

QTL analysis of cadmium and zinc accumulation in the heavy metal hyperaccumulator *Thlaspi caerulescens*

A. X. Deniau · B. Pieper · W. M. Ten Bookum ·
P. Lindhout · M. G. M. Aarts · H. Schat

Received: 14 September 2005 / Accepted: 21 June 2006 / Published online: 19 July 2006
© Springer-Verlag 2006

Abstract *Thlaspi caerulescens* (*Tc*; $2n = 14$) is a natural Zn, Cd and Ni hyperaccumulator species belonging to the Brassicaceae family. It shares 88% DNA identity in the coding regions with *Arabidopsis thaliana* (*At*) (Rigola et al. 2006). Although the physiology of heavy metal (hyper)accumulation has been intensively studied, the molecular genetics are still largely unexplored. We address this topic by constructing a genetic map based on AFLP[®] markers and expressed sequence tags (ESTs). To establish a genetic map, an F₂ population of 129 individuals was generated from a cross between a plant from a Pb/Cd/Zn-contaminated site near La Calamine, Belgium, and a plant from a compa-

rable site near Ganges (GA), France. These two accessions show different degrees of Zn and, particularly, Cd accumulation. We analyzed 181 AFLP markers (of which 4 co-dominant) and 13 co-dominant EST sequences-based markers and mapped them to seven linkage groups (LGs), presumably corresponding to the seven chromosomes of *T. caerulescens*. The total length of the genetic map is 496 cM with an average density of one marker every 2.5 cM. This map was used for Quantitative Trait Locus (QTL) mapping in the F₂. For Zn as well as Cd concentration in root we mapped two QTLs. Three QTLs and one QTL were mapped for Zn and Cd concentration in shoot, respectively. These QTLs explain 23.8–60.4% of the total variance of the traits measured. We found only one common locus (LG6) for Zn and Cd (concentration in root) and one common locus for shoot and root concentrations of Zn (LG1) and of Cd (LG3). For all QTLs, the GA allele increased the trait value except for two QTLs for Zn accumulation in shoot (LG1 and LG4) and one for Zn concentration in root (LG1).

Communicated by O. Savolainen.

A. X. Deniau (✉) · W. M. Ten Bookum · H. Schat
Ecology and Physiology of Plants,
Vrije Universiteit Amsterdam,
de Boelelaan 1087, 1081 HV Amsterdam,
The Netherlands
e-mail: antoine.deniau@falw.vu.nl

B. Pieper · M. G. M. Aarts
Laboratory of Genetics, Wageningen University,
Droevendaalsesteeg 1, 6708 PB Wageningen,
The Netherlands

P. Lindhout
De Ruiter Seeds R&D NL BV, leeuwenhoekweg 52,
2660 BB Bergschenhoek, The Netherlands

Present Address:

B. Pieper
Department of Plant Breeding and Genetics,
Max Planck Institute for Plant Breeding Research,
Carl-von-Linné-Weg 10,
50829 Köln, Germany

Introduction

Thlaspi caerulescens is a heavy metal hyperaccumulator plant species that is able to accumulate extremely high levels of zinc (Zn) and cadmium (Cd) in its shoots (30,000 µg Zn g⁻¹ dry weight (DW) and 10,000 µg Cd g⁻¹ DW). For many years, it has been the subject of intensive research to better understand the mechanism of heavy metal hyperaccumulation and tolerance. *T. caerulescens* represents a potential source of genes for engineering heavy metal phytoremediation in plants (Assunção et al. 2003b; Cobbett 2003).

A major advantage of *T. caerulescens* over the other hyperaccumulator plant species is the available genetic variation for metal specificity with regard to accumulation, translocation and tolerance traits between different *T. caerulescens* accessions (Meerts and van Isacker 1997; Escarré et al. 2000; Lombi et al. 2000; Schat et al. 2000). This intraspecific variation permits a genetic analysis of these traits in segregating populations generated from intraspecific crosses, including quantitative trait locus (QTL) mapping analysis (Alonso-Blanco and Koornneef 2000). As the genus *Thlaspi* is closely related to the genus *Arabidopsis* (Koch et al. 2001), it is possible to assess the synteny between these species, such as previously done with *Arabidopsis lyrata* and *Arabidopsis thaliana* (Kuittinen et al. 2004).

AFLP analysis has been shown to be well-suited for genotyping and map construction in many plant species. We constructed an AFLP-based comprehensive genetic linkage map based on the F₂ progeny of the inter-accession cross between La Calamine (LC) and Ganges (GA). The GA accession combines the properties of extreme Cd accumulation and Cd tolerance, while the LC accession exhibits a much lower accumulation and a slightly lower tolerance. The accumulation and tolerance of Zn are comparable in both accessions. Despite the similarities, the mechanisms of both Zn and Cd accumulation seem to differ between these accessions (Lombi et al. 2000; Zhao et al. 2002; Zha et al. 2004). In addition to constructing a linkage map, we used the population to identify QTLs controlling Cd or Zn accumulation in order to further understand heavy metal hyperaccumulation, tolerance and homeostasis in *T. caerulescens*.

Materials and methods

Plant origin and crossing scheme

A cross was made between a plant grown from *T. caerulescens* J. & C. Presl seeds collected at a strongly Pb/Cd/Zn-enriched site near La Calamine, Belgium, and a plant grown from seeds collected at a similar calamine site near Ganges, France (Zhao et al. 2002). The LC plant was used as a mother. Based on former studies maternal inheritance was not expected (Zha et al. 2004). One F₂ family, derived from a single self-pollinated F₁ plant, was sown and about 130 F₂ plants were analyzed.

Plant culture and vernalization

Plants were grown from seeds sown on moist peat. Three-week-old seedlings were transferred to 1-l

polyethylene pots (two seedlings per pot) filled with modified half-strength Hoagland's nutrient solution: 3 mM KNO₃, 2 mM Ca(NO₃)₂, 1 mM NH₄H₂PO₄, 0.5 mM MgSO₄, 1 μM KCl, 25 μmol H₃BO₃, 2 μM ZnSO₄, 2 μM MnSO₄, 0.1 μM CuSO₄, 0.1 μM (NH₄)₆Mo₇O₂₄, 20 μM Fe(Na)ethylene-diamine-tetraacetic acid. The pH buffer 2-(*N*-morpholino)ethanesulfonic acid was added to a final concentration of 2 mM, and the pH was set at 5.5 using KOH. The solutions were replaced twice a week. All the crossings and experiments were performed in a climate chamber (20/15°C day/night; photon flux density 250 μmol m⁻² s⁻¹ at plant level; 14 h day⁻¹; 75% relative humidity, RH). To induce flowering, pots with 5-week-old plants were transferred to a cold growth cabinet (4°C day and night; 200 μmol m⁻² s⁻¹ at plant level; 12 h day⁻¹; ± 60% RH) where they remained for 5–6 weeks, after which they returned to the climate chamber.

Cd and Zn accumulation in parental accessions and F₂

Three-week-old seedlings from a sample of bulked seeds of the parental accessions LC and GA (49 LC plants and 56 GA plants), and 129 F₂ plants were grown in nutrient solution (two seedlings per pot) containing 2 μM ZnSO₄ and supplemented with 5 μM CdSO₄ (Assunção et al. 2003a). This concentration level was found to yield the highest relative difference in Cd accumulation between LC and GA (Assunção et al. 2003c). The nutrient solutions were replaced by a fresh one twice a week, and after 2 weeks the leaves grown during metal exposure and half of the root system were harvested. Plants survive this treatment and are able to complete their life cycle. The roots were desorbed for 30 min in 5 mM ice-cold Pb(NO₃)₂. Roots and shoots were dried overnight in a stove at 70°C, wet-ashed in a 4:1 mixture of HNO₃ (65%) and HCl (37%) in Teflon bombs at 140°C for 7 h and analyzed for Cd and Zn using flame atomic absorption spectrometry (Perkin Elmer 1100B, Perkin Elmer Nederland, Nieuwerkerk a/d IJssel, the Netherlands).

Statistical analysis

ANOVA was used to test the statistical significance of the differences in metal concentrations between the parental accessions. Correlations between elemental concentrations in F₂ plants were tested using the Pearson correlation coefficient. Data were transformed logarithmically before statistical analysis to obtain homogeneity of variances. Transgression, i.e., segregation beyond the limits of the phenotype distributions of

the parental controls, was tested for by comparing the observed and expected frequencies of F_2 s in the lowest and the highest 2.5%-area sections of the parental control distributions. The 2.5%-area limits were calculated as the antilog of the values representing the means $\pm t \times SD$. Expected frequencies were calculated assuming that 75% of the F_2 s were distributed like the parental distribution to be tested against. The probabilities of observed frequencies were considered to equal the relative frequencies of the Poisson distributions with the expected frequencies as the means. We tested each of the markers for normal Mendelian segregation using a χ^2 -test at a significance level of 0.05 followed by a correction according to the Bonferroni–Holm sequential method (Rice 1989).

DNA extraction

DNA was extracted from two freshly harvested leaf disks of 2-week-old plants following the CTAB protocol of Qi and Lindhout (1997). The original parents were lost and by taking four individuals of each original parent accession, we expect most of the alleles present in the original parents to be represented in the pool. DNA was extracted and pooled from four individuals per accession.

AFLP markers

The AFLP technique was performed as described by Vos et al. (1995), with some minor modifications as described by Qi and Lindhout (1997), using the enzyme combinations *EcoRI/MseI* (*E/M*). The number of AFLP fragments (<100 bands) and their polymorphism rates between the parental lines GA and LC (>20%) were used as selection criteria for the initial pre-screening of 24 *E/M* primer combinations (PCs) (E32M11–22 and E35M11–22; Table 1). PCs with at least 15 easily scorable polymorphic AFLP fragments were selected. The 19 *E/M* PCs to be used on the F_2 population were E32M11, E32M12, E32M13, E32M14, E32M15, E32M16, E32M17, E32M18, E32M19, E32M20, E32M21, E32M22, E35M12, E35M13, E35M14, E35M17, E35M19, E35M20 and E35M22 (Table 2). The adapter and primer sequences employed were based on the core primer design described by Vos et al. (1995). Gel images were electronically scanned with a Licor machine (Westburg, Leusden) and AFLP markers were dominantly scored using Quantar-Pro software (Keygene, Wageningen, the Netherlands). Each polymorphic AFLP band was identified by a code referring to the PC, followed by the estimated size of the DNA fragments in nucleo-

Table 1 List of AFLP primers and adapters. DNA sequences are always given in the 5' to 3' orientation unless indicated otherwise

| Primers/adapters | Sequences |
|--|---|
| <i>EcoRI</i> adapter | 5'-CTCGTAGACTGCGTACC-3'; 3'-CTGACGCATGGTTAA-5' |
| E00 (universal primer) | GACTGCGTACCAATTC |
| <i>EcoRI</i> + 1 selective nucleotide | |
| E01 | E00 + A |
| <i>EcoRI</i> + 3 selective nucleotides | |
| E32 | E00 + AAC |
| E35 | E00 + ACA |
| <i>MseI</i> adapter | 5'-GACGATGAGTCCTGAG-3'; 3'-TACTCAGGACTCAT-5' |
| M00 (universal primer) | GATGAGTCCTGAGTAA |
| <i>MseI</i> + 0 selective nucleotide | |
| M00 | M00 |
| <i>MseI</i> + 2 selective nucleotides | |
| M11 | M00 + AA |
| M12 | M00 + AC |
| M13 | M00 + AG |
| M14 | M00 + AT |
| M15 | M00 + CA |
| M16 | M00 + CC |
| M17 | M00 + CG |
| M18 | M00 + CT |
| M19 | M00 + GA |
| M20 | M00 + GC |
| M21 | M00 + GG |
| M22 | M00 + GT |

tides (e.g., E32M15-143.5) with reference to the SequaMark 10-bp DNA ladder (Research Genetics, Huntsville, AL, USA). In order to score co-dominant markers, Quantar-Pro software was used. This program can detect co-dominant markers based on intensity differences of the corresponding AFLP bands. For each marker detected with the Quantar-Pro program, the reliability was checked by comparison with the marker profiles of the parent accessions, and the co-dominant scoring of the mapping population was checked by analysis of the monogenic segregation ratios.

CAPS and indels

The sequences of *T. caerulea* ESTs (Rigola et al. 2006) of which the Arabidopsis orthologues are evenly distributed over the *At*-genome, have been used to develop a set of 13 co-dominant genetic markers (Table 3). Each EST sequence was compared to the *At*-genome using BLASTN (Altschul et al. 1990) and intron flanking primers were designed. PCR-amplified fragments for LC and GA were sequenced and polymorphisms were determined using Vector NTI Suite 9™ (Invitrogen, Breda, The Netherlands). Cleaved

Table 2 Number of bands and polymorphism detected by the 19 selected AFLP primer combinations

| Primer combination | Selective nucleotides | Total bands | Polymorphic bands | Percent polymorphism |
|--------------------|-----------------------|-------------|-------------------|----------------------|
| E32M11 | AAC/AA | 65 | 19 | 29.23 |
| E32M12 | AAC/AC | 85 | 25 | 29.41 |
| E32M13 | AAC/AG | 65 | 19 | 29.23 |
| E32M14 | AAC/AT | 72 | 19 | 26.39 |
| E32M15 | AAC/CA | 53 | 16 | 30.19 |
| E32M16 | AAC/CC | 65 | 21 | 32.31 |
| E32M17 | AAC/CG | 72 | 16 | 22.22 |
| E32M18 | AAC/CT | 66 | 17 | 25.76 |
| E32M19 | AAC/GA | 66 | 17 | 25.76 |
| E32M20 | AAC/GC | 62 | 19 | 30.65 |
| E32M21 | AAC/GG | 57 | 16 | 28.07 |
| E32M22 | AAC/GT | 59 | 20 | 33.90 |
| E35M12 | ACA/AC | 64 | 18 | 28.13 |
| E35M13 | ACA/AG | 55 | 16 | 29.09 |
| E35M14 | ACA/AT | 51 | 17 | 33.33 |
| E35M17 | ACA/CG | 44 | 17 | 38.64 |
| E35M19 | ACA/GA | 65 | 15 | 23.08 |
| E35M20 | ACA/GC | 60 | 17 | 28.33 |
| E35M22 | ACA/GT | 59 | 19 | 32.20 |
| Total | | 1,185 | 343 | |
| Average | | 62.37 | 18.05 | 29.26 |

Amplified Polymorphic Sequence (CAPS) markers were designed if the polymorphism removed or created a restriction site. Large (>15 bp) insertions or deletions (indels) were scored on 2–3% agarose or metaphor gels, whereas acrylamide gels were used to separate small indels (<15 bp) (Table 3).

Mapping analysis

Marker segregation data were obtained by analyzing the entire F_2 population for AFLP markers using 19 selected E/M PCs together with the CAPS and indel markers. Linkage analysis was performed with the computer software package JOINMAP 3.0 (Stam 1993; Stam and van Ooijen 1996). Kosambi's mapping function (Kosambi 1944) was used to convert recombination frequencies into map distances (cM). All dominant and co-dominant markers were used for the linkage map construction. The pair-wise analysis obtained from JOINMAP was used to assign markers to linkage groups (LGs) with Log of Odds (LOD) score 4. The LOD score calculated by JOINMAP for the recombination frequency is based on the G^2 statistic for independence in a two-way contingency table: $G^2 = 2 \sum o \ln(o/e)$ with o the observed and e the expected number of individuals in a cell, \ln the natural logarithm and \sum the sum over all cells. Under the null hypothesis the statistic has a chi-square distribution with the number of rows minus one multiplied by the number of columns minus one as degrees of freedom

(df). The test for independence is not affected by segregation distortion like the LOD score employed normally in linkage analysis (i.e., the 10-log likelihood ratio comparing the estimated value of recombination frequency with 0.5), thus leading to less incidence of spurious linkage (Stam 1993; Stam and van Ooijen 1996). When the integrated map was compared with each of the parental maps, in general a similar marker order was found (data not shown).

QTL mapping

Potential QTLs for each trait were identified using the MAPQTL 5.0 package (van Ooijen 2004). Kruskal–Wallis and interval mapping analyses were initially performed to find regions with potential QTL effects. If the markers directly flanking the QTL do not provide complete information, genetic information from five markers surrounding the assumed QTL map position is used to calculate the component probabilities. AFLP markers are dominant which means that the genotype is just partially known). MAPQTL tries to resolve these so-called incomplete genotypes as much as possible, by taking into account the markers beyond it on the map. In the case that all of the linked marker genotypes are unknown and cannot be used to calculate the component probabilities, then their average values over the population are used (van Ooijen 2004). Markers with the highest LOD value were then used in various combinations as co-factors in multiple QTL models in Multiple

Table 3 List of EST-based markers and their *A. thaliana* orthologous genes

| EST/gene | Primers | <i>At</i> ortholog | Common name | Description | Marker type |
|-------------|--|--------------------|-------------|--|----------------------|
| RR6nr032 | F 5'-GGATGATAAGATCGAACCTGAAGAGGC-3'; R 5'-TGGTTCATACCAACCAAGAGACTTGG-3' | None | None | None | Indels—180 bp |
| RR1nr059 | F 5'-ACCAACTCCAACCTCTCTCTCTCC-3'; R 5'-AAGAAATCGGATGGTGGACTGAAGC-3' | At1g09560 | GLP5 | Germin-like protein | CAPS— <i>DraI</i> |
| <i>ZNT1</i> | F 5'-CATCGCCGATCTTTTGGAAATC-3'; R 5'-TCGTCTCGACAGAGTCTGATTCG-3' | At1g10970 | ZIP4 | Metal transporter ZIP4, similar to <i>TcZNT1</i> | CAPS— <i>MseI</i> |
| <i>ZNT2</i> | F 5'-CGTAAGACCCCAATGTTCTCATCG-3'; R 5'-TGAAGAAGCAGCCATTGTCTCTGG-3' | At1g60960 | <i>IRT3</i> | Metal transporter <i>IRT3</i> , similar to <i>TcZNT2</i> | CAPS— <i>MseI</i> |
| RR7nr016 | F 5'-TCGGATCTCCAGCAAACCGG-3'; R 5'-AATAGCTTCGAGCTTGGCGTCG-3' | At1g78080 | TFRAP2,4 | AP2 domain transcription factor RAP2.4 | CAPS— <i>BcgI</i> |
| RR19nr015 | F 5'-AAAGGCTTTTCTGCTTCAACACTGTC-3'; R 5'-TCAGGATGAGAAATCGATCATTGG-3' | At2g36540 | NIF | NLI interacting factor family protein | CAPS— <i>DdeI</i> |
| RR11nr025 | F 5'-GTGGTAAACATCACTCTCTCCCTGG-3'; R 5'-AAGCATTTAGCACTCCTACTCCGGC-3' | At2g36830 | GAMMA-TIP1 | Major intrinsic protein family | Indels—15 bp |
| RR4nr003 | F 5'-TGTTCCTTTTACAAGTCGGCG-3'; R 5'-TCCCTGCTCTCTGTAATCGAAC-3' | At3g19820 | DWARF1 | Cell elongation protein | Indels—22 bp |
| RR22nr089 | F 5'-TCCGTTGTAGTCCAGGC-3'; R 5'-CTTCGTCCTTGTGGTTGATCG-3' | At3g26310 | Cyt.P450 | Cytochrome P450 family | CAPS— <i>HindIII</i> |
| RR3nr075 | F 5'-TCAGAGTTCAAGGAAGCGTTAGCC-3'; R 5'-CATCACGGTCCCAAGCTCTTC-3' | At5g21274 | CAM6 | Calmodulin-6 | CAPS— <i>Eco32I</i> |
| RR25nr081 | F 5'-TACCTCGAGTCTGAAATCCTTCCAAG-3'; R 5'-GGCCTAAACAAATGTGCTAATGGAG-3' | At5g23140 | ClpP2 | ATP-dependent Clp protease | CAPS— <i>TruI</i> |
| RR1nr041 | F 5'-GGGCCCTGGAATACTCTTCTTC-3'; R 5'-GCTCTTTTGTCACTTGCAGAAACCG-3' | At5g25090 | PLDC | Proteolytic subunit of plastocyanin-like domain containing protein | Indels—32 bp |
| RS26nr082 | F 5'-AATCGCCGGTACCGGAAAG-3'; R 5'-TTACAATCCAGCCAGAGGAATC-3' | At5g67330 | NRAMP4 | NRAMP metal ion transporter 4 | CAPS— <i>BfaI</i> |

F forward, *R* reverse

QTL Model analysis (MQM analysis also performed with MAPQTL). Using MAPQTL software, permutation tests (1,000 permutations) calculated LOD score thresholds for each trait which was applied to declare the presence of a QTL. This corresponds to a general genome-wide significance of $P < 0.05$ for normally distributed data, as was determined by extensive simulation experiments (van Ooijen 1999). The QTL graphs were prepared with MAPCHART 2.2 (Voorrips 2002).

Results

Characterization of Zn and Cd accumulation in LC, GA and the F₂ progeny

The ranges of shoot and root Zn concentrations for the GA accession (1.5–5.1 and 0.5–2.6 $\mu\text{mol g}^{-1}$ DW,

respectively, $n = 56$) were broadly overlapping with those for the LC accession (2.1–10.0 and 0.3–1.6 $\mu\text{mol g}^{-1}$ DW, respectively; $n = 49$) (Fig. 1). We observed only little overlap for Cd concentration in shoot (Fig. 2; 0.4–23.7 $\mu\text{mol g}^{-1}$ DW for GA and 0.1–1.3 $\mu\text{mol g}^{-1}$ DW for LC) and no overlap for root Cd concentrations [GA (3.2–12.4 $\mu\text{mol g}^{-1}$ DW), LC (0.2–1.8 $\mu\text{mol g}^{-1}$ DW)]. On average, shoot and root Cd concentrations were 6.7-fold ($P < 0.001$) and 12.1-fold ($P < 0.001$) higher in GA than in LC. The Zn concentrations were 0.6-fold ($P < 0.001$) and 2.3-fold higher ($P < 0.001$), respectively (Table 4).

The set of 129 F₂ plants clearly segregated for Cd and Zn shoot and root concentrations (Figs. 1, 2). The frequency distribution of Zn concentration in shoot showed significant transgression beyond the higher limits ($P < 0.001$) of the distribution of the parental accessions. Also, there was a clear transgression

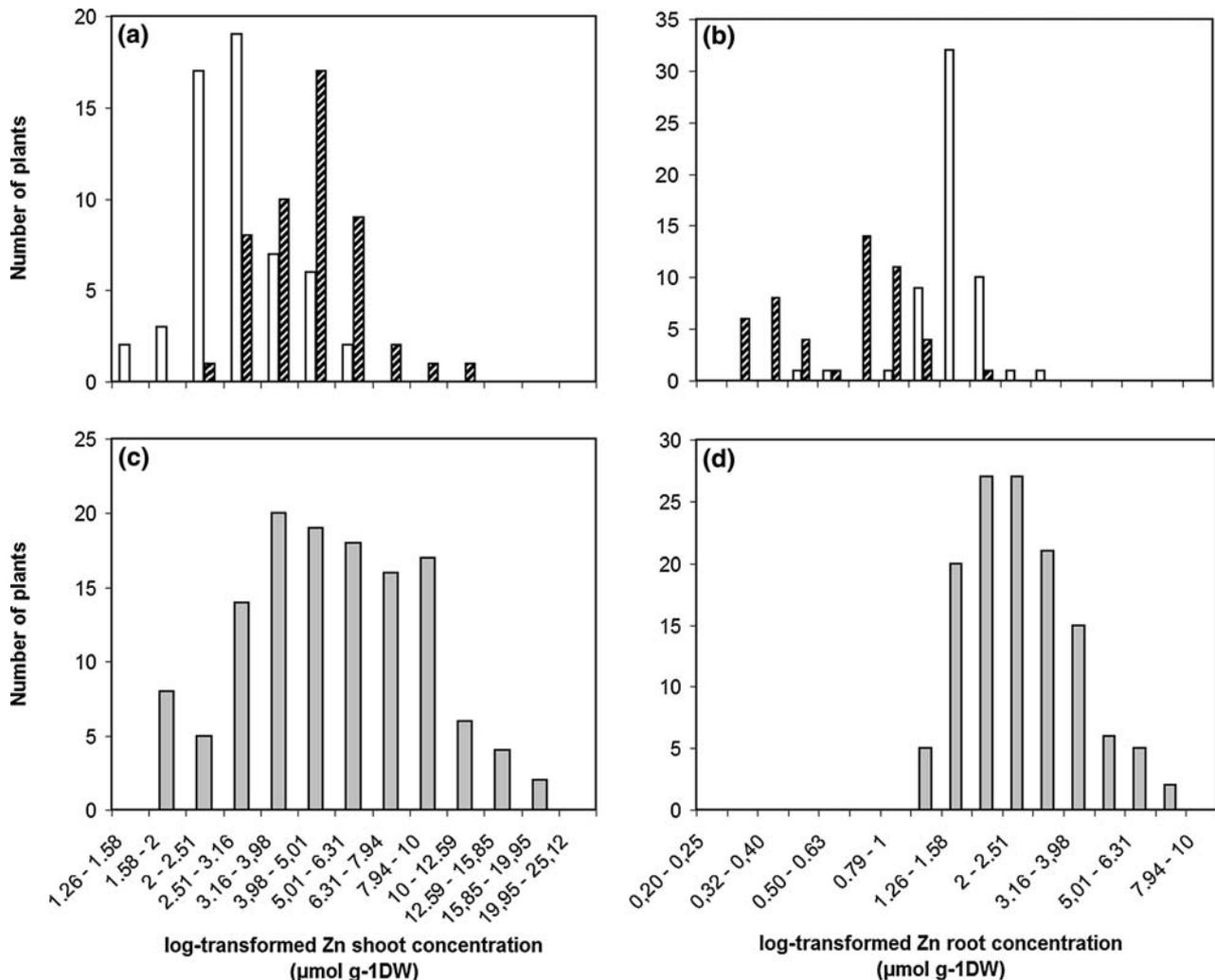


Fig. 1 Frequency distributions of log-transformed Zn shoot concentration (a), root concentration (b) in the LC (dashed bars) and GA (white bars) parental controls, and in the F₂ (gray bars) from

the cross LC \times GA; (c) and (d), respectively. *T. caerulea* plants were grown hydroponically with 5 $\mu\text{mol CdSO}_4$ and 2 $\mu\text{mol ZnSO}_4$

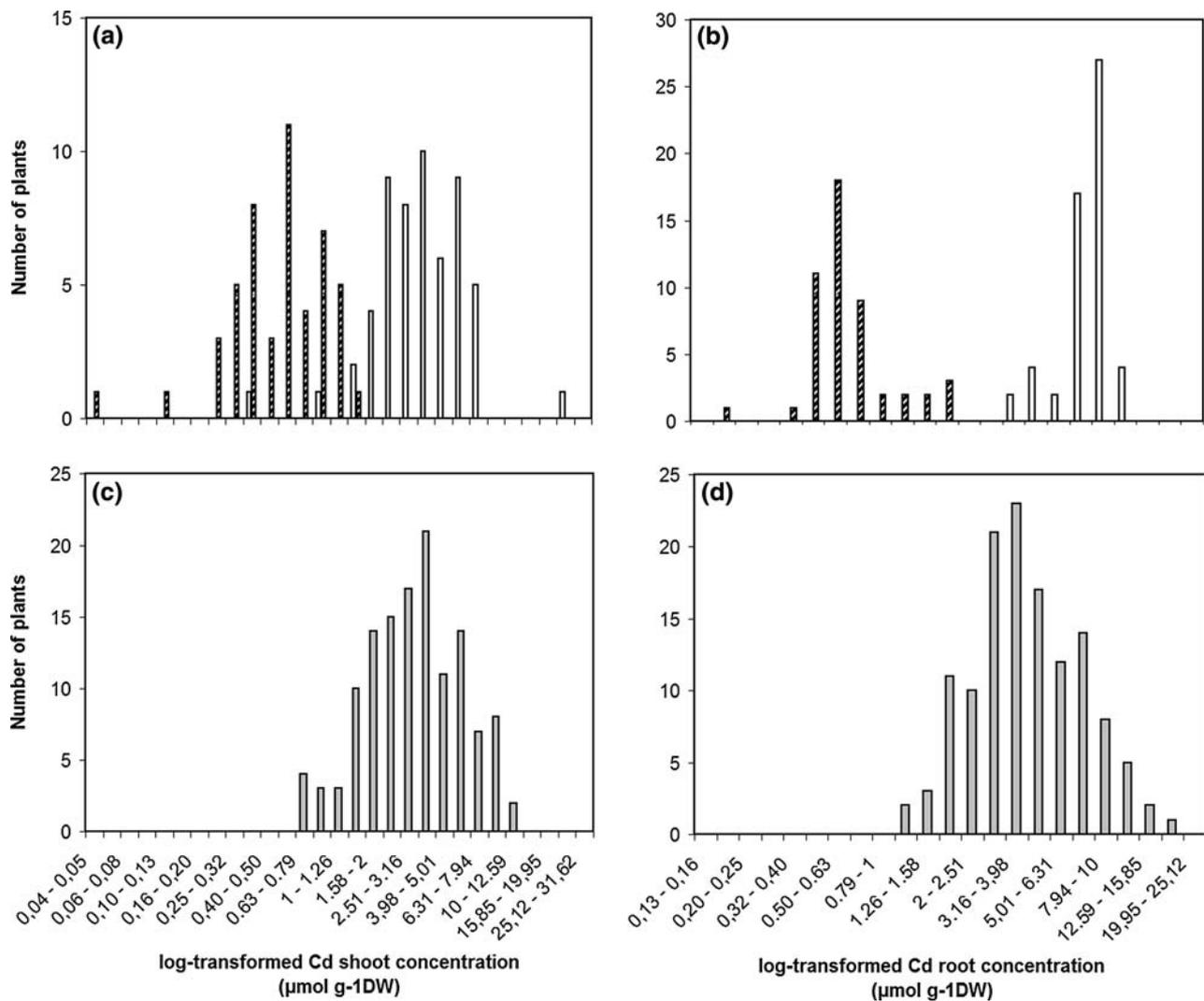


Fig. 2 Frequency distributions of log-transformed Cd shoot concentration **(a)**, root concentration **(b)** in the LC (*dashed bars*) and GA (*white bars*) parental controls, and in the F₂ (*gray bars*) from

the cross LC × GA; **(c)** and **(d)**, respectively. *T. caerulea* plants were grown hydroponically with 5 μmol CdSO₄ and 2 μmol ZnSO₄

toward high levels of root Zn concentration, beyond the upper limit of the GA distribution ($P < 0.001$). However, there was no significant transgression beyond the upper and lower limits of the parental distributions for any Cd-related trait. Cd and Zn concentrations were significantly ($P < 0.01$), but far from strictly, correlated both in shoot and in root (Fig. 3).

Genetic linkage map

Based on a screening for the highest polymorphism rates between both parental accessions, 19 *E/M* PCs were chosen. The size of the fragments generated ranged from 80 to 600 bp, with most of the fragments smaller than 500 bp. The average number of well-amplified bands varied from 44 to 85, with an average of 62.37 bands, and the average polymorphism rate

between LC and GA was 29.26%. Each selected PC generated on average 18.05 polymorphisms, ranging from 15 (E35M19) to 25 (E32M12). A total of 343 polymorphic AFLP bands were generated (157 LC-specific and 182 GA-specific) and four co-dominant AFLP markers were scored. In addition, nine CAPS and four indel markers were scored in the F₂ mapping population. Some of these co-dominant markers correspond to genes known to be involved in metal homeostasis, such as *TcZNT1* (Pence et al. 2000; Assunção et al. 2001), *TcZNT2* (Assunção et al. 2001) and the *T. caerulea* orthologue of *AtNRAMP4* (Thomine et al. 2000). An integrated map was constructed using both dominant and co-dominant markers. Of the total of 355 markers, 353 were assigned to seven LGs (LOD score 4.0), while only two dominant AFLP markers could not be fitted to any of these LGs. Markers that

Table 4 Median trait levels and 95% confidence intervals for LC, GA and F₂s

| Trait | Population/ accession | Median (in $\mu\text{mol g}^{-1}$ DW) | 95% confidence intervals (in $\mu\text{mol g}^{-1}$ DW) |
|-------|--------------------------|---|---|
| ZnS | LC | 4.24 | 3.87–4.65 |
| | GA | 2.72 | 2.51–2.94 |
| | F ₂ | 4.93 | 4.48–5.43 |
| ZnR | LC | 1.18 | 1.13–1.23 |
| | GA | 2.71 | 2.08–3.52 |
| | F ₂ | 2.29 | 2.12–2.46 |
| CdS | LC | 0.51 | 0.44–0.60 |
| | GA | 3.41 | 2.94–3.95 |
| | F ₂ | 3.80 | 3.40–4.24 |
| CdR | LC | 0.62 | 0.55–0.71 |
| | GA | 7.52 | 6.98–8.09 |
| | F ₂ | 3.95 | 3.57–4.37 |

ZnS zinc concentration in shoot, ZnR zinc concentration in root, CdS Cd concentration in shoot, CdR concentration in root

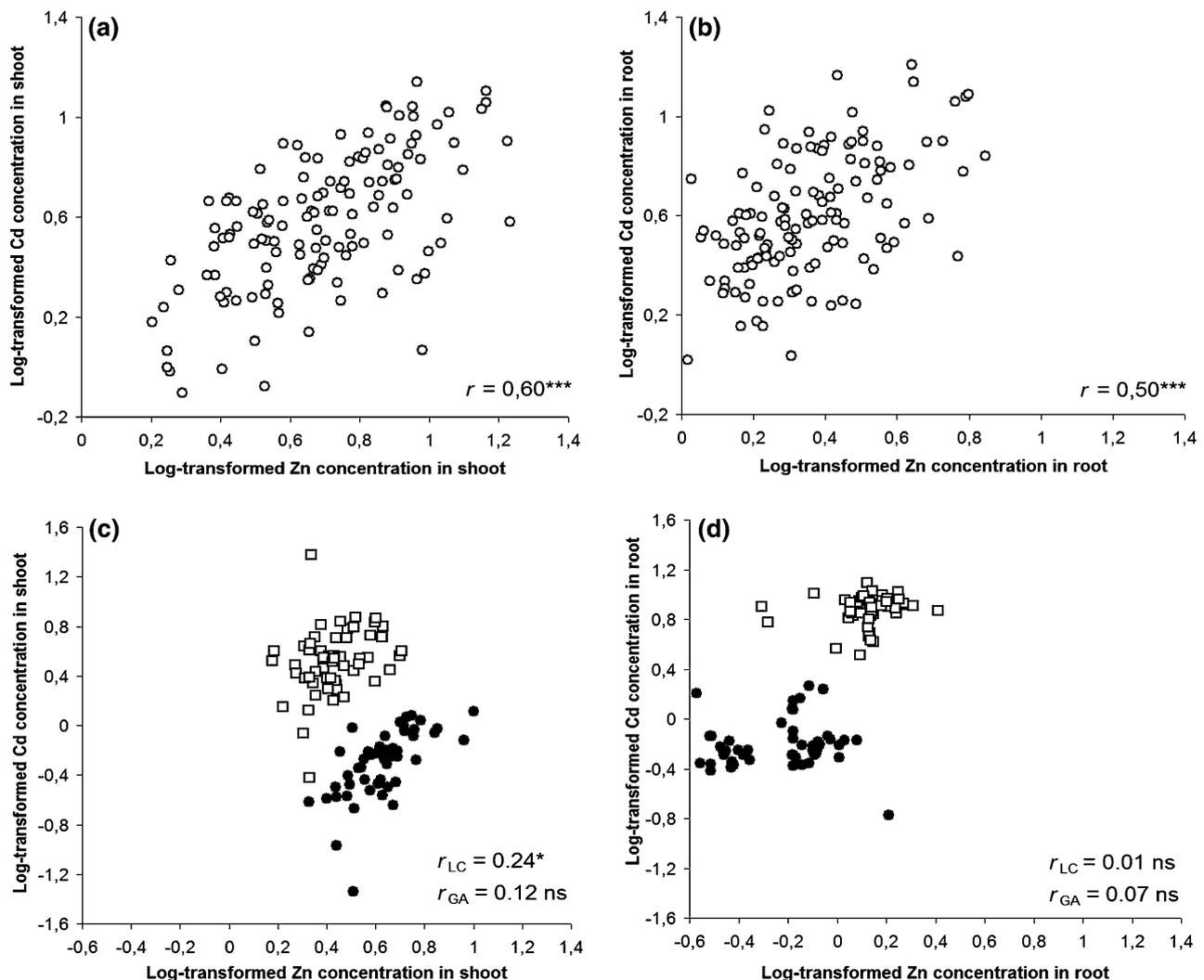


Fig. 3 Pearson correlation plot of log-transformed Cd/Zn concentrations in shoot (a) and in root (b) in the F₂ (open circles) and in the parental controls (LC, closed circles; GA, open squares) in

caused a large increase of the χ^2 value for the map of the individual chromosomes were removed from the data set. In addition when more than three AFLP markers co-segregated, the ones with the least missing data were maintained in the data set. Finally, 198 markers (181 AFLP and 17 co-dominant markers) were used to calculate an integrated map presented in Fig. 4. This map consisted of seven LGs, corresponding to the number of haploid chromosomes, each with total lengths between 56 and 92 cM. The seven LGs were labeled 1–7 in arbitrary order of grouping from JOIN-MAP 3.0. The total map length was 496 cM with an average of one marker per 2.5 cM and a largest genetic distance of 15 cM (LG5). The marker density per centi-Morgan was not significantly different between LGs. In each of the LGs, there was a cluster of linked markers, probably representing the centromeric regions. Most of the loci showed genotype ratios as expected for a

shoot (c) and in root (d). Significance of the correlation—ns not significant; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

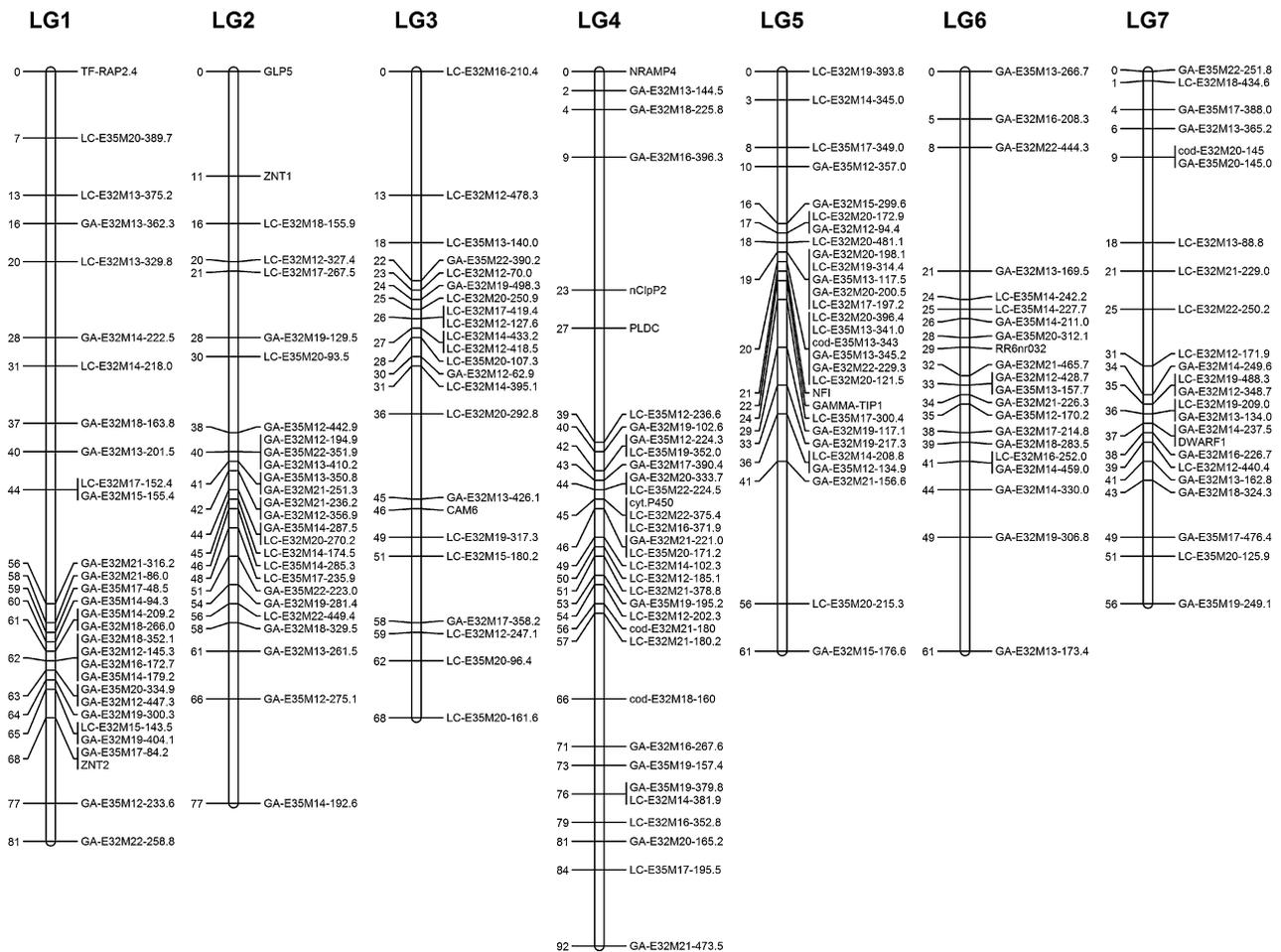


Fig. 4 The seven LGs of the integrated genetic map of the *T. caeruleus* LC × GA F₂ mapping population. The genetic distances are given in cM

segregating F₂ population (3:1 for dominant and 1:2:1 for co-dominant loci). For 22 markers this ratio was significantly distorted ($P < 0.05$), with a cluster of three loci mapping on LG4 between position 0 and 2 cM. For these loci the GA alleles were overrepresented. The remaining markers with a disturbed segregation were not linked to any other markers with distorted ratios. However, after correction of the significance levels according to the Bonferroni–Holm sequential method (Rice 1989), none of the markers showed a significant distortion from the expected Mendelian segregation.

QTL mapping

By combining molecular marker data with data on the Cd or Zn concentration in shoot and in root, QTL mapping was performed for the four traits analyzed. The location of all the significant QTLs and their 1-LOD and 2-LOD support interval are presented in Fig. 5. We found three significant QTLs for Zn concentration in shoot located on LG1 (ZnS1), LG4 (ZnS2) and LG7

(ZnS3), explaining 14.7, 9.6 and 18.1% of the phenotypic variance respectively, with a multiple QTL PVE (Percentage Variance Explained) of 44.5% (Table 5). Two QTLs were mapped for Zn concentration in root on LG1 (ZnR1) and LG6 (ZnR2). These explained 14.9 and 54.4% of the trait variance and with a multiple QTL PVE of 60.8%. ZnS1 and ZnR1 are co-locating. Only one QTL was found for Cd concentration in shoot (CdS), located on LG3. This QTL explained 23.8% of the variance. For Cd concentration in root we mapped two significant QTLs, on LG3 (CdR1) and on LG6 (CdR2), explaining 9.6 and 33.1% of the trait variance, respectively. The multiple QTL PVE was 55.5%. CdR1 and CdR2 are co-locating with CdS and ZnR2, respectively. Interestingly, both ZnR2 and CdR2 explain the bigger part of variances of Zn and Cd concentration in root (54.4 and 33.1%, respectively). Due to the lack of co-dominant markers with sufficient linkage to the identified QTLs, we were unable to establish any epistatic interactions between QTLs. Assuming that overdominance was absent (see Sect. "Discussion"), the

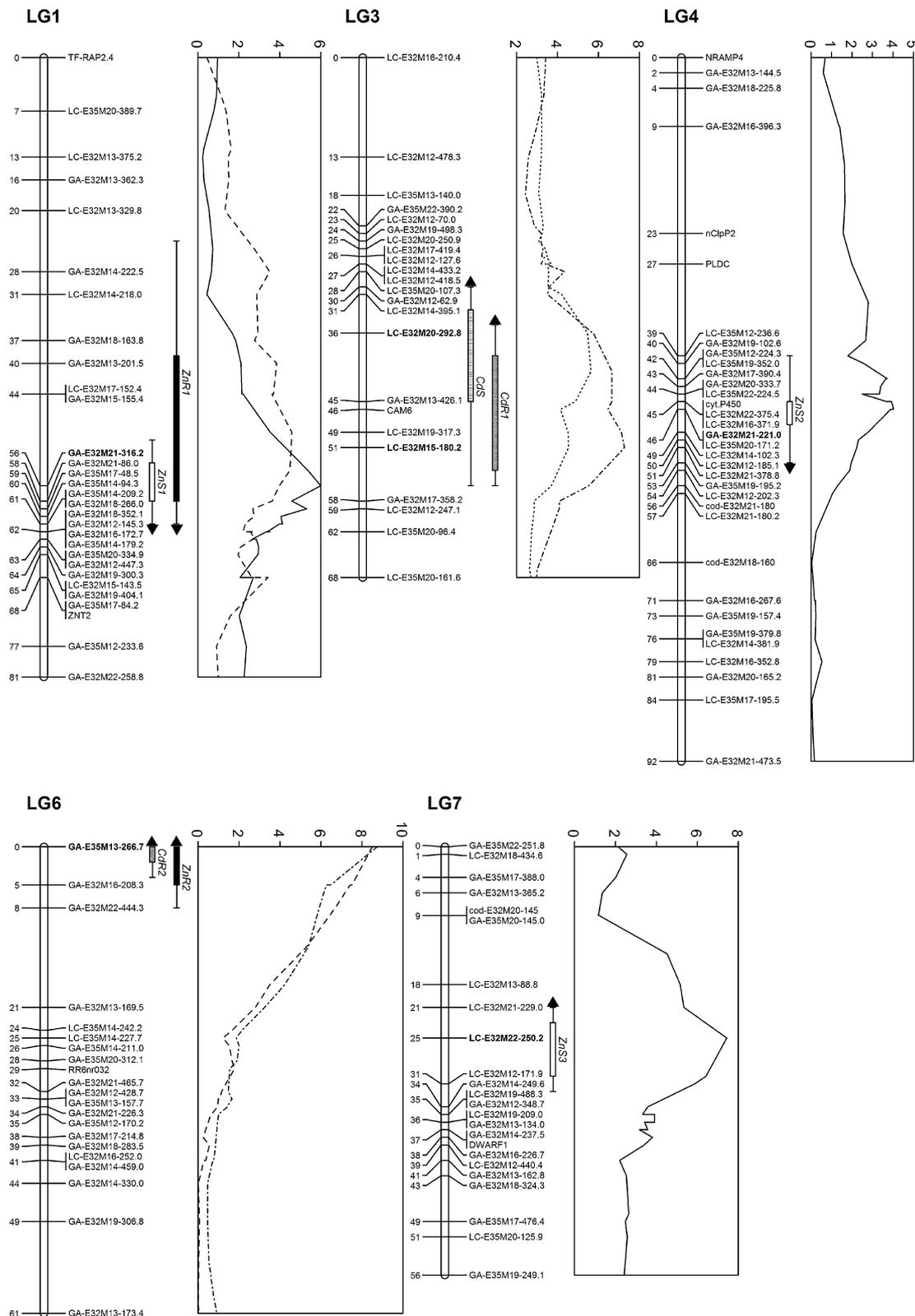


Fig. 5 Location of all the significant QTLs for Zn and Cd accumulation in shoot and in root, respectively, supported by their 1-LOD (bars) and 2-LOD (lines) support interval. The positions and origins of QTLs are indicated by upward (GA) and downward (LC) arrows. The LOD profile of each phenotype is shown

on the left of the LGs (Zn accumulation in shoot, dark line; Zn accumulation in root, dashed line; Cd accumulation in shoot, dotted line; Cd accumulation in root, dashed-dotted line). The most closely associated molecular marker loci are indicated in bold

Table 5 QTLs detected for Zn and Cd accumulation in shoot and root, respectively

| Trait | Description | GW LOD threshold | QTL | LG | MML | LOD | % PVE | Multiple QTL PVE (%) | Trait-enhancing allele |
|-------|---------------------------|------------------|------|----|-----------------|-----|-------|----------------------|------------------------|
| ZnS | Zn concentration in shoot | 3.4 | ZnS1 | 1 | GA-E32M21-316.2 | 6.0 | 14.7 | 44.5 | LC |
| | | | ZnS2 | 4 | GA-E32M21-221.0 | 4.0 | 9.6 | | LC |
| | | | ZnS3 | 7 | LC-E32M22-250.2 | 7.4 | 18.1 | | GA |
| ZnR | Zn concentration in root | 3.4 | ZnR1 | 3 | GA-E32M21-316.2 | 4.0 | 14.9 | 60.4 | LC |
| | | | ZnR2 | 6 | GA-E35M13-266.7 | 8.5 | 54.4 | | GA |
| CdS | Cd concentration in shoot | 3.3 | CdS | 3 | LC-E32M20-292.8 | 5.5 | 23.8 | 23.8 | GA |
| CdR | Cd concentration in root | 3.3 | CdR1 | 3 | LC-E32M15-180.2 | 7.3 | 9.6 | 55.5 | GA |
| | | | CdR2 | 6 | GA-E35M13-266.7 | 8.7 | 33.1 | | GA |

QTLs are named according to trait abbreviations and followed by a number to distinguish QTLs mapping affecting the same trait. Trait-enhancing allele = direction of the effect. GW LOD threshold indicates genome wide log-likelihood threshold determined by MapQTL5 permutation test ($n = 1,000$); multiple QTL PVE indicates total variance explained by all the QTLs detected

LG linkage group, MML most closely associated molecular marker loci, LOD maximum LOD score, % PVE percent phenotypic variation explained

trait-enhancing alleles for ZnS1, ZnS2 and ZnR1 appeared to originate from LC. For all the other QTLs, the trait-enhancing alleles originated from GA. In case of co-locating QTLs, the trait-enhancing alleles were always derived from the same parent.

Discussion

Segregation for Zn and Cd hyperaccumulation traits

We observed that GA accumulates more Cd ($7.52 \mu\text{mol g}^{-1}$ DW in GA, $0.62 \mu\text{mol g}^{-1}$ DW in LC) and Zn ($2.71 \mu\text{mol g}^{-1}$ DW in GA, $1.18 \mu\text{mol g}^{-1}$ DW in LC) in root and more Cd ($3.41 \mu\text{mol g}^{-1}$ DW in GA, $0.51 \mu\text{mol g}^{-1}$ DW in LC) in shoot than LC. The Zn concentration in shoot is higher in LC ($2.72 \mu\text{mol g}^{-1}$ DW in GA, $4.24 \mu\text{mol g}^{-1}$ DW in LC). Previously Zha et al. (2004) compared shoot Zn and Cd accumulation in GA and Prayon (PR). Zn or Cd accumulation in the PR accession is very similar to that in the LC accession (H. Schat, unpublished data). At a fivefold excess of Zn over Cd in the nutrient solution, they found similar shoot Zn accumulation in GA and PR, but at equimolar exposure PR accumulated over two times more Zn in the shoot than GA. This compares well with our observations. Overall, they found two to three times more shoot Cd accumulation in GA than in PR, which is less different than in our study (sevenfold). The discrepancy may be due to the fact that we used a 2.5-fold excess of Cd over Zn, whereas Zha et al. (2004) supplied Zn and Cd at equimolar concentrations, or supplied Zn at a fivefold higher concentration than Cd. Zha et al. (2004) proposed that two uptake systems with distinctive affinities for Cd and Zn are differently expressed in the two ecotypes. Their evidence suggested that there is an uptake

system with high affinity for Zn and lower affinity for Cd, which is predominant in PR, and one with high affinity for Cd but lower affinity for Zn, which is predominant or exclusive in GA. The latter system is likely to be responsible for Cd uptake in GA but can also contribute to Zn accumulation. When the Cd:Zn molar ratio in the nutrient solution is high (e.g., 2.5 in our experiment and 1.0 in Zha et al.'s experiment), LC and PR accumulate more Zn than GA.

The frequency distributions of the F_2 population for Cd or Zn concentration in shoot and root are not bimodal suggesting that Cd and Zn accumulation are quantitatively inherited traits. Similar findings were described in two other studies (Assunção et al. 2003b; Zha et al. 2004). The Zn and Cd accumulation in shoot and in root were significantly correlated, which was also found for the shoot concentrations in the study of Zha et al. (2004). In both cases the correlations are far from strict however, which could result from the contribution of mechanisms with different affinity patterns for both metals such as suggested by Zha et al. (2004).

We observed a significant transgression beyond the higher limit ($P < 0.001$) of the Zn parental distributions indicating that the trait-enhancing alleles from both parents at the different QTLs add both to the phenotype. Alternative explanations for transgression are overdominance and epistasis. Overdominance is unlikely both for Cd and Zn QTLs. In all cases of interecotypic *Thlaspi* crosses reported so far, F_1 plants showed accumulation rates that were intermediate between the parents, or dominance was directed toward low accumulation (Frérot et al. 2003, 2005; Assunção et al. 2006). Epistasis cannot be tested for because we were not able to distinguish between homozygote and heterozygote genotypes for the QTLs in a number of individuals sufficient for statistical anal-

ysis. In contrast, there was no significant transgression for Cd accumulation, which is again in agreement with Zha's results. This could be due to the absence of the high affinity Cd uptake system in LC, such as postulated by Zha et al. (2004).

The identification of QTLs for Zn accumulation in shoot and in root with trait-enhancing alleles originating from both parents and the significant transgression found for Zn accumulation is in agreement with the hypothesis of two systems for Zn accumulation as suggested by Zha et al. (2004). This does not seem to be the case for Cd accumulation where the trait-enhancing alleles originate from GA.

Clustering of markers and map reliability

While the overall distribution of AFLP markers appeared satisfactory, some of the *E/M* clustered in distinct regions with higher marker densities. This high degree of clustering of these markers at central regions is a notable feature in many species such as barley (Qi et al. 1998), tomato (Tanksley et al. 1992) or wheat (Chao et al. 1989) and *Arabidopsis* (Alonso-Blanco et al. 1998). Genetic maps are calculated from the recombination rates between loci as a result of chromosome crossovers at meiosis. So, assuming that markers are randomly distributed over the length of the chromosome, centromeric suppression of recombination is most likely the main reason for the clustering of markers (Tanksley et al. 1992; Frary et al. 1996). This enables us to locate putative centromeres of the *T. caerulescens* chromosomes.

The current map *T. caerulescens* compares very well to the genetic map made for another population using a common parent [e.g., Lellingen (LE) × LC in Assunção et al. 2006], for which a similar total map length and comparable size of the corresponding LGs was observed. A comparison based on AFLP markers with LC specific bands showed the same map position in both maps. The low number of common markers did not allow an integration of these two genetic maps. The small number of orthologous markers does not enable us to make further assumptions on the whole genome synteny and to establish possible chromosomal rearrangements as was done for another zinc hyperaccumulator *Arabidopsis halleri* (G. Willems and P. Saumitou-Laprade, personal communication) and the related *A. lyrata* (Kuittinen et al. 2004).

QTLs analysis

The QTL analysis confirmed that all the studied traits, except for Cd concentration in shoot, are controlled by

more than one gene, as expected from the phenotypic segregation patterns. In case of Cd concentration in shoot we found only one QTL explaining not more than 22.8%, suggesting that minor QTLs still remain to be detected. The multiple QTL PVE in this study ranged from 23.8 to 60.4% (Table 5). As we only had single plants for each F₂ genotype, we cannot calculate the heritability of the traits in this population.

We observed two cases of co-location of shoot and root concentration QTLs (Fig. 5): one for Zn on LG1 (ZnS1 and ZnR1) and another one for Cd on LG3 (CdS and CdR1). The trait-enhancing alleles for the ZnS1/ZnR1 originate from LC and those for CdS/CdR1 originate from GA. Apparently there are genetic determinants enhancing metal accumulation in both shoot and root. Other QTLs specifically apply to the accumulation in shoot or in root. The genes underlying these QTLs might be involved in uptake, but also sequestration or in plant-internal metal transport. Interestingly, there was one case of co-location of QTLs for Zn and Cd, ZnR2 and CdR2 on LG6 with both trait-enhancing alleles originating from GA. These QTLs explained most of the trait variances, both for Zn and Cd. This could be the result of a Zn/Cd transporter with relatively high affinity for Cd, such as proposed to be present in the GA parent (Zha et al. 2004). Such a transporter might explain the significant correlation between Zn and Cd accumulation. However, the other QTLs are metal-specific, suggesting the existence of more metal-specific transport systems. This would explain the relatively low degree of correlation of Zn and Cd accumulation.

The Zn transporter genes *ZNT1* (LG2) and *ZNT2* (Pence et al. 2000; Assunção et al. 2001) (LG1) or the iron deficiency induced transporter gene *NRAMP4* (Thomine et al. 2000) (LG4) have been described as putatively involved in metal (hyper)accumulation. Their location on the genetic map, however, did not coincide with the location of any QTL. It is possible that these genes do not segregate in this inter-accession cross because they are similarly expressed in different accessions (Assunção et al. 2001).

Assunção et al. (2006) mapped QTLs for Zn accumulation in a *T. caerulescens* F₃ cross between a plant from a non-metallicolous population (Lellingen, Luxemburg) and a plant from LC. The LE ecotype is characterized by a relatively high Zn accumulation, compared to LC. In their study no significant QTLs for Zn accumulation in shoot were found. However, there were two significant QTLs for Zn in root, one with the trait-enhancing allele derived from the LE parent and the other with the trait-enhancing allele from the LC parent, comparable to the results obtained in the pres-

ent study. Remarkably, the QTL with the trait-enhancing allele originating from LE, named LE3, co-segregated with a co-dominant marker (indel47/48), representing At5g21274 (calmodulin 6). In our study CdR1/CdS co-segregates with markers neighboring CAM6, which represents the same gene. It is not clear whether these genes underlying these QTLs are different or not. The second QTL for Zn accumulation in root found by Assunção et al. (2006) was mapped to chromosome 5 (LG5 in our study), where we did not find any QTL. This could mean that LC and GA have similarly effective alleles at this particular locus. Also, LE and LC might similar alleles at the loci underlying the QTLs ZnR1 (LG1) and ZnR2 (LG6). Taken together the results of both studies, it is evident that there are at least four loci determining the inter-accession variation in Zn accumulation in root. Dependent on the origin of the parents, either of these loci may or may not contribute to the segregation of the trait in inter-accession crosses.

Overall, this study has proven that AFLP markers are adequate for map construction and QTL analysis in *T. caerulescens*, which agrees with the findings of a recent study by Assunção et al. (2006). The present molecular map and the accompanying mapping population open new avenues for the identification of genes involved in Zn and Cd (hyper)accumulation, which will enable us to unravel the molecular genetics of these complex traits. EST-based markers which link apparently orthologous genes from *Arabidopsis* and *T. caerulescens* will be used for comparative mapping and colinearity analysis. Eventually this will facilitate the selection of candidate genes for *T. caerulescens* QTL based on DNA information from the corresponding genome region in *Arabidopsis*.

Acknowledgments We are very grateful to Diana Rigola (Lab of Genetics, Wageningen University, The Netherlands) for providing the *Thlaspi caerulescens* EST data prior to publication and to Fangjie Zhao (Agriculture and Environment Division, IACR-Rothamsted, UK) for providing us with seeds of the Ganges accession. We thank Maarten Koornneef for critically reading the manuscript and Petra van den Berg, Fien Meijer-Dekens and Richard Flinkers for their technical advice and assistance. This research was supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek-Genomics (NWO) Grant (050-10-166) and the European Union-PHYTAC project (QLRT-2001-00429).

References

- Alonso-Blanco C, Koornneef M (2000) Naturally occurring variation in *Arabidopsis*: an underexploited resource for plant genetics. *Trends Plant Sci* 5(1):22–29
- Alonso-Blanco C, Peeters AJM, Koornneef M, Lister C, Dean C, Van den Bosch N, Pot J, Kuiper MTR (1998) Development of an AFLP based linkage map of Ler, Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a Ler/Cvi recombinant inbred line population. *Plant J* 14:259–271
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–410
- Assunção AGL, Da Costa Martins P, De Folter S, Vooijs R, Schat H, Aarts MGM (2001) Elevated expression of metal transporter genes in three accessions of the metal hyperaccumulator *Thlaspi caerulescens*. *Plant Cell Environ* 24:217–226
- Assunção AGL, Ten Bookum WM, Nelissen HJM, Vooijs R, Schat H, Ernst WHO (2003a) A co-segregation analysis of zinc (Zn) accumulation and Zn tolerance in the Zn hyperaccumulator. *New Phytol* 159:383–390
- Assunção AGL, Schat H, Aarts MGM (2003b) *Thlaspi caerulescens*, an attractive model species to study heavy metal hyperaccumulation in plants. *New Phytol* 159:351–360
- Assunção AGL, Ten Bookum WM, Nelissen HJM, Vooijs R, Schat H, Ernst WHO (2003c) Differential metal-specific tolerance and accumulation patterns among *Thlaspi caerulescens* populations originating from different soil types. *New Phytol* 159:411–419
- Assunção AGL, Pieper B, Vromans J, Lindhout P, Aarts MGM, Schat H (2006) Construction of a genetic linkage map of *Thlaspi caerulescens* and quantitative trait loci analysis of zinc accumulation. *New Phytol* 170:21–32
- Chao S, Sharp PJ, Worland AJ, Warham EJ, Koebner RMD, Gale MD (1989) RFLP-based genetic maps of wheat homoeologous group 7 chromosomes. *Theor Appl Genet* 78:495–504
- Cobbett C (2003) Heavy metals and plants model systems and hyperaccumulators. *New Phytol* 159:289–293
- Escarré J, Lefebvre C, Gruber W, Leblanc M, Lepart J, Riviere Y, Delay B (2000) Zinc and cadmium hyperaccumulation by *Thlaspi caerulescens* from metalliferous and nonmetalliferous sites in the Mediterranean area: implications for phytoremediation. *New Phytol* 145:429–437
- Frary A, Presting GG, Tanksley SD (1996) Molecular mapping of the centromeres of tomato chromosomes 7 and 9. *Mol Genet Genomics* 250:295–304
- Frérot H, Petit C, Lefebvre C, Gruber W, Collin C, Escarré J (2003) Zinc and cadmium accumulation in controlled crosses between metallicolous and nonmetallicolous populations of *Thlaspi caerulescens* (Brassicaceae). *New Phytol* 157:643–648
- Frérot H, Lefebvre C, Petit C, Collin C, Dos Santos A, Escarré J (2005) Zinc tolerance and hyperaccumulation in F₁ and F₂ offspring from intra and interecotype crosses of *Thlaspi caerulescens*. *New Phytol* 165:111–119
- Koch M, Haubold B, Mitchell-Olds T (2001) Molecular systematics of the Brassicaceae: evidence from coding plastidic matK and nuclear Chs sequences. *Am J Bot* 88:534–544
- Kosambi D (1944) The estimation of map distances from the recombination values. *Ann Eugen* 12:172–175
- Kuittinen H, de Haan AA, Vogl C, Oikarinen S, Leppala J, Koch M, Mitchell-Olds T, Langley C, Savolainen O (2004) Comparing the linkage maps of the close relatives *Arabidopsis lyrata* and *Arabidopsis thaliana*. *Genetics* 168:1575–1584
- Lombi E, Zhao FJ, Dunham SJ, McGrath SP (2000) Cadmium accumulation in populations of *Thlaspi caerulescens* and *Thlaspi goesingense*. *New Phytol* 145:11–20
- Meerts P, van Isacker N (1997) Phytoremediation of metals, metalloids and radionuclides. *Adv Agron* 75:1–56

- van Ooijen JW (1999) LOD significant thresholds for QTL analysis in experimental populations of diploid species. *Heredity* 89:803–811
- van Ooijen JW (2004) MapQTL[®] 5, software for the mapping of quantitative trait loci in experimental populations. Kyazma B.V., Wageningen
- Pence NS, Larsen PB, Ebbs SD, Letham DLD, Lasat MM, Garvin DF, Eide D, Kochian LV (2000) The molecular physiology of heavy metal transport in the Zn/Cd hyperaccumulator *Thlaspi caerulescens*. *Proc Natl Acad Sci USA* 97:4956–4960
- Qi X, Lindhout P (1997) Development of AFLP markers in barley. *Mol Genet Genomics* 254:330–336
- Qi X, Stam P, Lindhout P (1998) Use of locus-specific AFLP markers to construct a high-density molecular map in barley. *Theor Appl Genet* 96:376–384
- Rice WR (1989) Analyzing tables of statistical tests. *Evolution* 43:223–225
- Rigola D, Fiers M, Vurro E, Aarts MGM (2006) The heavy metal hyperaccumulator *Thlaspi caerulescens* expresses many species-specific genes as identified by comparative EST analysis. *New Phytol* 170:753–766
- Schat H, Llugany M, Bernhard R (2000) Metal-specific patterns of tolerance, uptake, and transport of heavy metals in hyperaccumulating and non-hyperaccumulating metallophytes. In: Terry N, Banuelos G (eds) *Phytoremediation of contaminated soils and water*. CRC, Boca Raton, pp. 171–188
- Stam P (1993) Construction of integrated linkage maps by means of a new computer package: JOINMAP. *Plant J* 3:739–744
- Stam P, van Ooijen JW (1996) JOINMAP version 3.0: software for the calculation of genetic linkage maps. CPRO-DLO, Wageningen
- Tanksley SD, Ganal MW, Prince JP, de-Vicente MC, Bonierbale MW, Broun P, Fulton TM, Giovannoni JJ, Grandillo S, Martin GB, Messeguer R, Miller JC, Miller L, Paterson AH, Pineda O, Roder MS, Wing RA, Wu W, Young ND (1992) High-density molecular linkage maps of the tomato and potato genomes. *Genetics* 132:1141–1160
- Thomine S, Rongchen W, 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* 97:4991–4996
- Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. *Heredity* 93:77–78
- Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M (1995) AFLP[®]: a new technique for DNA fingerprinting. *Nucleic Acids Res* 23:4407–4414
- Zha HG, Jiang RF, Zhao FJ, Vooijs R, Schat H, Barker JHA, McGrath SP (2004) Co-segregation analysis of cadmium and zinc accumulation in *Thlaspi caerulescens* interecotypic crosses. *New Phytol* 163:299–312
- Zhao FJ, Hamon RE, Lombi E, McLaughlin MJ, McGrath SP (2002) Characteristics of cadmium uptake in two contrasting ecotypes of the hyperaccumulator *Thlaspi caerulescens*. *J Exp Bot* 53:535–543