*, hinders this challenge. In this context, comparative genome analyses, which provide the opportunity to study genome colinearity at macro- and microsynteny levels, will be very useful.

The complete annotation of the *A. thaliana* genome sequence (The Arabidopsis Genome Initiative 2000) provides a solid foundation for comparative mapping studies within the Brassicaceae family, and the genome organization of *A. thaliana* has already been compared with those of several species (Koch and Kiefer 2005) like *A. lyrata lyrata* (Yogeeswaran et al. 2005), *A. lyrata petraea* (Kuittinen et al. 2004) and *Capsella rubella* (Boivin et al. 2004). A similar comparison remains to be done in the metallicolous species *A. halleri*.

To this end, the 61 previously genotyped markers anchored on *A. thaliana* (Willems et al. 2007) have been complemented with 20 new ones. An updated *Ah* × *Alp* genetic linkage map was generated and the interval of the QTLs for Zn tolerance recalculated. The genome colinearity was assessed empirically. The three QTL confidence intervals for Zn tolerance were transferred to the *A. thaliana* physical map, and the microsynteny level was assessed. The effective position of 23 genes involved in metal homeostasis was verified; among them, 3 genes were localized in the QTL regions.

Materials and methods

Plant material and DNA extraction

Genomic DNA of the four parental genotypes and the 336 progeny segregating for Zn tolerance from the BC1 of *A. halleri* × *A. l. petraea* described in Willems et al. (2007) was extracted for marker analysis using a slightly modified Dellaporta method (Saumitou-Laprade et al. 1999).

Genetic marker analysis

To increase the density of the *Ah* × *Alp* genetic map, we designed primers homologous to the *A. thaliana* genomic sequence using the public *Arabidopis* databank (TAIR database at http://www.arabidopsis.org/). The newly defined markers targeted genes among functional groups related to metal homeostasis and/or in accordance with their position in *A. thaliana* (Table 1). Most of the primers amplifying these newly defined markers were designed in exons (maximizing cross-species amplification of the markers), to amplify fragments displaying a size between 600 and 850 base pairs (bp) and to scan for polymorphism between species, within at least one intron. The sequences of

Table 1 Markers anchored in *A. thaliana* used in the linkage map construction and for the synteny establishment

Localization in *A. thaliana* (kb)	Locus name	Identification in *A. thaliana*	Type of polymorphism	Forward primer	Reverse primer	PCR (°C; s)
Linkage group LG1						
Chr.1-00240	AthACS	At1g01660	Willems et al. 2007			
Chr.1-01019	Ah–PCS2	At1g03980	*Eco*RI	CGATCTCAACATATCATAGGCAA	CCTTCAAATACTTTGCAATGCTG	50; 60
Chr.1-01498	AXR1	At1g05180	Willems et al. 2007			
Chr.1-02985	Ah–NAS1	At1g09240	*Hae*III	CTTCGATCTTGACCTCTTACC	ACCAACATTCCCAATTGAAAC	50; 75
Chr.1-03095	PhyA	At1g09570	Willems et al. 2007			
Chr.1-03665	Ah–ZNT1	At1g10970	*Ind*eI–LiCor	TTCGTGCTCATGCAGCTCAC[a]	GACACAATCCCAAGCTCC[a]	50; 60
Chr.1-04266	ICE10	F12F1	Willems et al. 2007			
Chr.1-04516	ICE13	At1g13220	Willems et al. 2007			
Chr.1-05639	Tc–up3	At1g16500	*Ind*eI–LiCor	GAGGAGATCGCGAGTCATGAG[a]	CTGCCTAACGTACCGCATAACTG[a]	50; 60
Chr.1-08062	GI	At1g22770	Willems et al. 2007			
Chr.1-10858	ATTS0392	At1g30630	Willems et al. 2007			
Chr.1-16487	VIP1	At1g43700	Willems et al. 2007			
Chr.1-17268	At1g46768	At1g46768	Willems et al. 2007			
Linkage group LG2						
Chr.1-22940	F19K23	At1g62050	Willems et al. 2007			
Chr.1-21141	Ah–NAS2	At1g56430	*Rsa*I	CTTCAATCTTAACCTCTTTTG	ATGAACAACACCATTAGCATC	50; 60
Chr.1-24870	SLL2	At1g66680	Willems et al. 2007			
Chr.1-25548	Lyr132	At1g68150	Willems et al. 2007			
Chr.1-26613	Nga111	F28P22	Willems et al. 2007			
Chr.1-28476	At1g75830	At1g75830	Willems et al. 2007			
Chr.1-28980	ADH1	At1g77120	Willems et al. 2007			
Chr.1-30378	Ah–NRAMP1	At1g80830	*Ind*eI, LiCor	GTCTCAATTCATTTCCAGTTCC	GATCCACACCGGTATGCTAAAG	50; 60
Linkage group LG3						
Chr.3-00303	Tc–up2	At3g01860	*Ind*eI–LiCor	TGAGAAGAGGAGACACAGGAA[a]	CACTTACCAAATCGAAAACTGCT[a]	50; 60
Chr.3-01936	Ah–FRD3	At3g08040	Willems et al. 2007			
Chr.3-02889	Ah–MT2a	At3g09390	*Eco*RI[b]	Zhou and Goldsbrough 1994		
Chr.3-04052	Ah–ZIP1	At3g12750	*Eco*RI[b]	Grotz et al. 1998		
Chr.3-04577	MDC16	MDC16	Willems et al. 2007			
Chr.3-08098	DMC1	At3g22880	Willems et al. 2007			
Chr.3-08806	At3g24300	At3g24300	*Alu*I	CTCTTTGGTTACGCCTTTGC	AAACGTCCGTTACGTTCCAG	50; 60
Chr.2-04317	At2g10940	At2g10940	Courbot et al. 2007			
Chr.2-06252	At2g14620	At2g14620	Courbot et al. 2007			
Chr.2-07228	Ah–CCH	At2g18196	Courbot et al. 2007			
Chr.2-08286	Ah–HMA4	At2g19110	Willems et al. 2007			
Linkage group LG4						
Chr.2-09133	Con	At2g21320	Willems et al. 2007			
Chr.2-09533	At2g22430	At2g22430	Willems et al. 2007			

Table 1 (continued)

Localization in *A. thaliana* (kb)	Locus name	Identification in *A. thaliana*	Type of polymorphism	Forward primer	Reverse primer	PCR (°C; s)
Chr.2-09834	CK2alpha2	At2g23080	Willems et al. 2007			
Chr.2-09863	Ah–NRAMP3	At2g23150	Willems et al. 2007			
Chr.2-11066	ELF3	At2g25930	Willems et al. 2007			
Chr.2-12324	At2g28700	At2g28700	Willems et al. 2007			
Chr.2-12846	Ah–ZIP6	At2g30080	*Ind*eI–LiCor	GACGGAGGGATGGAGTATATGC	GCACATATACACCACTGTCCCT	50; 60
Chr.2–14018	At2g33010	At2g33010	Willems et al. 2007			
Chr.2–15786	ACT3	At2g37620	Willems et al. 2007			
Chr.2–15997	Ah–CAX1	At2g38170	*Eco*RV	TGACAGAGCCGTGGTCAGTAGC	AGCTAAGTCCAAATATCCAAGG	50; 60
Chr.2-16299	ICE12	At2g39010	Willems et al. 2007			
Chr.2-16779	At2g40140	At2g40140	Willems et al. 2007			
Chr.2-17894	At2g43010	At2g43010	Willems et al. 2007			
Chr.2-19112	ICE11	At2g46530	Willems et al. 2007			
Chr.2-19245	Ah–MTP1A	At2g46800	Willems et al. 2007			
Chr.2-19598	Ah–MTPc1	At2g47830	*Sph*I	GACTGAAAACAGGACACCAA	TGTCTGAGACAGTGGCTTCA	50; 60
Chr.2-19531	Ah–MHX1	At2g47600	*Eco*RI	GACTGAAAACAGGACACCAA	TGTCTGAGACAGTGGCTTCA	50; 60
Linkage group LG5						
Chr.2-00684	Nga1145	At2g02540	Willems et al. 2007			
Chr.3-10526	At3g28220	At3g28220	Willems et al. 2007			
Chr.3-14096	At3g33530	At3g33530	Willems et al. 2007			
Chr.3-16844	At3g45810	At3g45810	Willems et al. 2007			
Chr.3-19036	F3H	At3g51240	Willems et al. 2007			
Chr.3-21509	Ah–MTPc3	At3g58060	Willems et al. 2007			
Chr.3-23179	Nga112	At3g62650	Willems et al. 2007			
Linkage group LG6						
Chr.4-04240	At4GA2	T18A10	Willems et al. 2007			
Chr.4-01123	LD	At4g02560	Willems et al. 2007			
Chr.4-06346	AthDET1	At4g10180	Willems et al. 2007			
Chr.5-08819	ICE2	K18P6	Willems et al. 2007			
Chr.5-08352	AthCDPK9	MQM1	Willems et al. 2007			
Chr.5-04824	Ah–HMADP2	At5g14910	*Ind*eI–LiCor	CGCAATTGCAGTTTTGCTGGGT	CCTGAACCTTTAAATTGGAAAC	50; 40
Chr.5-02626	At5g08160	At5g08160	Willems et al. 2007			
Ch .5-00980	ATHCTR1	CTR1.2	*Ind*eI–LiCor	TCTATCAACAGAAACGCACCGAG	CCACTTGTTTCTCTCTCTAG	55; 35
Chr.5-00506	Ah–MT2b	At5g02380	Willems et al. 2007			
Linkage group LG7						
Chr.4-17925	At4g38220	At4g38220	Willems et al. 2007			
Chr.4-17540	Ah–HMA1	At4g37270	*Ind*eI–LiCor	GTCTACGTGCTGTCGAAGATCA	TGCAGAATCTGGATTACTCTCT	50; 60
Chr.4-15994	At4g33160	At4g33160	Willems et al. 2007			
Chr.4-13587	TSB2	At4g27070	Willems et al. 2007			
Chr.4-11384	SRK	At4g21366	Willems et al. 2007			

Chr.4-10703	Ah–IRT2	At4g19680	*Ind*eI–LiCor	CATGGTGTTGTGCTAGCAAC[a]	GAGATAGTCCAATGACCACAG[a]	50; 60
Chr.4-09449	HAT4	At4g16780	*Nco*I	CGTGAACAGACCACCCGTC	AGCGTCAAAAGTCAAGCCGT	65; 60
Chr.4-09207	FCA	At4g16280	Willems et al. 2007			
Chr.5-14025	PhyC	At5g35840	Willems et al. 2007			
Chr.5-15022	ATTSO191	At5g37780	Willems et al. 2007			
Chr.5-16149	ICE9	At5g40340	Willems et al. 2007			
Linkage group LG8						
Chr.5-17648	ATCLH2	At5g43860	Willems et al. 2007			
Chr.5-20841	EMF2	At5g51230	Willems et al. 2007			
Chr.5-24008	Ah–ZIP2	At5g59520	*Eco*RV[b]	Grotz et al. 1998		
Chr.5-24702	MHJ24		Willems et al. 2007			
Chr.5-26879	Ah–NRAMP4	At5g67330	Willems et al. 2007			

The position in the *A. thaliana* physical map in kilobases, the name and the identification of the loci in *A. thaliana* are shown. The positions of genes involved in metal homeostasis are underlined. The type of amplified fragment length polymorphism is shown: *Ind*eI–LiCor, insertion/deletion detected on an acrylamide gel on an automated sequencer LiCor 4200L; *Eco*RI, *Eco*RV, *Nco*I, *Sph*I, fragment length polymorphisms detected on an agarose gel after PCR-fragment-digestion with the specified restriction enzyme. The primer sequences, the annealing temperature and the extension time used for PCR, if newly defined markers, are indicated.

[a] Primers were designed in Basic and Besnard (2006).

[b] Genomic DNA was digested with restriction enzyme *Eco*RI or *Eco*RV and hybridized with the full complementary DNA corresponding to clone 110G24T7 for Ah–MT2a, AF033535 for Ah–ZIP1 and AF033536 for Ah–ZIP2.

the primers Tc-up3, Tc-up2, Ah-IRT2 were designed in *Thlaspi caerulescens* by Basic and Besnard (2006).

To obtain labelled PCR products detectable on the automated genotyper Li–Cor 4200 (Li-Cor–Science-Tec), the forward and reverse primer contained a 5′-tail of 19 bp (forward primer) or 20 bp (reverse primer) homologous to the universal consensus M13 primer sequence, followed by the locus-specific sequence (Oetting et al. 1995). For all the markers, PCR reactions were carried out in a total volume of 15 µl containing 20 ng of template DNA, 2 mM $MgCl_2$, 0.2 mg/ml BSA, 0.2 mM dNTP, 0.2 µM of each unlabelled tail-specific primer, 0.15 µM of the universal primer M13 fluorescently labelled (either IRD-700 or IRD-800), 20 mM Tris–HCl (pH 8.3), 50 mM KCl and 0.4 units of AmpliTaq® DNA Polymerase (Applied Biosytems). PCR was performed on a Perkin–Elmer Gene–Amp system 9700 (94°C for 5 min, followed by locus-specific amplification: 94°C for 45 s, annealing temperature for 45 s, extension temperature at 72°C for 38 cycles, and a final extension 72°C for 7 min). The annealing temperature and the extension time of the locus-specific amplification varied, respectively, between 50 and 65°C, and between 35 and 75 s depending on the locus (Table 1).

Size polymorphisms of the PCR products – *Ah–ZNT1*, *Tc–up3*, *Ah–NRAMP1*, *Tc–up2*, *Ah–ZIP6*, *Ah–HMADP2*, *AtHCTR1*, *Ah–IRT2* (*Ah*, *Arabidopsis halleri*; *Tc*, *Thlaspi caerulescens*; *At*, *Arabidopsis thaliana*) were estimated by electrophoresis using a Li–Cor genotyper, the Li–Cor 50–700 bp sizing standard as a molecular weight marker, and the Saga–GT software (Table 1). For markers that did not present a size polymorphism (*Ah–PCS2*, *Ah–NAS1*, *Ah–NAS2*, At3g24300, *Ah–CAX1*, *Ah–MTPc1*, *Ah–MHX1*, *HAT4*), sequence analysis was performed in the parental individuals to identify nucleotide variation at specific restriction sites. Polymorphism was detected using PCR-restriction fragment length polymorphism (RFLP; Table 1). Restriction was carried out in a total volume of 20 µl containing 10 µl of PCR product, 0.2 mM spermidine, 1× specific enzyme buffer provided by the supplier and 1 unit of restriction enzyme, for at least 4 h at the appropriate temperature in a Perkin–Elmer Gene–Amp system 9700. Polymorphism was revealed on an agarose gel stained with ethidium bromide.

For the *Ah–MT2a*, *Ah–ZIP1* and *Ah–ZIP2* markers, polymorphism was detected by Southern blot analy-

sis: 10 μg of genomic DNA was digested with 40 U of *Eco*RI or *Eco*RV (Roche), respectively, overnight at 37°C and separated on a 1% TAE–agarose gel, and the gel and blot processed according to standard procedures (Saumitou-Laprade et al. 1999). The DNA was transferred onto an uncharged nylon membrane (Biodyne A from Pall Biosupport, USA) using the vacuum blot system of Pharmacia and UV cross-linked (1.2 J/cm^2) to the nylon support. The clones referenced in GenBank as 110G24T7, AF033535, and AF033536, respectively containing *A. thaliana MT2a*, *ZIP1* and *ZIP2*, were labelled with digoxigenin-*d*-UTP (PCR DIG probe Synthesis Kit from Roche). The probes were hybridized overnight at 68°C as recommended by the supplier, and RFLP variation was revealed by immunological detection and chemiluminescence using CDP Star from Roche (Saumitou-Laprade et al. 1993).

Increasing the density of the linkage map, and QTL analysis

The *Ah* × *Alp* linkage map presented here was constructed from 336 BC1 individuals, using 81 markers anchored in the *A. thaliana* sequence: 61 sequence-based markers previously published (Willems et al. 2007) and 20 newly defined markers to increase the density of the genetic map. Linkage map construction was performed using the Joinmap 3.0 program (Van Ooijen and Voorrips 2001). Linkage groups were obtained at a logarithm-of-odds (LOD) score threshold of 5. Markers along each linkage group were ordered using the sequential method implemented in Joinmap applying goodness-of-fit thresholds of 0.5 and 1 for the linkage groups and the loci, respectively. Kosambi's mapping function was used to obtain map distances (Kosambi 1944). Two-LOD support intervals were obtained on the newly obtained *Ah* × *Alp* map through the use of the MapQTL 4.0 program (Van Ooijen et al. 2002) and the Mapchart program (Voorrips 2002) as described in Willems et al. (2007).

Comparative analysis of the *A. halleri* × *A. l. petraea* genetic linkage map with the *A. thaliana* genome, and transfer of the QTL intervals

The chromosomal position in kilobases (kb) in *A. thaliana* of the markers used in the mapping experiment was obtained using the *A. thaliana* database (http://www.tair.org). Chromosomal rearrangements between the *A. thaliana* linkage map and the *Ah* × *Alp* map, and transfer of the QTL intervals were graphically represented using the Mapchart program (Voorrips 2002).

Results

Increasing linkage map density, the QTL intervals and mapping of potential candidates

The 20 newly defined markers mapped to the 8 linkage groups of the *Ah* × *Alp* linkage map previously published (Willems et al. 2007). For the linkage groups LG2, LG4 and LG6, the goodness-of-fit values were 0.896, 0.541 and 0.837, respectively, and thus slightly exceeded the threshold of 0.5. Regarding the total number of 83 loci detected, only 4 (Chr 1-02985, Chr 1-28476, Chr 5-02626, Chr 5-04824) showed a goodness-of-fit value exceeding the threshold of 1 (Chr 1.255, Chr 1.627, Chr 1.137, Chr 1.196, respectively). Linkage group lengths varied from 53 to 80.2 cM, summing to a total map length of 550 cM. The average distance between two markers corresponded to 6.6 cM, ranging from 0.4 to 23.5 cM.

The three QTLs previously identified (Willems et al. 2007) were confirmed on linkage groups LG3, LG4 and LG6. The higher density of markers around the QTL regions allows recalculation of these QTL intervals, and, consequently, increases the accuracy of these regions by changing the intervals from 24, 4 and 13 cM (Willems et al. 2007) to 5, 3 and 12 cM for Zntol-1, Zntol-2 and Zntol-3 respectively. Five markers colocalized with the LOD support intervals: markers At2g18196 (Chr 2-07228) and *Ah–HMA4* (Chr 2-08286) mapped in the LOD support interval of Zntol-1, markers At2g46530 (Chr 2-19112) and *Ah–MTP1a* (Chr 2-19245) colocalized with Zntol-2, and marker *Ah–MTP1b* colocalized with Zntol-3.

On the total number of 83 markers mapped on the *Ah* × *Alp* linkage map, 23 corresponded to genes reported in *A. thaliana* to be involved in metal homeostasis (Table 1). All the 23 genes mapped at the expected position according to the synteny (Table 2). Twenty are localized outside of the QTL regions and three (*Ah–CCH* and *Ah–HMA4*, and *Ah–MTP1*) are localized within the QTL intervals. In accordance with the synteny, *Ah–MTP1a* colocalizes

Table 2 General features of 23 genes related to metal homeostasis and mapped on the *A. lyrata* × *A. halleri* genetic map

Localisation in *A. thaliana* (kb)	Gene code	*A. thaliana* homologous	Putative function of the encoded protein in *A. thaliana*
Chr 1-00240	Ah–PCS2	At1g03980	Phytochelatin synthase 2
Chr 1-02985	Ah–NAS1	At1g09240	Nicotiamine synthase related protein
Chr 1-03665	Ah–ZNT1	At1g10970	Metal transporter (ZIP4)
Chr 1-21141	Ah–NAS2	At1g56430	Nicotiamine synthase related protein
Ch r1-30378	Ah–NRAMP1	At1g80830	Metal transporter (NRAMP1)
Chr 2-07228	Ah–CCH	At2g18196	Copper Chaperone related protein
Chr 2-08286	Ah–HMA4	At2g19110	Cd/Zn transporting ATPase (HMA4)
Chr 2-09863	Ah–NRAMP3	At2g23150	Metal transporter (NRAMP3)
Chr 2-12846	Ah–ZIP6	At2g30080	Metal transporter (ZIP6)
Chr 2-15997	Ah–CAX1	At2g38170	Ca exchanger (CAX1)
Chr 2-19245	Ah–MTP1a	At2g46800	Zinc transporter (ZAT1, MTP1)
Chr 2-19531	Ah–MHX1	At2g47600	Magnesium/proton exchanger (MHX1)
Chr 2-19598	Ah–MTPc1	At2g47830	Chloroplastic metal transporter putative/metal tolerance protein (MTPc1)
Chr 3-01936	Ah–FRD3	At3g08040	MATE efflux family protein; similar to FRD3 (ferric reductase defective 3)
Chr 3-02889	Ah–MT2a	At3g09390	Metallothionein (MT2a)
Chr 3-04052	Ah–ZIP1	At3g12750	Metal transporter (ZIP1)
Chr 3-21509	Ah–MTPc3	At3g58060	Cation efflux family protein/metal tolerance protein, putative (MTPc3)
Chr 4-10703	Ah–IRT2	At4g19680	iron-responsive Fe/Zn transporter (IRT2)
Chr 4-17540	Ah–HMA1	At4g37270	Cd/Zn transporting ATPase (HMA1)
Chr 5-00506	Ah–MT2b	At5g02380	Metallothionein (MT2b)
Chr 5-04824	Ah–MMADP2	At5g14910	Heavy-metal-associated domain-containing protein
Chr5 -24008	Ah–ZIP2	At5g59520	Zn transporter (ZIP2)
Chr5 -26879	Ah–NRAMP4	At5g67330	Metal transporter (NRAMP4)

The position in the *A. thaliana* physical map in kilobases, the name, the identification loci and the putative function of the encoded protein in *A. thaliana* are shown. Positions of genes mapped in the QTL intervals for Zn tolerance are in bold.

with QTL region Zntol2. However, two duplication events of the *MTP1* gene were detected in *A. halleri* (*MTP1b* and *MTP1c*) (Dräger et al. 2004). *Ah–MTP1b* colocalizes with the QTL region Zntol3, whereas *Ah–MTP1c* is mapped outside of the QTLs of Zn tolerance, in LG1.

Comparative analysis of the *A. halleri* × *A. l. petraea* linkage map and the *A. thaliana* genome

Similarly to the *A. lyrata* subspecies (Koch et al. 2005) and as previously described in Willems et al. (2007), the reduction from eight chromosomes in *A. halleri* to five chromosomes in *A. thaliana* probably involved five major events consisting of three fusions and two reciprocal translocations (Fig. 1). The fusions involved the linkage groups LG1 and LG2 generating *At* Chr.I, the linkage groups LG3 and LG4 resulting in *At* Chr.II, and finally the linkage groups LG7 and LG8 generating the lower part of *At* Chr.V. One translocation event occurred between LG3 and LG5, generating the upper arm of *At* Chr.II and *At* Chr.III, and the other one implicated the linkage groups LG6 and LG7, generating *At* Chr.IV and the upper part of *At* Chr.V. In addition to these important chromosomal rearrangements, three small-scale inversions were detected on the *Ah* × *Alp* map. On the upper arm of LG2, an inversion was identified between the markers Chr 1-22940 and Chr 1-21141 (Fig. 1, I1). The comparison of *At* Chr.II and the linkage groups LG3, LG4 and LG5 of the *Ah* × *Alp* map showed one inversion (Fig. 1, I2) between the markers positioned on the extremity of the lower arm of linkage group LG4. The third inversion (Fig. 1, I3) was identified between the markers Chr 4-04240 and Chr 4-01123, located on the lower arm of LG6. Given their larger

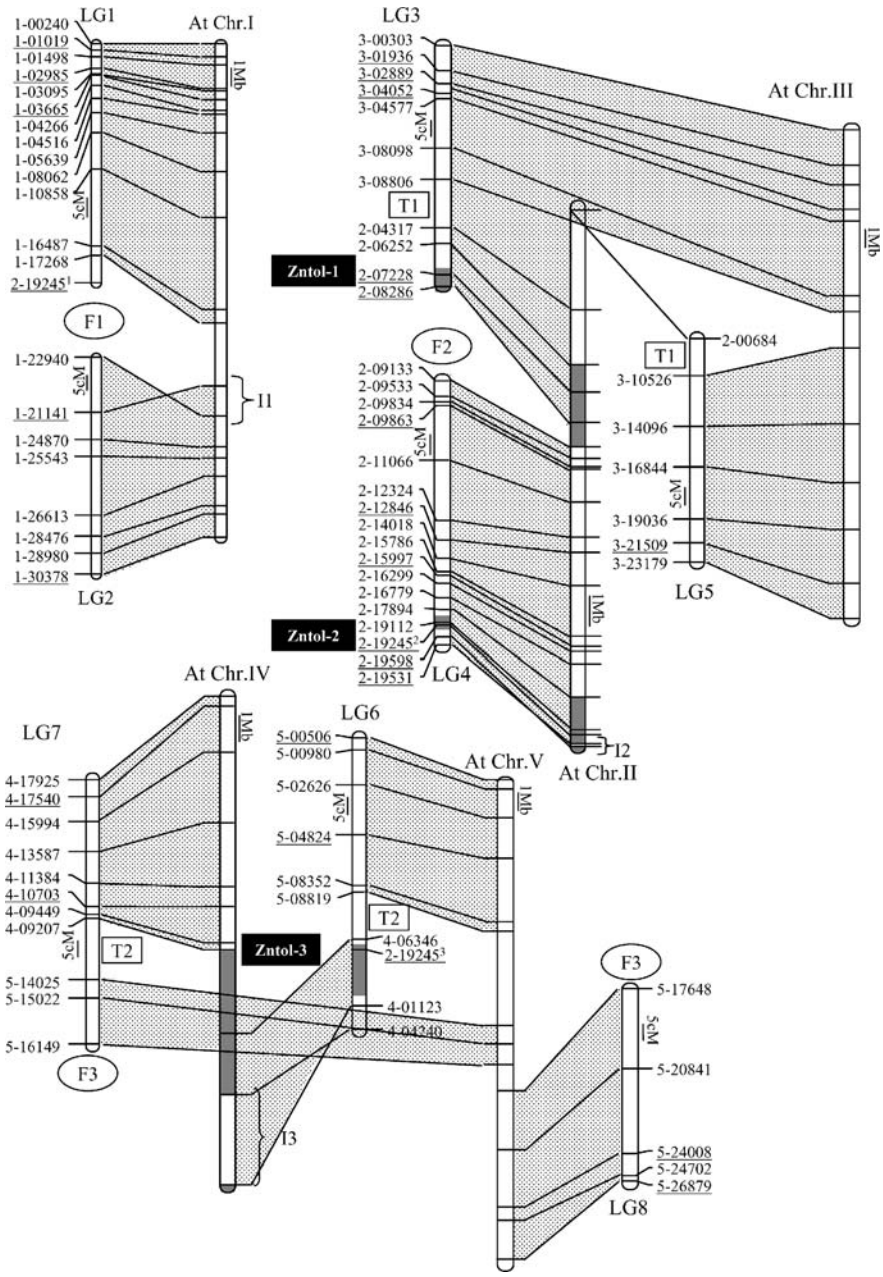

Fig. 1 Identification of the chromosomal rearrangements through a comparative analysis of the *Ah* × *Alp* genetic linkage map (linkage groups *LG1 to 8*) and the *A. thaliana* physical map (*At Chr. I to V*). Markers are labelled according to their localisation (in kilobases) in *A. thaliana*. Underlined markers correspond to genes involved in metal homeostasis. *T1* and *T2* indicate the approximate position of the breakpoints involved in the reciprocal translocation events between linkage groups *LG3* and *LG5*, and *LG6* and *LG7*, respectively. *F1*, *F2* and *F3* indicate the fusion events that implicated the linkage groups LG1 and LG2, LG3 and LG4, and LG7 and LG8, respectively. *I1*, *I2* and *I3* indicate the inversions observed on the *Ah* × *Alp* map compared with the *A. thaliana* physical map. The QTL support intervals Zntol-1, Zntol-2 and Zntol-3 (and their corresponding regions in *A. thaliana*) are highlighted by the *shaded region* on linkage groups LG3 (*At Chr. II*), LG4 (*At Chr. II*) and LG6 (*At Chr. IV*), respectively. Three copies of the MTP1 gene (2-19245 on the (*Ah* × *Alp)* map) were detected in *A. halleri*: [1]Ah–MTP1c (*LG1*), [2]Ah–MTP1a, the ortholog of the MTP1 gene in *A. thaliana* (*LG4*), and [3]Ah–MTP1b (*LG6*)

genome sizes, gene duplication events might be expected to have occurred in *A. thaliana* relatives. One duplication event was revealed on the *Ah* × *Alp* linkage map and concerned the marker 2-19245, corresponding to the metal homeostasis gene *MTP1* (formerly known as *ZAT*; van der Zaal et al. 1999). As previously reported, three copies – *MTP1a*, *MTP1b* and *MTP1c* – were detected in *A. halleri*, and mapped to the linkage groups LG4, LG6 and LG1, respectively, whereas only one copy, corresponding to *Ah–MTP1a*, is present in *A. l. petraea* and *A. thaliana* (Dräger et al. 2004; Willems et al. 2007).

Transfer of the QTL intervals for Zn tolerance from the *Ah* × *Alp* map to the *A. thaliana* map

With regard to the QTL Zntol-1 on the *Ah* × *Alp* map, the upper limit of the two-LOD support interval fell

between the markers Chr 2-07228 and Chr 2-06252 (Fig. 1), whereas the lower limit was situated between the locus Chr 2-08286 and the end of linkage group LG3. On the *A. thaliana* map, the QTL Zntol-1 is thus limited by Chr 2-09133 (the first marker of linkage group LG4).

The upper limit of the confidence interval associated with the QTL Zntol-2 on the *Ah* × *Alp* map was positioned between the markers Chr 2-17894 and Chr 2-19112, and the lower limit localized between the loci Chr 2-19245 and Chr 2-19598, on the linkage group LG4. Because the marker Chr 2-17894 was not implicated in any of the chromosomal rearrangements previously described, this marker was used to define the upper limit of Zntol-2 on the *A. thaliana* map. However, because the marker Chr 2-19598 was involved in an inversion, the QTL interval Zntol-2 was extended to the end of At Chr.II. The very small region between the markers Chr 2-19598 and Chr 2-19531, which was not included in the QTL Zntol-2, was consequently also excluded on the *A. thaliana* map from the QTL interval.

The upper and lower limits of the QTL support interval Zntol-3, on linkage group LG6 of the *Ah* × *Alp* map, fell between the markers Chr 4-01123 and Chr 4-06346. This region was involved in one of the reciprocal translocation events and included the duplication of the *MTP1* gene (*Ah–MTP1b*, Chr 2-19245). Moreover, on the *Ah* × *Alp* map an inversion (Fig. 1, I3) was observed between the marker Chr 4-01123 (that is, the lower limit of the QTL Zntol-3) and Chr 4-04240. Because the region situated between those two markers was not included in the QTL interval on the *Ah* × *Alp* map, it was also excluded from the QTL interval Zntol-3 in *A. thaliana*. The region between marker Chr 4-01123 and the end of the *At* Chr.IV, however, was again included in the QTL interval. Except for its location in proximity to the breakpoint involved in the reciprocal translocation between LG6 and LG7, no rearrangements were observed on the *Ah* × *Alp* map near marker Chr 4-06346. Consequently, the latter could be expected to define the lower limit of the QTL Zntol-3 on the *At* Chr.IV. However, a large inversion was identified in both *A. lyrata* subspecies near the lower limit of this QTL. Because the same event was observed in *Capsella rubella* (Boivin et al. 2004) it might represent the ancestral state of marker order (Yogeeswaran et al. 2005) and it should have occurred in the *A. thaliana* lineage after its divergence from the *A. lyrata* and *A. halleri* species. The first marker following Chr 4-06346 on *At* Chr.IV was thus adopted to define the lower limit of the QTL Zntol-3.

Discussion

Comparative mapping is a powerful tool that is useful not only for addressing questions pertaining to the evolution of genome structure (Kuittinen et al. 2004; Yogeeswaran et al. 2005), but that also enables the transfer of QTL intervals localized on the genetic map of a non-model organism to the physical map of a model species. However, before such an approach becomes feasible, the key question of the level of the macro- and microsynteny between the two related species has to be assessed.

In this study, we first show that primers designed on *A. thaliana* sequences can be used to amplify *A. halleri* and *A. lyrata*, in agreement with the 94% identity detected between the *A. halleri* and *A. thaliana* species (Becher et al. 2004). We also illustrated how genetic markers based on genes from the *A. thaliana* model may help to increase the density of genetic maps of the related non-model species *A. halleri* and aid in obtaining information about the localization of a specific gene. In particular, the higher density of markers in the QTL intervals provided by this work has drastically reduced the length of the QTL two-LOD support intervals from 24, 4 and 13 cM (Willems et al. 2007) to 5, 3 and 12 cM for Zntol-1, Zntol-2 and Zntol-3, respectively. Moreover, most of these primers can probably be used to amplify other Brassicaceae species, as shown by Basic and Besnard (2006), who reported that, using primers anchored on gene coding sequences, 70% of cross-species PCR amplification between species representative of Brassicaceae was successful. Consequently, the primers designed in this study can be further used as a source of transferable gene markers for comparative mapping analyses among species in the Brassicaceae family.

Second, we confirm through an extensive comparative analysis that the *Ah* × *Alp* map is highly collinear with the *A. thaliana* genome, except for three small-scale chromosomal rearrangements. The inversion (I1) observed on the upper arm of linkage group LG2 of the *Ah* × *Alp* map was also detected in

A. l. petraea (Hansson et al. 2006) and *C. rubella* (Boivin et al. 2004). As reported for these species, this marker order probably represents the ancestral state of marker order and thus probably arose in the *A. thaliana* lineage after divergence from the *Arabidopsis* relatives (Boivin et al. 2004; Hansson et al. 2006).

The inversion (I2) observed on LG4 when compared to At Chr.II was not reported for the other *Arabidopsis* relatives. The higher density of markers presented in our *Ah* × *Alp* map could explain the detection of this new inversion. Indeed, the proposed order of markers is robust because the goodness-of-fit values of the linkage group LG4 (0.541), as well as of the inverted markers, were near or below the thresholds applied during linkage map construction. Moreover, these values dramatically increased when the order of markers was fixed according to *A. thaliana* (0.857 for LG4, 3.12 for locus Chr 2-19531, and Chr 2.815 for locus Chr 2-19598).

The inversion (I3) identified on the lower extremity of LG6 on the *Ah* × *Alp* map can be related to the rearrangement described in the *A. l. petraea* map (Kuittinen et al. 2004). In this map, the five markers in the lower end of AL6, namely *GA1* (Chr 4-01238), *LD* (Chr 4-01129), *ICE7* (Chr 4-06945), *AthDET1* (Chr 4-06346) and *CRY1* (Chr 4-05724), were clearly in a different order to in the *A. thaliana* homologous region (*LD*, *GA1*, *CRY1*, *AthDET1*, *ICE7*). The authors inferred from their data at least one non-ambiguous inversion of the region (*CRY1*, *DET1*, *ICE7*) and mentioned with more uncertainty a second inversion concerning the (*GA1*, *LD*) region. Our data are in favour of the occurrence of this last inversion, which can be extended to the region between *AT4GA2* (Chr 4-04240) and *LD* (Chr 4-01124) markers on the basis of our mapping. The sequencing project on the *Arabidopsis* relatives *A. lyrata* and *C. rubella* http://www.jgi.doe.gov/sequencing/why/CSP2006/Alyrata-Crubella.html) should clarify this issue. Moreover, the non-ambiguous inversion of the region (*CRY1*, *DET1*, *ICE7*) reported for the *A. lyrata* subspecies, also concerns a specific duplication event in *A. halleri* at the origin of the *Ah–MTP1b* gene. The characterization of the extension of this duplication will require sequencing effort, specifically in *A. halleri*.

Third, this study demonstrated the possibility of transferring QTL intervals from a species with a poorly known genome structure but with interesting ecological traits to a model species; this transfer allows easier identification of candidate genes for these traits. More precisely, the QTL for Zn tolerance identified on the *Ah* × *Alp* map can now be transferred to the *A. thaliana* genome, rendering possible the selection of candidate genes belonging to the corresponding *A. thaliana* region. However, two major limits of this strategy for identifying candidate genes should be mentioned: (1) because local chromosomal rearrangements can remain undetected, the possibility of an erroneous transfer of QTL intervals should be kept in mind, and (2), as indicated by the genome size differences between *A. thaliana* and its wild relatives (Clauss and Koch 2006), some information is probably missing in the *A. thaliana* genome and therefore could remain unavailable through this approach.

Finally, given the small-scale rearrangements observed on the *Ah* × *Alp* map, this study highlighted the importance of mapping candidate genes to verify their effective presence or absence in the QTL intervals. On the *Ah* × *Alp* linkage map, 23 genes reported in *A. thaliana* to be involved in metal homeostasis were present at the expected position, according to synteny. Among them *Ah–CCH* and *Ah–HMA4* colocalize with Zntol-1, and *Ah–MTP1a* with Zntol-2. However, the mapping of *Ah–MTP1* genes highlights the limitation of using only comparative mapping to identify genes present in the QTL region of non-model organisms, because the presence of a duplicated gene, such as *Ah–MTP1b,* which colocalizes with Zntol-3, cannot be predicted.

In conclusion, we have demonstrated that comparative mapping is an important approach to not only addressing questions pertaining to the evolution of *A. halleri* genome, but also to transferring QTL intervals for Zn tolerance from the *Ah* × *Alp* genetic map to the physical map of the model species, *A. thaliana*. Our comparative mapping may help scientists to transfer to *A. thaliana* any QTL region of interest identified in *A. halleri*. Moreover, the analysis of genes present in the Zn tolerance QTL intervals transferred to *A. thaliana* may help to identify further candidates for specific metal tolerance characteristics.

Acknowledgement The authors thank Aude Bodin for technical support in the genetic mapping. This work was supported by funding from the Contrat de Plan Etat/Région Nord-Pas de Calais (PRC), from the European FEDER (contract no. 79/1769), from the BRG (contract no. 92), and from the INSU–CNRS program ACI ECCO (contract no. 04 2

9 FNS). G. Willems was funded by the European Research Training Network "Metalhome" (HPRN-CT-2002-00243), and N. Roosens by the Marie Curie intra-European Fellowship "Metolevol" (contract no. 024683 MEIF-CT-2005-0224683).

References

Assuncao AGL, Schat H, Aarts MGM (2003) *Thlaspi caerulescens*, an attractive model species to study heavy metal hyperaccumulation in plants. New Phytol 159:351–360

Basic N, Besnard G (2006) Gene polymorphisms for elucidating the genetic structure of the heavy-metal hyperaccumulating trait in *Thlaspi caerulescens* and their cross-genera amplification in Brassicaceae. J Plant Res 119:479–487

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 J 37:251–268

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 23:1741–1750

Boivin K, Acarkan A, Mbulu RS, Clarenz O, Schmidt R (2004) The Arabidopsis genome sequence as a tool for genome analysis in Brassicaceae. A comparison of the Arabidopsis and *Capsella rubella* genomes. Plant Physiol 135:735–744

Clemens S, Palmgrem MG, Kramer U (2002) A long way ahead: understanding and engineering plant metal accumulation. Trends Plant Sci 7:309–315

Chiang HC, Lo JC, Yeh KC (2006) Genes associated with heavy metal tolerance and accumulation in Zn/Cd hyperaccumulator *Arabidopsis halleri*: a genomic survey with cDNA microarray. Environ Sci Technol 40:6792–6798

Clauss MJ, Koch MA (2006) Poorly known relatives of *Arabidopsis thaliana*. Trends Plant Sci 11:449–459

Courbot M, Willems G, Motte P, Arvidsson S, Roosens N, Saumitou-Laprade P, Verbruggen N (2007) A major quantitative trait locus for cadmium tolerance in *Arabidopsis halleri* colocalizes with HMA4, a gene encoding a heavy metal. Plant Physiol 14:1052–1065

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 J 39:425–439

Filatov V, Dowdle J, Smirnoff N, Ford-Lloyd B, Newbury HJ, Macnair M (2006) Comparison of gene expression in segregating families identifies genes and genomic regions involved in a novel adaptation, zinc hyperaccumulation. Mol Ecol 15:3045–3059

Grotz N, Fox T, Connolly E, Park W, Guerinot ML, Eide D (1998) Identification of a family of zinc transporter genes from Arabidopsis that respond to zinc deficiency PNAS 95:7220–7224

Hansson B, Kawabe A, Preuss S, Kuittinen H, Charlesworth D (2006) Comparative gene mapping in *Arabidopsis lyrata* chromosomes 1 and 2 and the corresponding *A. thaliana* chromosome 1: recombination rates, rearrangements and centromere location. Genet Res 87:75–85

Koch MA, Kiefer M (2005) Genome evolution among cruciferous plants: a lecture from the comparison of the genetic maps of three diploid species *Capsella rubella*, *Arabidopsis lyrata* subsp. *petraea*, and *A. thaliana*. Am J Bot 92:761–767

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

Kuittinen H, de Haan AA, Vogl C, Oikarinen S, Leppala J, Koch M, Mitchell-Olds T, Langley CH, Savolainen O (2004) Comparing the linkage maps of the close relatives *Arabidopsis lyrata* and *A. thaliana*. Genetics 168:1575–1584

Macnair MR, Bert V, Huitson SB, Saumitou-Laprade P, Petit D (1999) Zinc tolerance and hyperaccumulation are genetically independent characters. Proc R Soc Lond B Biol Sci 266:2175

Oetting WS, Lee HK, Flanders DJ, Wiesner GL,. Sellers TA et al (1995) Linkage analysis with multiplexed short tandem repeat polymorphism using infrared fluorescence and M13 tailed primers. Genomics 30:450–458

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

Peer WA, Mamoudian M, Lahner B, Reeves RD, Murphy AS, Salt DE (2003) Identifying model metal hyperaccumulating plants: germplasm analysis of 20 Brassicaceae accessions from a wide geographical area. New Phytol 159:421–430

Reeves RD, Schwartz C, Morel JL, Edmondson J (2001) Distribution and metal-accumulating behavior of *Thlaspi caerulescens* and associated metallophytes in France. Int J Phytoremediat 3:145–172

Saumitou-Laprade P, Rouwendal GJA, Cuguen J, Krens FA, Michaelis G (1993) Different CMS sources found in *Beta vulgaris* ssp. *maritima*: mitochondrial variability in wild populations revealed by a rapid screening procedure. Theor Appl Genet 85:529–535

Saumitou-Laprade P, Piquot Y, Raspé O, Bernard J, Vrieling K (1999) Plant DNA fingerprinting and profiling. In: JT Epplen, T Lubjuhn (eds) DNA profiling and DNA fingerprinting. Birkhauser, Basel, pp 17–38

Van der Zaal BJ, Neuteboom LW, Pinas JE, Chardonnens AN, Schat H, Verkleij JAC (1999) Overexpression of a novel Arabidopsis gene related to putative zinc transporter genes from animals can lead to enhanced zinc resistance and accumulation. Plant Physiol 119:1047–1055

Van Ooijen JW, Voorrips RE (2001) Joinmap 3.0, Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, The Netherlands

Van Ooijen JW, Boer MP, Jansen RC, Maliepaard C (2002) MapQTL 4.0, Software for the calculation of QTL positions on genetic maps. Plant Research International, Wageningen, The Netherlands

Voorrips RE (2002) Mapchart: Software for the graphical presentation of linkage maps and QTLs. Heredity 93:77–78

Weber M, Harada E, Vess C, von Roepenack-Lahaye E, 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 J 37:269–281

Weber M, Trampczynska A, Clemens S (2006) Comparative transcriptome analysis of toxic metal responses in *Arabidopsis thaliana* and the Cd2+–hypertolerant facultative metallophyte *Arabidopsis halleri*. Plant Cell Environ 29:950–963

Willems G, Drager 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

Yogeeswaran K, Frary A, York TL, Amenta A, Lesser AH et al (2005) Comparative genome analyses of *Arabidopsis* spp.: inferring chromosomal rearrangement events in the evolutionary history of *A. thaliana*. Genome Res 15:505–515

Zhou J, Goldsbrough PB (1994) Functional homologs of fungal metallothionein genes from Arabidopsis. Plant Cell 6:875–884