MOLECULAR ECOLOGY

Molecular Ecology (2009) **18**, 3227–3239 doi: 10.1111/j.1365-294X.2009.04261.x

Adaptive differences in gene expression associated with heavy metal tolerance in the soil arthropod *Orchesella cincta*

DICK ROELOFS,* THIERRY K. S. JANSSENS,* MARTIJN J. T. N. TIMMERMANS,* BENJAMIN NOTA,* JANINE MARIËN,* ZOLTÁN BOCHDANOVITS,† BAUKE YLSTRA‡ and NICO M. VAN STRAALEN*

**Institute of Ecological Science, VU University Amsterdam, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands, †Department of Clinical Genetics, Section Medical Genomics, VU Medical Center, Van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands, ‡Microarray Facility CCA, VU Medical Center, De Boelelaan 1117, 1081 HV Amsterdam, The Netherlands*

Abstract

Field-selected tolerance to heavy metals has been reported for *Orchesella cincta* (Arthropoda: Collembola) populations occurring at metal-contaminated mining sites. This tolerance correlated with heritable increase in metal excretion efficiency, less pronounced cadmium (Cd)-induced growth reduction and overexpression of the metallothionein gene. We applied transcriptomics to determine differential gene expression caused by this abiotic stress in reference and Cd-tolerant populations. Many cDNAs responded to Cd exposure in the reference population. Significantly fewer clones were Cd responsive in tolerant animals. Analysis of variance revealed transcripts that interact between Cd exposure and population. Hierarchical cluster analysis of these clones identified two major groups. The first one contained cDNAs that were up-regulated by Cd in the reference culture but non-responsive or down-regulated in tolerant animals. This cluster was also characterized by elevated constitutive expression in the tolerant population. Gene ontology analysis revealed that these cDNAs were involved in structural integrity of the cuticle, anti-microbial defence, calcium channel-blocking, sulphur assimilation and chromatin remodelling. The second group consisted of cDNAs down-regulated in reference animals but not responding or slightly up-regulated in tolerant animals. Their functions involved carbohydrate metabolic processes, Ca^{2+}-dependent stress signalling, redox state, proteolysis and digestion. The reference population showed a strong signature of stress-induced genome-wide perturbation of gene expression, whereas the tolerant animals maintained normal gene expression upon Cd exposure. We confirmed the micro-evolutionary processes occurring in soil arthropod populations and suggest a major contribution of gene regulation to the evolution of a stress-adapted phenotype.

Keywords: adaptation, cadmium, collembola, micro-evolution, suppression subtractive hybridization, transcription profiling

Received 6 June 2008; revision received 7 April 2009; accepted 6 May 2009

Introduction

Before direct evidence was available, biologists already argued that gene expression is a major driver of evolution (Waddington 1939; Jacob & Monod 1961). With the development of new genetic tools to study transcription regulation, it is now possible to test whether regulatory mutations contribute to heritable changes in morphology, physiology and behaviour (Wray 2007). Recent studies indeed show that selection for altered transcriptional regulation can be a powerful mechanism for

Correspondence: Dick Roelofs, Fax: +31 20 5987123; E-mail: dick.roelofs@falw.vu.nl

micro-evolution (Schulte *et al.* 2000; Daborn *et al.* 2002; Gompel *et al.* 2005; Rockman *et al.* 2005; Shapiro *et al.* 2006; Whitehead & Crawford 2006). Adaptive transcriptional responses have also been implicated in the evolution of tolerance to natural (Knight *et al.* 2006) and anthropogenic (Fisher & Oleksiak 2007) stressors in the environment. Earlier, we presented data suggesting that evolution of metal tolerance in the soil-living invertebrate *Orchesella cincta* (Arthropoda: Collembola, springtails) involves altered transcriptional regulation. A tolerant population of springtails showed increased expression of the metal detoxifying gene metallothionein (*mt*) compared with a reference population (Sterenborg 2003; Timmermans *et al.* 2005; Roelofs *et al.* 2007). Moreover, in a reference population, cadmium (Cd)-induced *mt* overexpression contained a significant heritable component (Roelofs *et al.* 2006), suggesting that selection can act on elevated *mt* function in springtails at sites where Cd levels are a potential threat to their fitness (Timmermans *et al.* 2005; Janssens *et al.* 2007, 2008). The increase in transcriptional activation of *mt* may account for enhanced excretion efficiency that was observed in a tolerant population (Posthuma *et al.* 1993). However, it cannot explain other aspects of the metal-tolerant phenotype, such as earlier reproduction and the less outspoken growth reduction upon Cd exposure that was reported by Posthuma *et al.* (1993).

Differential gene expression patterns due to adaptation to heavy metals are poorly understood. *Orchesella cincta* is a soil-dwelling arthropod that has both sensitive as well as metal-tolerant populations and is thus an appropriate model to address the effect of heavy metals on the stress response system as well as identifying genes that are possibly involved in metal tolerance. However, being a non-genomic model organism, DNA sequence information is lacking for *O. cincta*. In such a case, it is efficient to construct cDNA libraries enriched for genes that potentially respond to the exposure under investigation, such as suppression subtractive hybridization (SSH; Diatchenko *et al.* 1996). For instance, Gracey *et al.* (2001) used SSH in combination with microarray analysis to investigate tissue-specific and temporal changes in gene expression levels in the non-genomic model organism *Gillichthys mirabilis* (long-jaw mudsucker) exhibiting tolerance to hypoxia. This study illustrates that it is possible to reconstruct a considerable fraction of the expressed genome even in organisms for which little genomic information existed before.

Recently, we (Roelofs *et al.* 2007) employed the SSH method to screen for genes that are differentially expressed in *O. cincta* upon Cd exposure. We used quantitative PCR to compare transcription profiles of the reference culture with those of Cd-tolerant animals originating from two locations near an abandoned lead/zinc mine near Plombières (Belgium). More insight was obtained into: (i) the mechanism of toxicity by identification of additional responsive genes and (ii) the adaptive mechanisms of Cd tolerance. The most intriguing result was the general lack of a stress response in the two tolerant cultures. Most of the assayed genes, differential between populations, showed no significant up-regulation upon Cd exposure in the two Plombières populations. In fact, three genes showed significant down-regulation. We suggested a mechanism of adaptation through constitutive detoxification of Cd preventing the increase in free metal ion concentrations which would harm cellular homeostasis and integrity. It may involve constitutively high levels of *mt* mRNA present in tolerant animals, providing sufficient protein to effectively chelate any free Cd (Sterenborg 2003; Timmermans *et al.* 2005). Five additional genes showed a similar pattern (low constitutive expression in reference populations and high constitutive expression in tolerant populations). Two of these genes, being a MAPK phosphatase and a protein tyrosine kinase homologue, were suggested to be involved in modulating metal transcription factor MTF-1 (which is the major transcription factor regulating *mt* transcription, Saydam *et al.* 2001, 2002), or some of its co-factors (Jiang *et al.* 2004), although the precise mechanism is far from clear. It was suggested that selection on constitutive regulation of stress response genes was important to develop metal tolerance (Stürzenbaum *et al.* 2001) in a polluted ecosystem. This mechanism was also proposed in plants living on metal-contaminated soils and seems therefore of broad relevance (Roelofs *et al.* 2008). To study the relevance of transcriptional regulation in stress tolerance, we applied transcriptomics to determine differential gene expression caused by an important environmental stressor (Cd) in a reference and Cd-tolerant population. Our results suggest that the Cd-induced stress response is dramatically diminished in tolerant animals. It suggests a major contribution of gene regulation to the evolution of a stress-adapted phenotype.

Materials and methods

Origin and exposure of animals

Orchesella cincta (Collembola) from the laboratory population at the Department of Animal Ecology, Vrije Universiteit Amsterdam, were taken as the reference group. This population originated from a pine forest from the reference area (Roggebotzand, the Netherlands, latitude 52°34′17″N, longitude 5°47′56″E) that contains on average < 0.5 μg Cd/gram litter and humus (Van Straalen

et al. 1987; Janssens *et al.* 2007). Tolerant animals were collected from randomly selected litter samples at two areas of the abandoned but still heavily polluted lead/zinc mining site of Plombières (Belgium, latitude 50°44′03″N, longitude 5°58′02″E): the locations contain an average Cd concentration between 10 and 52 μg/g soil (Van Straalen *et al.* 1987; Lock *et al.* 2003; Sterenborg 2003). To diminish putative environmental effects from the field, the animals were reared in a climate room (20 °C, 75% humidity, L:D 12 h:12 h) in PVC jars on a moist plaster of Paris layer feeding on algae present on twigs for at least three generations before experimental treatment.

Animals were exposed to Cd (nominal concentration of 112.4 μg Cd × g^{-1} dry algae) for 2 days, directly after moulting, via the food according to Roelofs *et al.* (2006). This exposure concentration is two times the no observed effect concentration (NOEC) for reproduction in chronically exposed animals (Van Straalen *et al.* 1989). The above mentioned Cd concentration in short-term exposure experiments results in internal body concentrations that are comparable with internal body concentrations measured in animals collected from the Plombières sites (Sterenborg 2003). Total RNA was extracted from eight pools of 12 animals per treatment using SV Total RNA Isolation system (Promega) and quantified on a Nanodrop ND-1000 Spectrophotometer (Nanodrop Technologies). Total RNA was visualized on a 1.5% agarose gel with ethidiumbromide to verify its integrity.

Microarrays, labelling and scanning

We amplified cDNAs from our previously obtained Cd-exposed SSH library (Roelofs *et al.* 2007) together with clones derived from a cDNA library of gut tissues from Cd-exposed animals (T.K.S. Janssens, P. Kille, S. Stürzenbaum and D. Roelofs, unpublished) with amine-linked primers. Amine-tagged cDNAs were spotted onto Codelink-activated slides (GE Healthcare). Each cDNA was printed twice on one array and two arrays were printed on one slide, so that each cDNA was four times technically replicated on a slide. Pins of the printer resampled the cDNAs before printing the replicate array, so that each slide contained two separately printed arrays. The eight biological replicates per treatment were hybridized to 16 different arrays.

A direct comparison was applied of all treatments and populations in a closed loop experimental design (Fig. 1). About 20 μg of total RNA was labelled by the incorporation of 5-(3-aminoallyl)-dUTP (Ambion) during cDNA synthesis with superscript II reverse transcriptase (Invitrogen) and an oligo(d)T_{15} probe. Subsequently, amine-labelled cDNA was treated with Fluorolink Cy-3 or Cy-5 monofunctional dyes (Amersham).

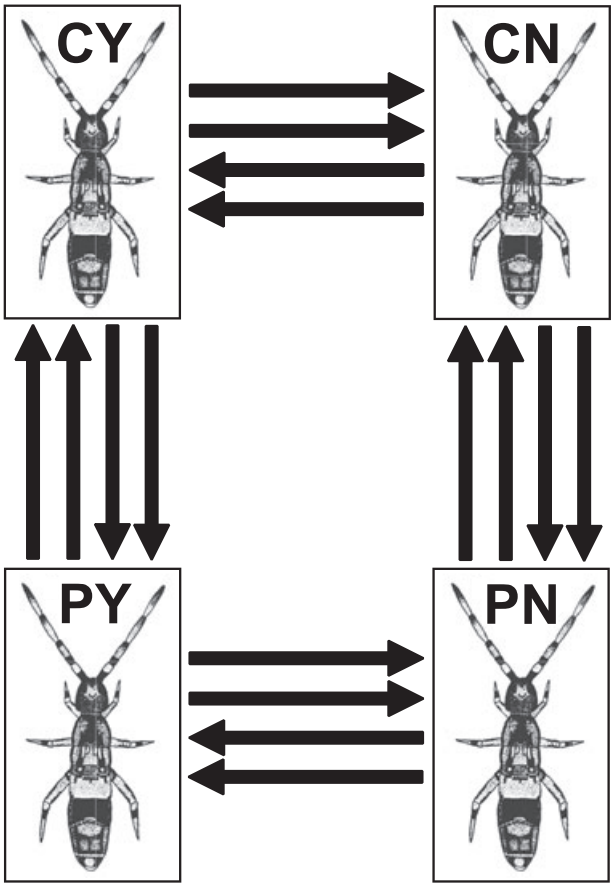


Fig. 1 Experimental design. C: Labculture (reference population) and P: Plombières culture (cadmium tolerant). Y: 112.4 μg Cd × g^{-1} dry food, N: clean food. Each arrow depicts an array hybridization. Arrow base: Cy-3, arrow point: Cy-5.

After removal of unincorporated fluorescent dyes, with a PCR clean-up column (Qiagen), labelled cDNA was eluted in nuclease free water and stored at −80°C until hybridization.

Hybridization of labelled cDNAs was performed according to Bergman *et al.* (2005) with a modified hybridization buffer (0.14 g/mL dextrane sulphate, 75% formamide, 3XSSC, 60 μg t-RNA and 24 μg human cot-1 DNA; Amersham) at 37 °C in a HybArray 12 (Perkin-Elmer) for 14 h. Wash steps were also performed in the HybArray 12 and included one step with 50% formamide/2XSSC, one step with phosphate buffer followed by three steps with SSC (0.2X, 0.1X and 0.01X).

Labelled cDNAs bound to microarray slides were quantified with a DNA microarray scanner (Agilent). Images were analysed using Bluefuse microarray analysis software (version 3.2; BlueGnome) in order to calculate background corrected $\log_2$ gene expression ratios of Cy3/Cy5 signals. From the 1920 cDNAs, 262 were omitted from further analysis because of aberrant spot morphology.

Statistics and bioinformatics

^{2}log-transformed Cy3/Cy5 ratios were analysed using R/Linear Models for Microarray Data (LIMMA; Wettenhall & Smyth 2004). First, data were smoothed by applying the global lowess correction method. Subsequently, cDNAs were obtained that were significantly affected by Cd exposure ($P < 0.05$). Significance values were adjusted for multiple testing using the Benjamini Hochberg algorithm (Benjamini & Hochberg 1995). ANOVA analysis was performed using the Microarray Analysis of Variance package in the R environment (R/MAANOVA; Wu & Churchill 2006). First, a lowess correction was conducted as described above to normalize for dye effect. A mixed-model ANOVA was fitted to the data (average intensities from the four replicated probes for each cDNA) and *F*-ratio test statistics were calculated for the two main effects and the interaction term with no differential expression as null hypothesis. The model is described by the following formula: $y_{ijkl} = \mu + P_i + C_j + (P \times C)_{ij} + D_k + A_l + \varepsilon_{ijkl}$, where y_{ijkl} is the signal from the sample derived from the *i*th population treated with the *j*th Cd treatment labelled with the *k*th dye and hybridized to the randomly assigned *l*th array. The parameter μ is the overall mean, P_i is the population effect (tolerant or sensitive), C_i is the Cd treatment effect (yes or no) and $(P \times C)_i$ is the population × Cd treatment interaction, D_i the dye effect (Cy3 or Cy5), A_i is the random effect of the array to which the sample has been assigned to and ε_{ijkl} is the stochastic error. Significance testing was performed using one-step FWER-adjusted empirical *P*-values based on 1000 permutations of the data. The MAANOVA input data-, design- and script files are available as Supporting Information files (Appendix S1: MaanovascriptCadOC.txt, MaanovainputCadOC.txt, MaanovaDesignCadOC.txt). The resulting tabulated *P*-values were adjusted for a false discovery rate of 0.1%.

Significantly differentially expressed cDNAs were hierarchically clustered using complete linkage within the Pearson uncentred distance method and visualized by TIGR Multi Experiment Viewer software (Saeed *et al.* 2003). The significant clones were sequenced and analysed using the Partigene (Parkinson *et al.* 2004) Perl scripts. In summary, the different sequences were assembled into groups of unique gene fragments (unigenes). Subsequently, the sequences within each cluster were assembled using Phrap (Gordon 2008). The obtained contig-sequences were then subjected to a BLAST (McGinnis & Madden 2004) search (BLASTX) against the non-redundant database (GenBank Release 163). In addition, all gene clusters were assigned Gene ontology (GO) terms using the Annot8r Perl script that used BLAST searches against the UniProt database (Schmid & Blaxter 2008). All EST sequences are submitted to GenBank (accession numbers) and are additionally available on Collembase (Timmermans *et al.* 2007; http://www.collembase.org). The function getGOGraph of the R package GOSIM (Frohlich *et al.* 2007) was applied to generate a Gene ontology hierarchy for several clusters.

The microarray data, as well as EST accession numbers, discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (Edgar *et al.* 2002) and are accessible through GEO Series accession number GSE14076 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE14076).

Results

Cadmium-dependent transcriptional response in each population

We first assessed the effect of Cd in each population separately. Figure 2 shows the results for each population, graphically represented by a volcano plot (Jin *et al.* 2001). Gene expression ratios from the reference culture reflect a severe transcriptional response to Cd: 1091 cDNA clones showed a significant (adjusted $P < 0.05$) response to the exposure (Fig. 2a). The SSH procedure enriches for up-regulated clones upon treatment. Indeed, approximately two-thirds of the significant cDNAs were up-regulated in reference population and the range of up-regulation was broad, reaching a maximum of 270-fold up-regulation of clone Oc_SSHCd_07A04 (Table S1). The remaining down-regulated genes showed a more narrow range with 19-fold down-regulation of clone Oc_SSHCd_12F8 (Table S1) as the most down-regulated clone. A total of 306 clones showed significant regulation higher than a $-^{10}\log P > 10$.

A smaller fraction of cDNAs (791) was significantly affected by Cd exposure in the tolerant population from Plombières (Fig. 2b). Moreover, the responses were less significant in tolerant animals when compared with reference animals: only 63 clones showed a response with significance of $-^{10}\log P > 10$. Finally, the amplitude of regulation was much less in tolerant animals, ranging from sixfold down-regulation (Oc_SSHCd_19A3, Table S1) up to 28-fold up-regulation (Oc_SSHCd_19E7, Table S1). Reference animals obviously showed a much more profound regulatory change upon Cd exposure, both in maximum fold change, as well as in significance of the transcriptional response, while tolerant animals from Plombières were less affected by Cd.

Direct comparison between Cd-sensitive and -tolerant populations

We hybridized not only Cd-exposed with non-exposed animals for each population but also directly compared

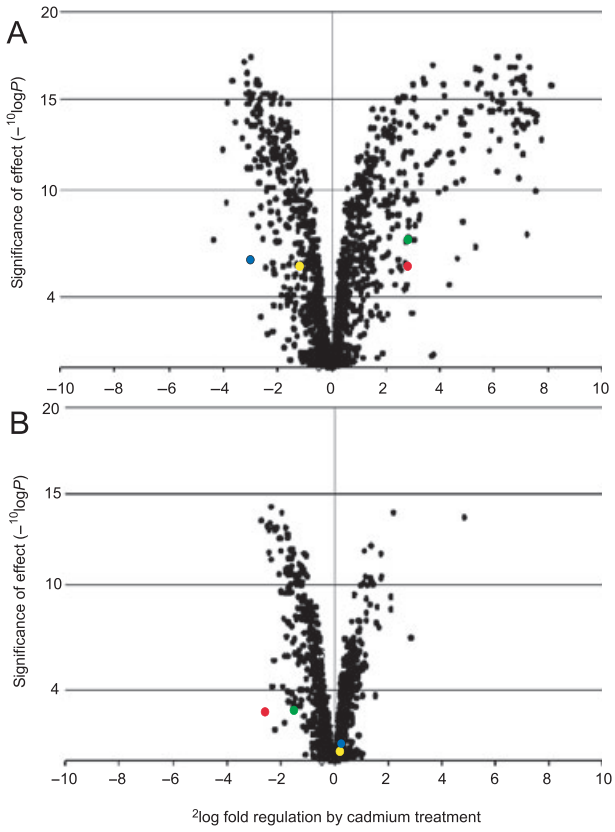


Fig. 2 Volcano plot of cadmium exposure. (A) Differential gene expression of cadmium-exposed reference culture. (B) Differential gene expression of cadmium-exposed tolerant culture from Plombières. *X*-axis: ^{2}log gene expression ratios (cadmium-exposed vs. non-exposed animals), *Y*-axis: level of significance ($-^{10}$log *P*) of differential expression. Coloured dots represent relevant genes showing an interaction between Cd treatment and population: Diapausin Occ00249 (red), Hsp40 Occ00117 (green), phosphoinositide-specific phospholipase C Occ00495 (yellow), Trypsin Occ00335 (blue).

non-exposed tolerant animals with non-exposed reference animals and Cd-exposed tolerant animals with Cd-exposed reference animals in a loop design (Fig. 1). Thus, we are capable of not only determining which genes are affected by Cd exposure or genetic background (population) but also identifying genes that respond differently to Cd exposure depending on the genetic origin of the animals (interaction between Cd and population). ANOVA analysis was performed on the four technically replicated spots of each cDNA as well as the collapsed mean value for each cDNA. Both analyses yielded equal numbers of significant cDNAs (data not shown). A significant dye effect was observed in two clones (Oc_SSHCd_12A04 and Oc_SSHCd_05A08, Table S1), which were omitted from further analysis. When *P*-values were adjusted for false discovery rate of 0.1%, 224 clones were significantly affected by Cd, 221 clones were significantly affected by population and 606 clones showed a significant interaction between Cd and population (Table S2).

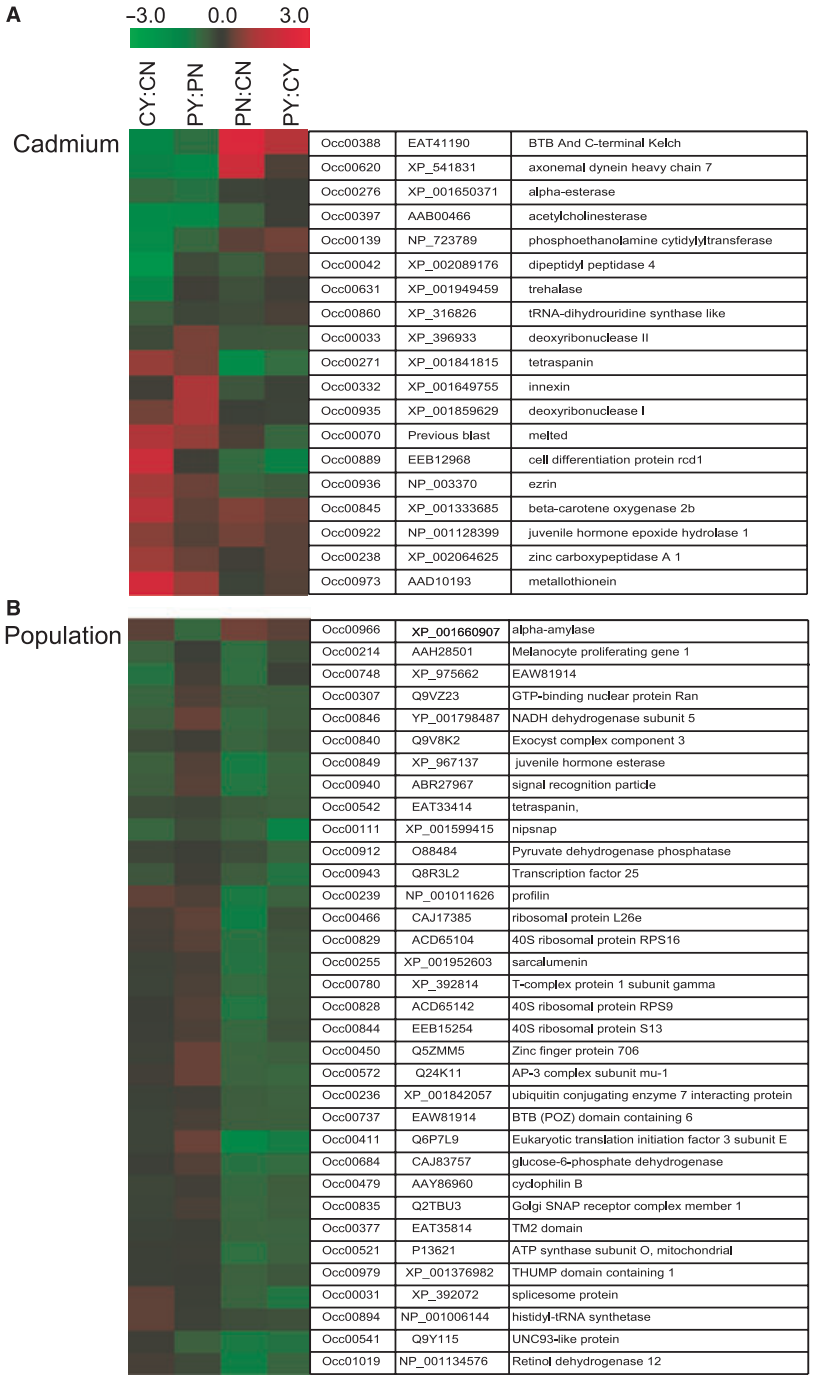
Gene expression affected by Cd treatment or population

The Cd-responding clones clustered in two groups: down-regulated by Cd and up-regulated by Cd (Fig. 3). The 224 significant clones assembled in 90 unique genes (unigenes); about two-thirds of the clones were down-regulated and the remaining clones were up-regulated due to Cd exposure. We retrieved BLAST hits for only 19 unigenes and Gene ontology analysis revealed that 14 biological processes were affected by Cd. One unigene, involved in phospholipid and fatty acid biosynthesis (Occ00139 and Occ00165), was down-regulated upon Cd exposure. Moreover, the stress biomarker acetylcholine esterase (Occ00397) was down-regulated upon Cd exposure. Innexin (Occ00332), belonging to a gene family involved in gap junction formation between cells, was up-regulated. A *melted* orthologue, with a pleckstrin domain (Occ00070), and a metallothionein (Occ00973) were up-regulated as well.

The 221 clones that showed a population-specific effect were assembled in 121 unigenes, of which 37 showed homology to sequences in the public databases (Fig. 3, Supporting Information file 2). Gene ontology analysis revealed 28 biological processes, of which six comprised general gene expression processes like RNA splicing, translation and protein folding. Furthermore, one unigene (Occ00572) contains a transmembrane gene of which the product is residential of lysosomes and is implicated in phagocytosis. Additionally, we identified a unigene that is involved in actin binding and the formation of stress fibres (profilin, Occ00239). Figure 3b shows that all these clones revealed a lower expression in the animals from the tolerant population compared with reference animals except one clone (unigene Occ00966, containing an α-amylase), which exhibited constitutively elevated levels in the tolerant population.

Interaction between Cd treatment and population

From the 606 interacting clones, 231 were up-regulated by Cd in the reference population and down-regulated or unaffected in the tolerant population, whereas the basal expression was generally higher in tolerant animals than in reference (Fig. S1). These 231 cDNAs assemble into 138 unigenes, of which 34 showed homology to known gene sequences in public databases. The most important genes within specific functional classes are summarized in Fig. 4 (cluster A). Eight unigenes showed high homology to cuticle



Occ00388	EAT41190	BTB And C-terminal Kelch
Occ00620	XP_541831	axonemal dynein heavy chain 7
Occ00276	XP_001650371	alpha-esterase
Occ00397	AAB00466	acetylcholinesterase
Occ00139	NP_723789	phosphoethanolamine cytidylyltransferase
Occ00042	XP_002089176	dipeptidyl peptidase 4
Occ00631	XP_001949459	trehalase
Occ00860	XP_316826	tRNA-dihydrouridine synthase like
Occ00033	XP_396933	deoxyribonuclease II
Occ00271	XP_001841815	tetraspanin
Occ00332	XP_001649755	innexin
Occ00935	XP_001859629	deoxyribonuclease I
Occ00070	Previous blast	melted
Occ00889	EEB12968	cell differentiation protein rcd1
Occ00936	NP_003370	ezrin
Occ00845	XP_001333685	beta-carotene oxygenase 2b
Occ00922	NP_001128399	juvenile hormone epoxide hydrolase 1
Occ00238	XP_002064625	zinc carboxypeptidase A 1
Occ00973	AAD10193	metallothionein
Occ00966	XP_001660907	alpha-amylase
Occ00214	AAH28501	Melanocyte proliferating gene 1
Occ00748	XP_975662	EAW81914
Occ00307	Q9VZ23	GTP-binding nuclear protein Ran
Occ00846	YP_001798487	NADH dehydrogenase subunit 5
Occ00840	Q9V8K2	Exocyst complex component 3
Occ00849	XP_967137	juvenile hormone esterase
Occ00940	ABR27967	signal recognition particle
Occ00542	EAT33414	tetraspanin,
Occ00111	XP_001599415	nipsnap
Occ00912	O88484	Pyruvate dehydrogenase phosphatase
Occ00943	Q8R3L2	Transcription factor 25
Occ00239	NP_001011626	profilin
Occ00466	CAJ17385	ribosomal protein L26e
Occ00829	ACD65104	40S ribosomal protein RPS16
Occ00255	XP_001952603	sarcalumenin
Occ00780	XP_392814	T-complex protein 1 subunit gamma
Occ00828	ACD65142	40S ribosomal protein RPS9
Occ00844	EEB15254	40S ribosomal protein S13
Occ00450	Q5ZMM5	Zinc finger protein 706
Occ00572	Q24K11	AP-3 complex subunit mu-1
Occ00236	XP_001842057	ubiquitin conjugating enzyme 7 interacting protein
Occ00737	EAW81914	BTB (POZ) domain containing 6
Occ00411	Q6P7L9	Eukaryotic translation initiation factor 3 subunit E
Occ00684	CAJ83757	glucose-6-phosphate dehydrogenase
Occ00479	AAY86960	cyclophilin B
Occ00835	Q2TBU3	Golgi SNAP receptor complex member 1
Occ00377	EAT35814	TM2 domain
Occ00521	P13621	ATP synthase subunit O, mitochondrial
Occ00979	XP_001376982	THUMP domain containing 1
Occ00031	XP_392072	splicesome protein
Occ00894	NP_001006144	histidyl-tRNA synthetase
Occ00541	Q9Y115	UNC93-like protein
Occ01019	NP_001134576	Retinol dehydrogenase 12

Fig. 3 Clusters of GO-annotated clones that show significant cadmium or population effect (0.1% FDR corrected) arranged according to functional group. (A) Cadmium effect; (B) population effect. Red: up-regulation, green down-regulation. Labels are cluster IDs (see Table S3). Each lane represents the mean ratios of four biological replicates of hybridizations. CY:CN, ratio cadmium-treated laboratory culture (CY) versus untreated laboratory culture (CN). Ratio PY:PN, cadmium-treated tolerant Plombières culture (PY) versus untreated tolerant Plombières culture (PN). PN:CN, ratio untreated tolerant Plombières culture versus untreated laboratory culture. PY:CY, ratio cadmium-treated tolerant Plombières culture versus cadmium-treated laboratory culture.

protein and GO analysis assigned these to the biological process 'structural constituent of the cuticle'. In addition, an endochitinase (Occ00140) was identified, suggesting that exoskeleton maintenance is severely affected by Cd. Two unigenes were identified as putative diapausin (Occ00233 and Occ00249) and one unigene was highly homologous to the chitin-binding peritrophin (Occ00173). Both diapausin and

perithrophin are involved in anti-microbial defence in the midgut (Kouno *et al.* 2007). Furthermore, the genes coding for heat shock protein 40 (Occ0011, involved in protein refolding) and sulphotransferase (Occ00895, involved in sulphate assimilation) were induced by Cd in the reference population. Moreover, both unigenes Occ00310 (homologous to the transcriptional regulator protein HCNGP) and Occ00327 (Histone acetyltransferase p300) are involved in chromatin remodelling by histone acetylation. HCNGP acts as a repressor by attracting histone deacetylase, while p300 is a transcriptional activator via its intrinsic histone acetylase activity. They are both up-regulated by Cd in the reference population.

The remaining 375 clones (Fig. S1) that show a population-specific response to Cd are characterized as down-regulated in the reference population upon Cd exposure and exhibit generally lower constitutive expression in tolerant animals (Fig. 4, cluster B). They assembled into 241 unigenes, of which 99 showed significant BLAST hits to known gene sequences. Gene ontology analysis revealed that diverse biological processes are down-regulated in reference animals but are less responsive in tolerant animals upon Cd exposure. Overall, general metabolic and translation processes seem to be affected; at least 16 unigenes were identified to be involved in energy metabolism and seven in translation (Fig. 4, Cluster B). Furthermore, a signature of digestive inhibition due to Cd toxicity was identified in reference animals: seven unigenes showed homology to proteases. Again these genes were unaffected by Cd in tolerant animals, when compared with reference populations. Several chitin-binding proteins and another peritrophin (Occ00075) anti-microbial peptide were less down-regulated by Cd in tolerant animals compared with the reference population (Table S3).

Surprisingly, genes involved in a possible stress-activated protein kinase signalling pathway (SAPK, Fig. 5), involving phosphatidylinositol and calcium signalling (e.g. phosphatidylinositol transfer protein lipid-binding domain Occ00267, phosphoinositide-specific phospholipase C Occ00495, cAMP-dependent protein kinase Occ00705, protein kinase C mu Occ00149 and protein tyrosine phosphatase Occ00205) were down-regulated in reference animals but were non-responsive or slightly up-regulated in tolerant animals. Moreover, some more specific targets were transcriptionally affected in similar ways (for instance, glutathione-S-transferase, Occ00622). Finally, the *O. cincta* metallothionein cluster was identified in this group of interacting genes. As shown in Fig. 4, metallothionein is up-regulated both in reference and tolerance animals upon Cd exposure. This is in accordance with previous studies, so that its position in the group of down-regulated clones in the microarray analysis may be an artefact of the hierarchical clustering method. However, in contrast to previous quantitative RT-PCR data, basal expression levels of *mt* seem to be lower in tolerant animals when compared with reference animals.

Discussion

The presented microarray data show that regulatory change of gene expression is an important factor associated with adaptation to heavy metals in natural *O. cincta* populations occurring at metal-contaminated sites. It is tempting to suggest that genes responding to Cd in a population-dependent context may have been molecular targets by which field-selected metal tolerance evolved. However, not all expression differences are necessarily a direct result of Cd exposure and so not all of them can be considered primary targets of micro-evolution. Still, this study shows that several stress responsive genes are less inducible in tolerant animals but exhibit high basal expression levels. This suggests that the Cd detoxification mechanism is constitutively overexpressed in tolerant animals, implying that cellular homeostasis is not disturbed upon exposure to Cd. In addition, genes important in food uptake are down-regulated in reference animals due to environmental Cd, while these genes are hardly affected in tolerant animals. This may explain the observation by Posthuma *et al.* (1992) that Cd causes less adverse effects on the growth of tolerant animals. Finally, the stress-activated kinase pathway seems to be adapted in such a way that tolerant animals are capable of modulating the above-mentioned processes in order to withstand chronic Cd exposure.

Our study shows that SSH (Diatchenko *et al.* 1996) is an effective way to enrich for genes that are significantly up-regulated by Cd exposure. About half of the isolated clones were significantly affected by either Cd, population or showed a significant interaction between population and Cd. One drawback of the enrichment procedure is the redundancy introduced on the microarray. Indeed, only 590 unigenes were retrieved from the 1051 sequenced cDNAs and for instance cluster Occ00070 was represented by 22 identical cDNAs on our array (http://www.collembase.org). However, redundancy also indicates the importance of a particular gene/cluster in the Cd response (Shaw *et al.* 2007). We, finally, note that the SSH enrichment in general misses genes that are differentially expressed less than threefold (Ji *et al.* 2002), which implies that potentially important genes responding more subtle but significantly to Cd will be overlooked.

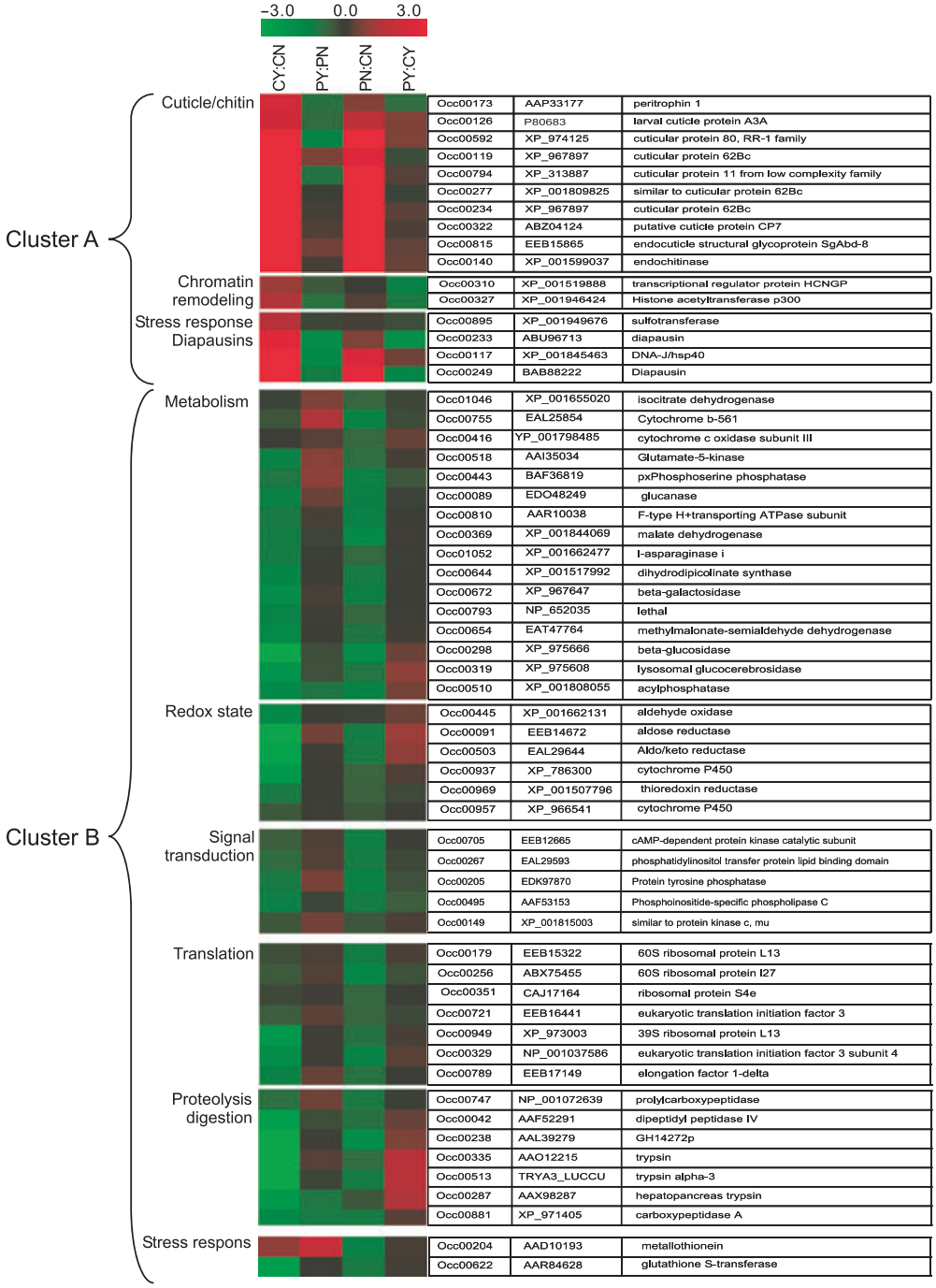


Fig. 4 Clusters of GO-annotated clones that show significant interaction between cadmium exposure and population (0.1% FDR corrected) arranged according to functional group. Red: up-regulation; green: down-regulation; labels are cluster IDs (Table S3). Cluster A: up-regulation by cadmium in reference culture and non-responsive or down-regulated in tolerant animals. Cluster B: down-regulation by cadmium in reference culture. Sample description as in Fig. 3.

Comparison of microarray data with previous qRT-PCR results

In a previous quantitative RT-PCR study, we determined the Cd response of a reference culture and tolerant animals from Plombières for 26 different unigenes under identical exposure conditions (Roelofs *et al.* 2007). These genes were also present on the microarray described in this experiment; however, the animals assessed were from cultures differing about 1 year (about two to three generations). For a substantial number of clones, the direction of differential gene

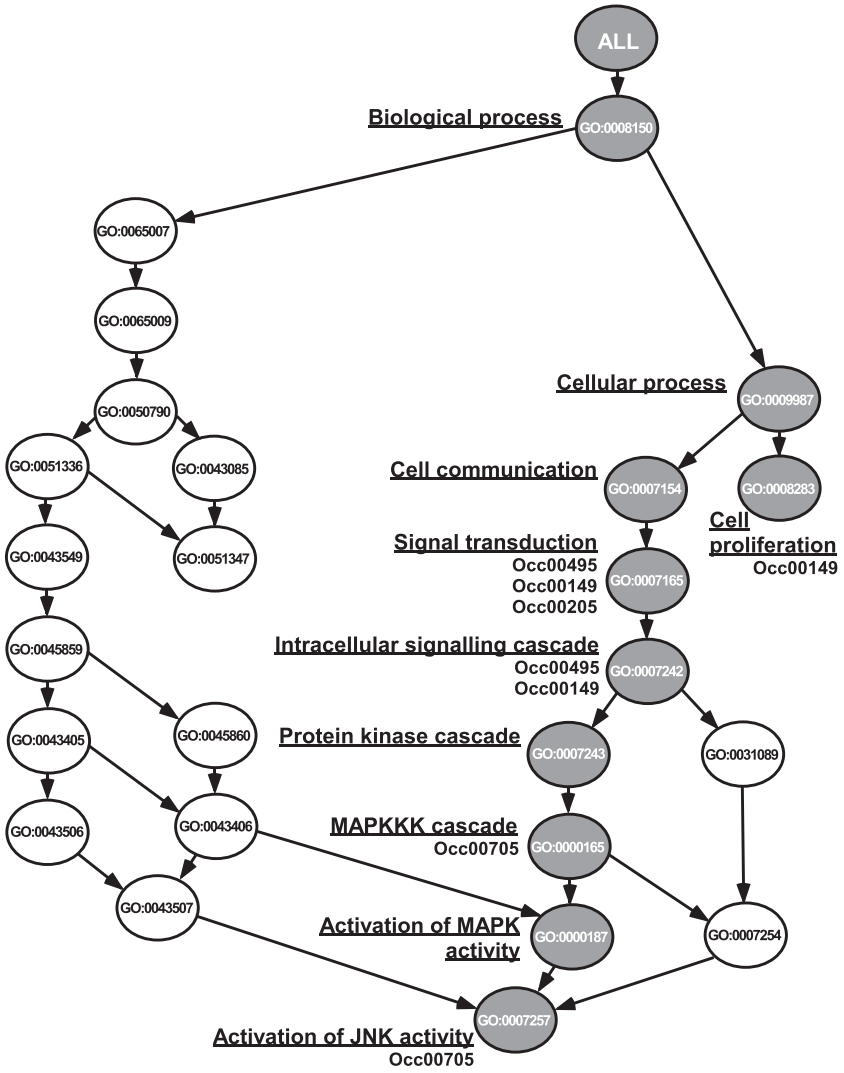


Fig. 5 GO hierarchical graph (getGOGraph in R package GOSIM) of biological process with clusters identified in the stress-activated protein kinase pathway (grey circles). BLAST annotation of identified clusters: Occ00149, protein kinase C mu; Occ00205, protein tyrosine phosphatase; Occ00495 phosphoinositide-specific phospholipase C; Occ00705, cAMP-dependent protein kinase catalytic subunit. Expression profile of the identified cluster numbers are shown in Fig. 4 in functional cluster 'signal transduction' (cluster B). Open circles indicate related biological processes that have not been identified in the current study.

expression was similar between the two quantification techniques, which resulted in a significant ($P < 0.01$) Pearson correlation coefficient of 0.6. Still, eight of the 26 clones showed opposite directions of gene regulation. Roelofs *et al.* (2007) reported seven clones that showed a significant interaction between Cd and population. We confirmed the significance of gene regulation for six of these clones. The clone (Oc_SSHCd_02E09, cluster Occ00058) homologous to an MAPK phosphatase (slingshot), however, showed an opposite response when compared with the quantitative RT-PCR data of Roelofs *et al.* (2007). According to these data, this gene was up-regulated in Cd-exposed reference animals, while it was down-regulated in Cd-exposed reference animals according to the microarray data. Another inconsistency was observed when constitutive levels of metallothionein (*mt*) gene expression levels were compared between reference and tolerant animals. Quantitative RT-PCR analysis from previous studies (Timmermans *et al.* 2005; Roelofs *et al.* 2007) has revealed a highly significant elevated constitutive expression of this gene in non-exposed tolerant Plombières animals when compared with non-exposed reference animals. The array data showed very slight but significant lower constitutive *mt* RNA abundance in unexposed tolerant samples when compared with reference samples. Explaining the above-described inconsistencies between the two platforms remains highly speculative. Transcriptional regulation

may have been changed during subsequent generations of culturing in the laboratory. Alternatively, the inconsistencies may be caused by a technical artefact, for instance, as a result of lowess correction of the raw microarray data.

A transcriptome-based model for heavy metal tolerance

Very recently, genome-wide transcriptional responses to Cd have been studied in *Daphnia* (Poynton *et al.* 2007; Shaw *et al.* 2007; Soetaert *et al.* 2007) and comparison of these studies with our findings show some striking similarities of predicted protein classes regulated by Cd exposure. First, metallothionein induction has been reported in *Daphnia magna* as well. Second, Poynton *et al.* (2007) reported significant down-regulation of genes coding for proteases and other digestive enzymes (glucanases). The *O. cincta* reference culture showed a similar response, in addition to down-regulated genes involved in energy metabolism and translation. However, the genes involved in these processes were hardly affected in the tolerant population. The absence of an inhibitory effect on digestive enzymes and the translational machinery could explain the smaller growth reduction upon Cd exposure observed in tolerant populations (Posthuma *et al.* 1992). Furthermore, the study of Shaw *et al.* (2007) showed that genes involved in cuticle formation were up-regulated by Cd, just as in the case of *O. cincta* reference culture. Remarkably, the tolerant population did not show induction of these genes upon Cd exposure, but instead a high basal transcriptional level was observed. Other cuticular proteins were overexpressed during Cd exposure in the tolerant population (Table S3). These observations support the hypothesis of a complex moulting cycle-regulated Cd excretion mechanism (Joosse & Buker 1979), in which midgut epithelium morphogenesis takes place (Bauer *et al.* 2001). This may also be related to the induction of an innexin (gap junction component) by Cd, by which epithelial cells are capable of communicating by the diffusion of small molecules (such as calcium). Induction of innexin expression may be necessary to compensate for the reduction in gap junction intercellular communication exerted by Cd (Fukumoto *et al.* 2001).

Cadmium uptake

The induction of innate anti-microbial defence mechanisms, such as diapausin and peritrophin, in the reference population (and the elevated constitutive expression in tolerant animals) are possibly due to damage of the cell membrane. Oxidative stress causes peroxidation of phospholipids (Valko *et al.* 2005) and this could make the midgut epithelium vulnerable to infections by micro-organisms. Alternatively to the N-type voltage gate blocking anti-fungal mechanism, diapausin structural studies of the leaf beetle *Gastrophysa atrocyanea* (the same accession as one of the BLAST hits in this study) revealed a calcium channel-blocking function (Kouno *et al.* 2007). Because the primary uptake route of Cd is through calcium channels (Fukuda & Kawa 1977; Craig *et al.* 1999), the induction of diapausins could represent a primary defence mechanism upon Cd exposure by blocking the animal's own channels. These diapausins exhibit a constitutive elevated expression in the tolerant population from Plombières, for which, due to the impaired calcium uptake, a calcium homeostasis related trade-off could exist. Moreover, diapausin is down-regulated upon Cd exposure in the tolerant animals, which suggests a need for calcium in the tolerance mechanism, such as replenishment of the calcium stores in the smooth endoplasmic reticulum. The induction of a sulphotransferase in the reference population, which is involved in the assimilation of sulphates, may be interpreted by the need for the synthesis of cysteine-containing molecules, such as metallothioneins and glutathione or even phytochelatin. The induction of sulphotransferases has been observed before in Cd-exposed plants (Nussbaum *et al.* 1988) and copper-exposed *Daphnia magna* (Poynton *et al.* 2007).

The stress-activated protein kinase pathway

The steady state or subtle induction by Cd in tolerant animals of genes, of which the protein products are involved in a putative SAPK cascade, may point to an adaptive micro-evolutionary process or may be merited by an increased first-line defence at the calcium channels. We identified five genes (of which four are represented in Fig. 5) potentially involved in a stress signalling pathway. In general, two endpoints, cell survival and cell death, are possible. Therefore, there are two separate pathways with the final effectors ERK and SAPK/JNK (Kuwabara *et al.* 2008). The pathway revealed is putatively involved in the apoptosis of the midgut epithelium in preparation of moulting and midgut regeneration. Although in vertebrates Cd stress signalling involves SAPK and down stream targets such as c-fos/c-jun and MTF-1 (Tibbles & Woodgett 1999; Saydam *et al.* 2002; Matsuoka *et al.* 2004), the exact mechanism and the mode of post-translational modifications remain elusive in *O. cincta*.

Metabolism and growth

The induction, in both reference and tolerant animals, of a *melted* orthologue with a pleckstrin domain, suggests a link between trace metal metabolism and the

conserved target of rapamycin (TOR) signalling pathway, which determines plastic organismal growth based on nutritional status and environmental cues. In yeast, *Drosophila* and mammalian cells, it is known to regulate about 5% of the genes (Reiling & Sabatini 2006). The product of *melted* in *D. melanogaster* enhances, together with other interacting proteins, the activation of TOR, a signalling pathway that influences metabolism and growth. The TOR pathway involves post-translational control of protein 4E-BP (elongation factor 4E-binding protein). TOR is inactivated by oxidative stress signalling from the sensor Tsc1/2 complex and this influences the metabolism and growth (insulin/PI3K pathway; Reiling & Sabatini 2006). The Tsc1/2 complex is recruited to the cell membrane region by *melted*, because the latter binds inositol phopholipids. By the inactivation of TOR, 4E-BP stays active and arrest growth and biosynthesis. Therefore, 4E-BP is known as a metabolic brake under environmentally adverse conditions (Teleman *et al*. 2005a,b). We suggest that induction of the *melted* orthologue by Cd in both reference and tolerant animals ensures a baseline activation of the TOR pathway, and hence lipid biosynthesis, in order to compensate the inactivation of TOR by oxidative stress (Reiling & Sabatini 2006) The interplay between stress pathways and the TOR pathway balances the gene expression between growth and stress-related genes, but the mechanisms which are responsible for this are still elusive (Lopez-Maury *et al*. 2008).

Anyhow, because of the large effect of Cd on lipid biosynthesis and the hypothesized TOR/insulin/PI3K pathway, and because of the population-specific expression upon Cd exposure of members from the putative SAPK signalling pathway, it seems that these mechanisms are intertwined and dependent on the status of the phospholipid bilayer, which is a site of action for ROS during conditions deviating from redox homeostasis. Moreover, proteins of the PI3K family play a pivotal role in between the two presumed pathways (Wymann & Pirola 1998) and also more downstream the insulin and SAPK pathways are known to interact in their response to adverse environmental stimuli (Wolf *et al*. 2008).

In conclusion, we have observed significant changes in transcription upon Cd exposure. Many transcriptional responses show a highly significant interaction between Cd exposure and source of the population, i.e. reference or tolerant. The tolerant Plombières population is much less affected by Cd exposure. Genes involved in stress response, exoskeleton maintenance, anti-microbial defence, calcium channel blockers and transcriptional regulation via histone acetylation- and deacetylation show a high basal expression and are less inducible in tolerant animals when compared with the reference culture. By contrast, genes involved in digestion, stress signalling and energy metabolism are less down- or up-regulated upon Cd exposure and show a lower basal expression in tolerant animals when compared with the reference culture. Our observations confirm the micro-evolutionary processes occurring in soil arthropod populations at historically heavy metal-contaminated sites (Timmermans *et al*. 2005; Janssens *et al*. 2007, 2008).

Acknowledgements

The authors thank Paul van den IJssel for spotting the cDNAs, Paul Eijk for performing the initial microarray hybridizations and Oscar Krijgsman for providing the initial LIMMA scripts. TKS Janssens was supported by a grant from the Dutch Research council (NWO no. 81304006). This research was in part financed by a grant from the Netherlands Genomics Initiative.

References

Bauer R, Lehmann C, Hoch M (2001) Gastrointestinal development in the *Drosophila* embryo requires the activity of innexin gap junction channel proteins. *Cell Communication and Adhesion*, **8**, 307–310.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate – a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Methodological*, **57**, 289–300.

Bergman AM, Eijk PP, van Haperen V *et al.* (2005) In vivo induction of resistance to gemcitabine results in increased expression of ribonucleotide reductase subunit M1 as the major determinant. *Cancer Research*, **65**, 9510–9516.

Craig A, Hare L, Tessier A (1999) Experimental evidence for cadmium uptake via calcium channels in the aquatic insect *Chironomus staegeri*. *Aquatic Toxicology*, **44**, 255–262.

Daborn PJ, Yen JL, Bogwitz MR *et al.* (2002) A single P450 allele associated with insecticide resistance in *Drosophila*. *Science*, **297**, 2253–2256.

Diatchenko L, Lau Y-FC, Campbell AP *et al.* (1996) Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 6025–6030.

Edgar R, Domrachev M, Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Research*, **30**, 207–210.

Fisher MA, Oleksiak MF (2007) Convergence and divergence in gene expression among natural populations exposed to pollution. *BMC Genomics*, **8**, 108.

Frohlich H, Speer N, Poustka A, Beissarth T (2007) GOSim – an R-package for computation of information theoretic GO similarities between terms and gene products. *Bmc Bioinformatics*, **8**, 166.

Fukuda J, Kawa K (1977) Permeation of manganese, cadmium, zinc, and beryllium through calcium channels of an insect muscle membrane. *Science*, **196**, 309–311.

Fukumoto M, Kujiraoka T, Hara M *et al.* (2001) Effect of cadmium on gap junctional intercellular communication in

primary cultures of rat renal proximal tubular cells. *Life Sciences*, **69**, 247–254.

Gompel N, Prud'Homme B, Wittkopp PJ, Kassner VA, Carroll SB (2005) Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature*, **433**, 481–487.

Gordon D (2008) *Phrap*. University of Washington TechTransfer, Washington, DC.

Gracey AY, Troll JV, Somero GN (2001) Hypoxia-induced gene expression profiling in the euryoxic fish *Gillichthys mirabilis*. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 1993–1998.

Jacob F, Monod J (1961) On regulation of gene activity. *Cold Spring Harbor Symposia on Quantitative Biology*, **26**, 193.

Janssens TKS, Marien J, Cenijn P *et al.* (2007) Recombinational micro-evolution of functionally different metallothionein promoter alleles from *Orchesella cincta*. *BMC Evolutionary Biology*, **7**, 88.

Janssens TKS, Lopez RDR, Marien J *et al.* (2008) Comparative population analysis of metallothionein promoter alleles suggests stress-induced microevolution in the field. *Environmental Science & Technology*, **42**, 3873–3878.

Ji W, Wright MB, Cai L, Flament A, Lindpaintner K (2002) Efficacy of SSHPCR in isolating differentially expressed genes. *Bmc Genomics*, **3**, 12.

Jiang HM, Fu K, Andrews GK (2004) Gene- and cell-type-specific effects of signal transduction cascades on metal-regulated gene transcription appear to be independent of changes in the phosphorylation of meta 1-response-element-binding transcription factor-1. *Biochemical Journal*, **382**, 33–41.

Jin W, Riley RM, Wolfinger RD *et al.* (2001) The contributions of sex, genotype and age to transcriptional variance in *Drosophila melanogaster*. *Nature Genetics*, **29**, 389–395.

Joosse ENG, Buker JB (1979) Uptake and excretion of lead by litter-dwelling collembola. *Environmental Pollution*, **18**, 235–240.

Knight CA, Vogel H, Kroymann J *et al.* (2006) Expression profiling and local adaptation of *Boechera holboellii* populations for water use efficiency across a naturally occurring water stress gradient. *Molecular Ecology*, **15**, 1229–1237.

Kouno T, Mizuguchi M, Tanaka H *et al.* (2007) The structure of a novel insect peptide explains its Ca^{2+} channel blocking and antifungal activities. *Biochemistry*, **46**, 13733–13741.

Kuwabara M, Asanuma T, Niwa K, Inanami O (2008) Regulation of cell survival and death signals induced by oxidative stress. *Journal of Clinical Biochemistry and Nutrition*, **43**, 51–57.

Lock K, Janssens F, Janssen CR (2003) Effects of metal contamination on the activity and diversity of springtails in an ancient Pb–Zn mining area in Plombières, Belgium. *European Journal of Soil Biology*, **39**, 25–29.

Lopez-Maury L, Marguerat S, Bahler J (2008) Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nature Reviews Genetics*, **9**, 583–593.

Matsuoka M, Igisu H, Nakagawa K, Katada T, Nishina H (2004) Requirement of MKK4 and MKK7 for $CdCl_2$- or $HgCl_2$-induced activation of c-Jun NH2-terminal kinase in mouse embryonic stem cells. *Toxicology Letters*, **152**, 175–181.

McGinnis S, Madden TL (2004) BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Research*, **32**, W20–W25.

Nussbaum S, Schmutz D, Brunold C (1988) Regulation of assimilatory sulfate reduction by cadmium in *Zea mays* L. *Plant Physiology*, **88**, 1407–1410.

Parkinson J, Anthony A, Wasmuth J *et al.* (2004) PartiGene – constructing partial genomes. *Bioinformatics*, **20**, 1398–1404.

Posthuma L, Hogervorst RF, Van Straalen NM (1992) Adaptation to soil pollution by cadmium excretion in natural populations of *Orchesella cincta* (L.) (Collembola). *Archives of Environmental Contamination and Toxicology*, **22**, 146–156.

Posthuma L, Hogervorst RF, Joosse ENG, Van Straalen NM (1993) Genetic variation and covariation for characteristics associated with cadmium tolerance in natural populations of the springtail *Orchesella cincta*. *Evolution*, **47**, 619–631.

Poynton HC, Varshavsky JR, Chang B *et al.* (2007) Daphnia magna ecotoxicogenomics provides mechanistic insights into metal toxicity. *Environmental Science & Technology*, **41**, 1044–1050.

Reiling JH, Sabatini DM (2006) Stress and mTORture signaling. *Oncogene*, **25**, 6373–6383.

Rockman MV, Hahn MW, Soranzo N *et al.* (2005) Ancient and recent positive selection transformed opioid cis-regulation in humans. *PLoS Biology*, **3**, 1–12.

Roelofs D, Overhein L, De Boer ME, Janssens TKS, Van Straalen NM (2006) Additive genetic variation of transcriptional regulation: metallothionein expression in the soil insect *Orchesella cincta*. *Heredity*, **96**, 85–92.

Roelofs D, Marien J, van Straalen NM (2007) Differential gene expression profiles associated with heavy metal tolerance in the soil insect *Orchesella cincta*. *Insect Biochemistry and Molecular Biology*, **37**, 287–295.

Roelofs D, Aarts MGM, Schat H, van Straalen NM (2008) Functional ecological genomics to demonstrate general and specific responses to abiotic stress. *Functional Ecology*, **22**, 8–18.

Saeed AI, Sharov V, White J *et al.* (2003) TM4: A free, open-source system for microarray data management and analysis. *BioTechniques*, **34**, 374.

Saydam N, Georgiev O, Nakano MY, Greber UF, Schaffner W (2001) Nucleo-cytoplasmic trafficking of metal-regulatory transcription factor 1 is regulated by diverse stress signals. *Journal of Biological Chemistry*, **276**, 25487–25495.

Saydam N, Adams TK, Steiner F, Schaffner W, Freedman JH (2002) Regulation of metallothionein transcription by the metal-responsive transcription factor MTF-1 – identification of signal transduction cascades that control metal-inducible transcription. *Journal of Biological Chemistry*, **277**, 20438–20445.

Schmid R, Blaxter ML (2008) annot8r: GO, EC and KEGG annotation of EST datasets. *BMC Bioinformatics*, **9**, 180.

Schulte PM, Glemet HC, Fiebig AA, Powers DA (2000) Adaptive variation in lactate dehydrogenase-B gene expression: role of a stress-responsive regulatory element. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 6597–6602.

Shapiro MD, Bell MA, Kingsley DM (2006) Parallel genetic origins of pelvic reduction in vertebrates. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 13753–13758.

Shaw JR, Colbourne JK, Davey JC *et al.* (2007) Gene response profiles for *Daphnia pulex* exposed to the environmental stressor cadmium reveals novel crustacean metallothioneins. *Bmc Genomics*, **8**, 477.

Soetaert A, Vandenbrouck T, van der Ven K *et al.* (2007) Molecular responses during cadmium-induced stress in Daphnia magna: Integration of differential gene expression with higher-level effects. *Aquatic Toxicology*, **83**, 212–222.

Sterenborg I (2003) *Molecular Physiology of Metal Tolerance in Orchesella cincta*: The Role of Metallothionein. Vrije Universiteit, Amsterdam.

Stürzenbaum SR, Winters C, Galay M, Morgan AJ, Kille P (2001) Metal ion trafficking in earthworms. Identification of a cadmium-specific metallothionein. *Journal of Biological Chemistry*, **276**, 34013–34018.

Teleman AA, Chen Y-W, Cohen SM (2005a) *Drosophila* melted modulates FOXO and TOR activity. *Developmental Cell*, **9**, 271–281.

Teleman AA, Chen YW, Cohen SM (2005b) 4E-BP functions as a metabolic brake used under stress conditions but not during normal growth. *Genes & Development*, **19**, 1844–1848.

Tibbles LA, Woodgett JR (1999) The stress-activated protein kinase pathways. *Cellular and Molecular Life Sciences*, **55**, 1230–1254.

Timmermans MJTN, Ellers J, Roelofs D, Van Straalen NM (2005) Metallothionein mRNA expression and cadmium tolerance in metal-stressed and reference populations of the springtail *Orchesella cincta*. *Ecotoxicology*, **14**, 727–739.

Timmermans MJTN, De Boer ME, Nota B *et al.* (2007) Collembase: a repository for springtail genomics and soil quality assessment. *BMC Genomics*, **8**, 341.

Valko M, Morris H, Cronin MTD (2005) Metals, toxicity and oxidative stress. *Current Medicinal Chemistry*, **12**, 1161–1208.

Van Straalen NM, Burghouts TBA, Doornhof MJ *et al.* (1987) Efficiency of lead and cadmium excretion in populations of *Orchesella cincta* (Collembola) from various contaminated forest soils. *Journal of Applied Ecology*, **24**, 953–968.

Van Straalen NM, Schobben JHM, De Goede RGM (1989) Population consequences of cadmium toxicity in soil microarthropods. *Ecotoxicology and Environmental Safety*, **17**, 190–204.

Waddington CH (1939) Physiological genetics. *Nature*, **144**, 817–818.

Wettenhall JM, Smyth GK (2004) LIMMA GUI: a graphical user interface for linear modeling of microarray data. *Bioinformatics*, **20**, 3705–3706.

Whitehead A, Crawford DL (2006) Neutral and adaptive variation in gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 5425–5430.

Wolf M, Nunes F, Henkel A, Heinick A, Paul RJ (2008) The MAP kinase JNK-1 of *Caenorhabditis elegans*: location, activation, and influences over temperature-dependent insulin-like signaling, stress responses, and fitness. *Journal of Cellular Physiology*, **214**, 721–729.

Wray GA (2007) The evolutionary significance of cis-regulatory mutations. *Nature Reviews Genetics*, **8**, 206–216.

Wu H, Churchill GA (2006) *R/MAANOVA: An Extensive R Environment for the Analysis of Microarray Experiments*, 1.2.1 edn. The Jackson Laboratory, Bar Harbor.

Wymann MP, Pirola L (1998) Structure and function of phosphoinositide 3-kinases. *Biochimica et Biophysica Acta – Molecular and Cell Biology of Lipids*, **1436**, 127–150.

We are interested in the molecular mechanisms underlying stress responses in populations of soil invertebrates. We study the link between genetic elements accounting for these eco-physiological responses and their impact on evolution. We mainly focus on polymorphisms in gene regulatory elements that contribute to phenotypes adapted to abiotic stress factors. From these studies we try to develop genetic markers that maybe useful for risk assessment of soil pollution.

Supporting information

Additional supporting information may be found in the online version of this article:

Appendix S1 The supplementary archive contains MAANOVA input data-, design- and script files.

Fig. S1 Hierarchical clustering of all 606 cDNA clones that show a significant interaction between cadmium and population.

Table S1 ^{2}log expression ratios for each cDNA clone on each of the 16 hybridized arrays

Table S2 ANOVA derived F-statistic (Fs, James-Stein estimation) of each cDNA clone for the term cadmium, population and the interaction between cadmium and population

Table S3 Gene clusters for which known gene functions were retrieved by BLASTX queries associated by GO terms for biological process

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.