

Research article

Open Access

Gene response profiles for *Daphnia pulex* exposed to the environmental stressor cadmium reveals novel crustacean metallothioneins

Joseph R Shaw^{*1,2,3,4}, John K Colbourne⁵, Jennifer C Davey^{2,3},
Stephen P Glaholt^{1,2,4}, Thomas H Hampton^{2,3}, Celia Y Chen^{1,2},
Carol L Folt^{1,2} and Joshua W Hamilton^{2,3}

Address: ¹Department of Biology, Dartmouth College, Hanover, New Hampshire 03755, USA, ²Center for Environmental Health Sciences at Dartmouth, Dartmouth Medical School, Hanover NH 03755, USA, ³Department of Pharmacology & Toxicology, Dartmouth Medical School, Hanover NH 03755, USA, ⁴The School of Public and Environmental Affairs, Indiana University, Bloomington, Indiana 47405, USA and ⁵The Center for Genomics and Bioinformatics, Indiana University, Bloomington, Indiana 47405, USA

Email: Joseph R Shaw* - joeshaw@indiana.edu; John K Colbourne - jcolbour@cgb.indiana.edu;
Jennifer C Davey - jennifer.c.davey@dartmouth.edu; Stephen P Glaholt - sglaholt@indiana.edu;
Thomas H Hampton - thomas.hampton@dartmouth.edu; Celia Y Chen - celia.y.chen@dartmouth.edu;
Carol L Folt - carol.l.folt@dartmouth.edu; Joshua W Hamilton - joshua.w.hamilton@dartmouth.edu

* Corresponding author

Published: 21 December 2007

Received: 19 May 2007

BMC Genomics 2007, 8:477 doi:10.1186/1471-2164-8-477

Accepted: 21 December 2007

This article is available from: <http://www.biomedcentral.com/1471-2164/8/477>

© 2007 Shaw et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Genomic research tools such as microarrays are proving to be important resources to study the complex regulation of genes that respond to environmental perturbations. A first generation cDNA microarray was developed for the environmental indicator species *Daphnia pulex*, to identify genes whose regulation is modulated following exposure to the metal stressor cadmium. Our experiments revealed interesting changes in gene transcription that suggest their biological roles and their potentially toxicological features in responding to this important environmental contaminant.

Results: Our microarray identified genes reported in the literature to be regulated in response to cadmium exposure, suggested functional attributes for genes that share no sequence similarity to proteins in the public databases, and pointed to genes that are likely members of expanded gene families in the *Daphnia* genome. Genes identified on the microarray also were associated with cadmium induced phenotypes and population-level outcomes that we experimentally determined. A subset of genes regulated in response to cadmium exposure was independently validated using quantitative-realtime (Q-RT)-PCR. These microarray studies led to the discovery of three genes coding for the metal detoxication protein metallothionein (MT). The gene structures and predicted translated sequences of *D. pulex* MTs clearly place them in this gene family. Yet, they share little homology with previously characterized MTs.

Conclusion: The genomic information obtained from this study represents an important first step in characterizing microarray patterns that may be diagnostic to specific environmental contaminants and give insights into their toxicological mechanisms, while also providing a practical tool for evolutionary, ecological, and toxicological functional gene discovery studies. Advances in *Daphnia* genomics will enable the further development of this species as a model organism for the environmental sciences.

Background

Recent advances in genomics and bioinformatics are revolutionizing the process of discovering genes whose regulation has important consequences to the fitness of individuals [1,2]. These resources for genetic investigations include functional genomic tools that compare the expression profiles of thousands of genes under different conditions. Microarrays are proving to be a particularly important modern resource for identifying genes in context of their complex regulation [3]. Until recently, the application of microarrays has largely focused on studies from a select set of organisms (*Saccharomyces cerevisiae*, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Danio rerio*, *Mus musculus*) that constitute some of the most well characterized laboratory models in the life sciences. While these model organisms are well suited for developmental, cellular and molecular studies – contributing a staggering amount of biological knowledge – it is difficult to relate environmental controls of gene regulation of these organisms (phenotype) to higher-level (population) responses because so little is known about their ecology. This chasm presents a challenge for toxicological genomic applications, especially those related to environmental toxicology where the goal is often to identify population and ecosystem-level responses in the context of environment change. Thus, the aim of discovering genes that are expressed as a function of ambient conditions (therefore, anchoring potential genetic biomarkers to biological/ecological functions) requires applying functional genomic approaches to keystone species with accessible natural populations and tractable ecologies, such as the ubiquitous aquatic micro-crustacean *Daphnia pulex*.

Commonly known as water fleas, *Daphnia* are familiar and ubiquitous inhabitants of ponds and lakes throughout the globe and have been the focus of study by limnologists for well over a century [4-6]. As a dominant member of the planktonic community, *Daphnia* play a central role in aquatic food webs, serving as the primary grazers of algae, bacteria and protozoans, and as the primary forage for fish. As a result, *Daphnia* are long recognized as a sentinel/indicator species in freshwater ecosystems and they are routinely used to determine the toxicity of aqueous solutions and to gauge the quality of inland waters [7]. Therefore, the gross-level responses of *Daphnia* to a number of environmental pollutants are well characterized [8], and these responses are referenced by environmental protection agencies (e.g., United States Environmental Protection Agency, Environment Canada, Organization for Economic Co-operation and Development) to define regulatory limits [9], monitor the quality of industrial and municipal effluents [10,11] and estimate the risk of environmental toxins on natural environments. For these reasons, the present study extends the "road-map" for applications of DNA microarrays to studies of

non-model organisms described by Gracey et al. [12] and later by others (reviewed in [13,14]), by characterizing the response of this critical aquatic indicator species to environmental stress. These investigations utilize *D. pulex*; a freshwater crustacean species that is ubiquitous throughout North America [15], sensitive to metals and metal mixtures [16], and supported by a major genome sequencing project [17]. They also complement recent efforts by others to develop a microarray platform for *D. magna* [18]; [19]; [20].

We focus our present investigations on cadmium as a model environmental stressor, but the technologies described are pertinent to a much wider range of ecological and toxicological applications. Cadmium is a ubiquitous environmental contaminant [21]. It ranks eighth on the Agency for Toxic Substances and Disease Registry [22] list of the top 50 priority pollutants, is one of the most common contaminants found in the U.S. EPA Superfund sites [23], and is highly toxic to aquatic life. Cadmium is extremely persistent in the environment and, as a result, bioaccumulates within food webs [24]. In aquatic animals, cadmium is a potent calcium antagonist that disrupts calcium uptake and homeostasis [25-27]. Cadmium also induces oxidative stress, resulting in lipid peroxidation, damage to membranes, impaired cellular functions, and tissue damage [28]. It is a Class B metal that has ten outer shell electrons in the *d* orbital and is highly reactive, tending to preferentially form covalent bonds with S>N>P>O [29,30]. Thus, complexation with anions controls exposure in the water column [31] and ultimately toxicity in the animal [32]. Within many organisms, the major ligands for cadmium are small metal binding proteins known as metallothionein (MT) [32,33].

The metallothioneins are a family of unusual and highly conserved small cytosolic proteins, characterized by their low molecular weight (i.e., 6000–7000 Daltons), lack of secondary structure in the absence of metal ligands, absence of histidine or aromatic residues, and high cysteine content (typically 30–33% of the protein). The MTs principally bind metals from groups 1B and 2B of the Periodic Table of Elements (e.g., cadmium, copper, mercury, silver, zinc; [34,35]), and in so doing form two metal-binding domains. The metal binding properties of MT are conferred by the large number, spacing and metal coordination of the cysteine residues within the protein. It is generally recognized that the physiological functions of MT are to regulate the intra-cellular concentrations of essential metals such as copper and zinc; to activate and deactivate zinc-regulated proteins; and to scavenge free radicals [35,36]. However, MT can also confer protection from metal poisoning by binding certain free metal ions or undergoing exchange reactions with metals bound to other ligands [35,37]. The synthesis of MT is also regu-

lated by metals with MT mRNA expression increasing in response to elevated metals in tissues [38,39]. Because of its high specificity and sensitivity for metal induction, MT levels have been successfully used to diagnose copper, zinc, and cadmium exposures in numerous studies ranging from temperate freshwater environments to tropical marine systems [40]. While MTs have been characterized in several Malacostraca crustacean species [41], with the exception of partial cDNA sequences of two MTs from *D. magna* [19] there is little sequence information available for these genes and their regulatory regions in the Branchiopoda, such as *Daphnia*.

We report a series of studies designed to test whether the exposure to sub-lethal chemical stressors results in identifiable changes in gene expression, exposing genes that respond to these conditions and providing a means of identifying potential biomarkers of response to specific exposures. Here, we developed a cDNA microarray platform for *D. pulex* to investigate differences in gene-response profiles for this aquatic indicator species following environmental perturbation by cadmium. Expression profiling successfully aided in the discovery of genes regulated in response to cadmium exposure, including the important metal biomarker, MT. Gene-response profiling provided mechanistic information that related to the observed cadmium-induced phenotypes and population-level responses. Finally, we identified and characterized the primary structure of the *D. pulex* MT sequences in context of their genomic architecture, translated sequence and phylogenetic relationship to other crustacean MTs.

Results

***Daphnia* response to cadmium**

To define sub-lethal exposure concentrations and better characterize cadmium-exposed phenotypes, preliminary experiments were conducted to determine the sensitivity of *D. pulex* to cadmium. For these studies, acute (48-h) toxicity tests were used to define median lethal (LC50) and sub-lethal (LC01) effects concentrations. Mortality in the control groups was less than 10% for all tests. The LC50 and LC01 values given as mean ± 95% confidence intervals were 74.6 (64.2 – 84.4) and 20.3 (9.7 – 30.1) µg/

L, respectively. Demographic (i.e., life-table) experiments that included longer-term exposures (21-d) to lower cadmium concentrations (1 to 2.5 µg Cd/L) were conducted to better define sub-lethal cadmium responses. Observed effects included a decline in individual fitness parameters of size and lipid content to ovary size index, and population-level endpoints of number of clutches, cumulative reproductive rate, and per capita birth rate (Table 1). These effects are highlighted by the two representative micrographs of control and cadmium-treated animals (Figure 1), which were taken at the end of the 21-d exposure period.

Microarray results

Microarrays were used to compare the gene-expression profiles of *D. pulex* maintained under control conditions with those exposed to 20 µg Cd/L for 48-h (i.e., sub-lethal concentration, ~LC01; [GEO:GSE9746]). These utilized RNA isolated from three independent and concurrently replicated exposures of *Daphnia* to cadmium and control conditions, applied to three replicate microarrays using a standard control vs. treated design that included dye swaps (for details see Methods). Gene expression log ratios ($M = \log_2 \text{treated/control}$) across the three microarrays were determined as described in Methods and plotted against log mean intensity values ($A = 1/2 \log_2 (\text{treated} * \text{control})$) as shown in Figure 2[42]. The M-values were distributed around zero, with print control elements that contained no cDNA possessing the lowest intensity (or A values). As expected, negative control elements, which contained non-*Daphnia* cDNA also, had low A-values. However, within these two groups (i.e., print control elements, negative control elements), gene expression (or M-values) was quite varied because of the considerable noise observed with these low intensity values. Positive control elements, which contained known *D. pulex* genes (cytochrome C, cytochrome B, actin, and ferritin) distributed with other *D. pulex* elements on the MA plot.

Empirical Bayes statistics using a p-value cut-off of 0.05 revealed 99 elements (2.9% of the array) for which expression was increased following cadmium treatment and 30 elements (1.1% of the array) for which expression

Table 1: Effects of 21-day cadmium exposure on *Daphnia pulex*

Cadmium (µg/l)	Length (mm)	Lipid-Ovary Index ^a	Number of Clutches ^a	Total neonates ^a	Per Capita Birth Rate (b) ^a
0	2.90 ± 0.11a	3.83 ± 1.50a	5.50 ± 0.55a	118.86 ± 29.94a	5.37 ± 0.89a
0.25	2.54 ± 0.10b	1.57 ± 1.13b	4.63 ± 0.52a	65.13 ± 16.60b	2.88 ± 0.66b
0.5	2.41 ± 0.12b	1.14 ± 0.69b	4.29 ± 0.76a	51.00 ± 25.66b	2.22 ± 1.12b
1	2.26 ± 0.16b, c	1.25 ± 0.96b	2.00 ± 1.00b	8.60 ± 3.85c	0.39 ± 0.18c
1.75	2.16 ± 0.12c	1.00 ± 0.93b	2.22 ± 1.09b	13.00 ± 8.04c	0.58 ± 0.34c
2.5	2.04 ± 0.14c	1.0 ± 0.83b	1.50 ± 1.05b	5.67 ± 5.32c	0.25 ± 0.23c

^aValues with different letters are significantly different (P < 0.05).

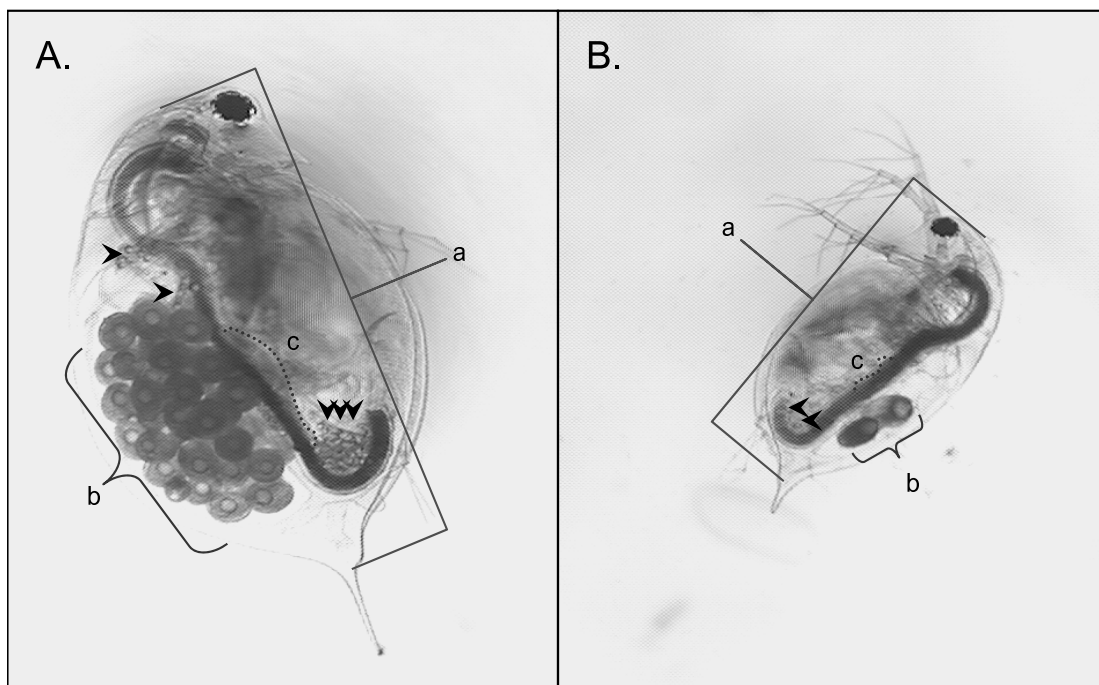


Figure 1
Cadmium induced phenotype. Representative micrographs of 21-d *Daphnia pulex* maintained under A) control conditions or B) exposed to cadmium (2.5 µg Cd/L). Images were collected at the same scale and are presented as raw image files. Differences were observed in a) body length and b) number of eggs in the brood chamber. The control animal also has a larger c) ovary and more pronounced lipid stores (represented by arrows).

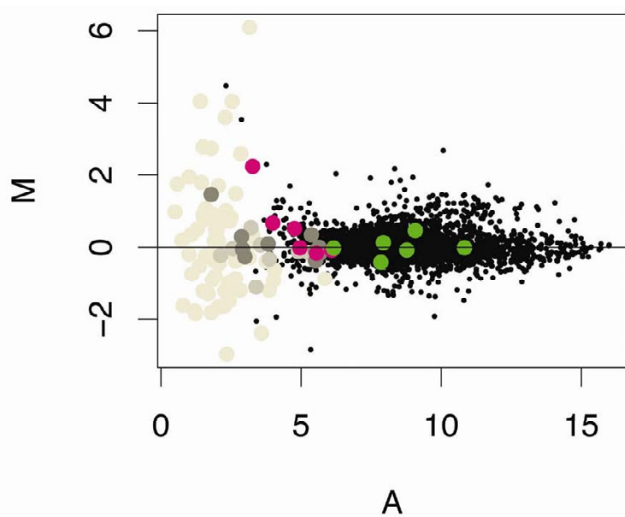


Figure 2
Cadmium effects on gene expression: Buffers, blanks, and controls. Gene expression data from control and cadmium (20 µg Cd/L for 48-h) treated *Daphnia pulex* [GEO:GSE9746]. Data were LOESS normalized; duplicate probes were averaged within LIMMA using gene-wise linear models fit to expression data, and gene expression log ratios ($M = \log_2 \text{treated/control}$) were plotted against log mean intensity values ($A = 1/2 \log_2 (\text{treated} * \text{control})$). Print control elements (buffers, blanks, betaine; grey shades); negative controls (non-daphniid cDNA; pink); and positive controls (*D. pulex* cytochrome b and c, actin, and ferritin; green) are highlighted.

decreased (Table 2, Figure 3). Of these elements, 95 were sequenced from the 5'-end. These ESTs clustered into 42 assembled sequences. Sequence analysis and alignment comparisons of these assembled sequences using the Blastx program and an expectation-value cut-off of 1×10^{-3} and 33 matched amino acids [43] against the non-redundant protein sequence inventory at the NCBI identified 27 unique cDNAs with sequence similarities to known genes. These included 22 up and five down-regulated genes. Fifteen assembled sequences were unidentifiable, sharing no similarities to known genes. The difference between the number of sequenced elements identified as cadmium responsive (i.e., 95) and the total number of likely unique cDNAs (i.e., 42) was due to redundancy on the microarray. Although many of the elements whose ESTs were clustered into assembled sequences shared similar expression levels, in some cases their expression levels deviated by as much as two-fold (Table 2). The observed differences among clustered elements were even more pronounced when all sequenced elements belonging to the set of differentially regulated genes (EST cluster) were included in the analysis. The variation was not a function of differences in their intensity values, because deviations from the mean value of A were negligible (data not shown). Without further data, we were unable to verify whether some assembled sequences were composed of alterna-

Table 2: Cadmium regulated microarray elements

Sig ^a	Seq ^b	EST Cluster	M average ^c	Up-regulated elements		P Value ^d	Putative GO annotations
				A average ^c	Hit Description		
8	9	Contig 262	1.02 ± 0.31	10.12 ± 1.88	2-domain hemoglobin protein subunit	< 0.001	GO:0015671, GO:0005833, GO:0005344
2	16	Contig 272	0.43 ± 0.21	9.06 ± 1.50	CG30045-PA	< 0.001	GO:0042302
1	1	Singlet 433	0.71	12.08	CG6305-PA	0.062	GO:0042302
1	1	Singlet 498	1.02	9.15	chitinase-1	0.028	GO:0008061, GO:0006030, GO:0004568, GO:0005576
3	4	Contig 213	0.82 ± 0.12	9.07 ± 0.53	chitinase-2	< 0.001	GO:0008061, GO:0006030, GO:0004568, GO:0005576
1	3	Contig 218	0.93 ± 0.19	9.64 ± 1.77	chitotriosidase	0.005	GO:0008061, GO:0006030, GO:0016798, GO:0005576
1	3	Contig 209	0.41 ± 0.17	12.21 ± 1.48	CUO6 BLACRCuticle protein 6 (BcNCPI4.9)	0.06	GO:0042302
4	6	Contig 241	1.07 ± 0.31	10.76 ± 2.07	cuticle protein-1	0.018	GO:0042302
5	8	Contig 257	1.03 ± 0.33	9.96 ± 1.54	cuticle protein-2	< 0.001	GO:0042302
13	28	Contig 273	0.91 ± 0.32	10.05 ± 1.65	cuticle protein-3	< 0.001	GO:0042302
1	3	Contig 149	0.72 ± 0.28	8.55 ± 0.92	cuticle protein-4	0.011	GO:0042302
5	8	Contig 261	0.88 ± 0.30	9.87 ± 1.42	cuticle protein-5	< 0.001	GO:0042302
1	1	Singlet 15	1.23	14.06	ERGA6350	0.009	
1	1	Singlet 469	0.80	7.99	helix-loop-helix transcription factor	0.055	GO:0006355, GO:0003677
1	1	Contig 21	1.45	10.03	hemoglobin-1	0.003	GO:0015671, GO:0005833, GO:0005344
1	1	Contig 16	1.08	9.81	hemoglobin-2	< 0.001	GO:0015671, GO:0005833, GO:0005344
1	1	Singlet 65	2.16	8.34	Hypothetical protein CBGI4247	< 0.001	
4	4	Contig 221	1.98 ± 0.47	8.88 ± 0.82	Metallothionein	< 0.001	GO:0046872, GO:0006875
1	5	Contig 220	0.31 ± 0.49	8.95 ± 0.62	myosin 2 light chain	0.08	GO:0005509
2	4	Contig 227	0.70 ± 0.04	12.59 ± 2.37	OPSC1 HEMSACCompound eye opsin BCRH1	0.036	GO:0007602, GO:0016021, GO:0001584, GO:0007186
1	1	Singlet 251	1.18	10.90	PREDICTED: similar to chitinase	0.019	GO:0008061, GO:0006030, GO:0016798, GO:0005576
1	1	Singlet 459	1.15	7.42	PREDICTED: similar to glutathione S-transferase	0.021	GO:0016740
1	1	Singlet 1	1.94	8.99	Unknown EST-1	0.009	
3	3	Contig 232	1.14 ± 0.37	8.27 ± 0.99	Unknown EST-2	< 0.001	
7	8	Contig 253	1.05 ± 0.44	9.99 ± 1.38	Unknown EST-3	< 0.001	
1	1	Singlet 166	0.87	8.20	Unknown EST-4	0.048	
1	1	Singlet 64	0.75	9.95	Unknown EST-5	0.059	
1	1	Singlet 49	0.74	8.77	Unknown EST-6	0.067	
1	5	Contig 237	0.33 ± 0.21	10.11 ± 1.55	Unknown EST-7	0.073	
1	1	Singlet 172	0.68	8.23	Unknown EST-8	0.071	
24					Unsequenced		

Sig ^a	Seq ^b	EST Cluster	M average ^c	Down-regulated elements		P Value ^d	Putative GO annotations
				A average ^c	Hit Description		
1	2	Contig 135	-0.36 ± 0.66	9.28 ± 0.08	carboxypeptidase A1	0.068	GO:0006508, GO:0004182
1	1	Singlet 97	-1.04	6.94	endo-1,4-mannanase	0.009	GO:0016985, GO:0000272
1	3	Contig 202	-0.38 ± 0.36	8.39 ± 1.02	PREDICTED: similar to CG31997-PA	0.036	
2	2	Contig 83	-1.29 ± 0.06	8.00 ± 0.40	PREDICTED: similar to monooxygenase	0.001	
1	3	Contig 162	-0.37 ± 0.30	8.98 ± 1.04	ribosomal protein S11-2	0.02	GO:0006412, GO:0003735, GO:0005840
1	2	Contig 191	-0.61 ± 0.27	10.76 ± 3.35	Unknown EST-1	0.019	
2	4	Contig 212	-0.72 ± 0.38	8.77 ± 0.97	Unknown EST-2	< 0.001	
4	5	Contig 229	-1.32 ± 0.11	9.96 ± 0.90	Unknown EST-3	< 0.001	
4	7	Contig 249	-1.08 ± 0.27	8.13 ± 0.98	Unknown EST-4	< 0.001	
1	2	Contig 56	-0.41 ± 0.36	9.25 ± 0.99	Unknown EST-5	0.056	
1	1	Contig 71	-0.71	6.21	Unknown EST-6	0.027	
1	1	Singlet 79	-1.08	8.90	Unknown EST-7	0.014	
10					Unsequenced		

^aNumber of microarray elements determined to be significant on the array using the empirical Bayes method to shrink the gene-wise error estimate in cadmium treated vs. control *Daphnia pulex* ($P < 0.05$). control *Daphnia pulex* ($P \leq 0.05$).

^bNumber of microarray elements selected for sequencing based on cadmium response determined in the current study; Eads et al., in review; and Colbourne et al., in review.

^cMean value (± SD) based on all sequenced microarray elements.

^dPermutation tests were performed for clones sharing common putative annotations. These tests included 1000 simulations using the mean expression values to determine the Probability of obtaining the M average for an EST cluster.

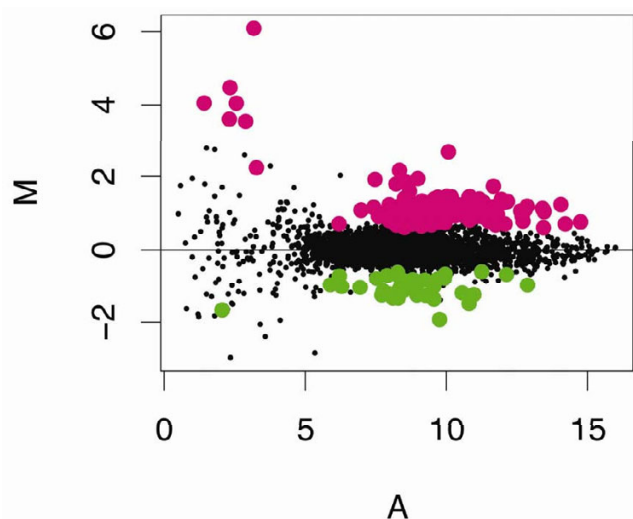


Figure 3

Cadmium effects on gene expression: Regulated genes. Gene expression data from control and cadmium (20 $\mu\text{g Cd/L}$ for 48-h) treated *Daphnia pulex* [GEO:GSE9746]. Microarray elements determined to significantly different ($p \leq 0.05$) using the empirical Bayes (ebayes) method to shrink gene-wise error estimate in cadmium treated vs. control animals are highlighted (up-regulated elements in pink; down-regulated elements in green).

tively spliced transcripts and/or recently duplicated loci that differed in their responses to cadmium.

Functional attributes of responsive genes

The molecular functions and biological processes of *Daphnia* genes responding to the cadmium treatment were investigated based on the putative assignment of Gene Ontology (GO) terms using Blast2GO [44]. Their functions spanned a defined set of gene ontologies (Figure 4a). A total of 23 genes (i.e. EST clusters) were assigned 49 molecular functional terms from the third level of the GO. The predominant terms included structural constituent of the cuticle (16%) and ion binding (14%). Indeed, 14 genes were altogether annotated as binding proteins, of which seven were annotated as metal ion binding (iron, calcium), four as carbohydrate (also listed as protein or chitin) binding proteins, three specifically as oxygen binding proteins (hemoglobins, also listed as tetrapyrrole binding proteins), one protein binding and one nucleic acid binding protein. This last gene (Singlet 469) is a transcriptional regulator whose best match to a *Drosophila* protein is *Similar to Deadpan* (FlyBase ID: FBgn0032741). Six genes also function as hydrolases (Figure 4a).

As well, a total of 17 genes were assigned 74 biological process terms from the fourth level of the GO. The majority of the terms were for metabolic processes ascribed to 11 genes, which included cellular metabolism (14%), primary metabolism (14%), macromolecule metabolism (11%), catabolism (7%), nitrogen compound metabo-

lism (5%), biosynthesis (3%) and regulation of metabolism (Figure 4b). Remarkably, a total of four genes were attributed roles in chitin metabolism, including chitinases and chitotriosidase. Another predominant biological process was related to the localization of cellular components (establishment of localization, transport, protein localization), which were ascribed to five genes. Three of these sequences coded for genes involved in oxygen transport (hemoglobins). A ferritin gene was also differentially regulated during the experiment, which binds iron and is involved in iron regulation and storage [45,46] and was ascribed the functions of cell and ion homeostasis. Finally, the remaining differentially regulated genes with annotations were likely involved in cell communication (such as signal transduction), development, and physiological processes (such as defence response) that specify a reaction to external stimuli, or stress.

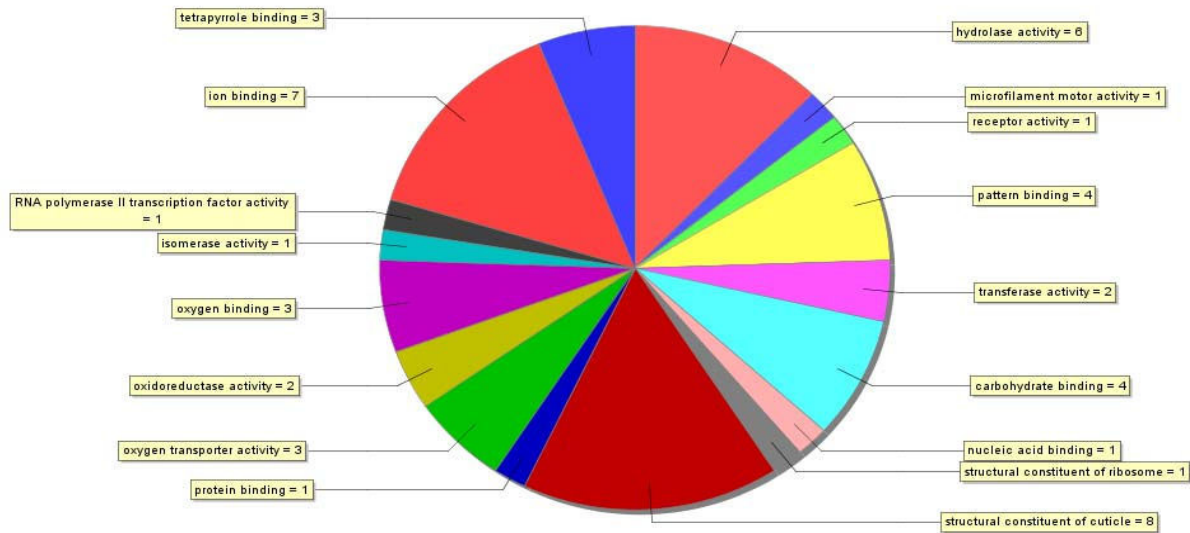
Confirmation of microarray result

Following sequence analysis, expression levels of a subset of genes identified as cadmium-responsive (Table 2) were measured to validate microarray output. This included three cadmium responsive genes (i.e., cuticle protein-2, Contig 257; 2-domain haemoglobin protein subunit, Contig 262; metallothionein, Contig 221) and one gene for which expression was not altered (i.e., serine-threonine kinase, Contig 274). Expression levels were confirmed by Q-RT-PCR (Table 3) using aliquots subsampled from pools of RNA that were used for microarray analysis (technical validation) and RNA collected from repeated, independent biological exposures (biological validation). In all instances, expression levels measured by Q-RT-PCR agreed with microarray output (Figure 5) both in terms of direction (M) and relative magnitude (A) of the response.

Characterization of *Daphnia metallothioneins*

Given its utility as a biomarker, we further characterized the putative *D. pulex* MT, which represents one of the first MT sequences identified from a non-malacostracan crustacean species. A putative MT was identified by the microarray following the sequencing of probes whose expression was up-regulated by cadmium exposure (Figure 6a). A sequence similarity search against the NCBI protein database using Blastx revealed little similarity to other MTs. The closest match at the time was directed to MT from the giant keyhole limpet, *Megathura crenulata* (E-value = 0.036), but this is well below 1×10^{-5} E-value considered to be significant [43]. However, the translated sequence revealed a 59 amino acid protein of high cysteine content (30.5%) that contained no aromatic amino acids or histidine residues, unique features that are hallmarks of MT. The 18 cysteine residues were arrayed in characteristic Cys-xaa-yaa-Cys (1), Cys-x-Cys (6), and Cys-Cys (2) motifs, establishing it as a class 1 MT (Fig. 5; [47]).

A.



B.

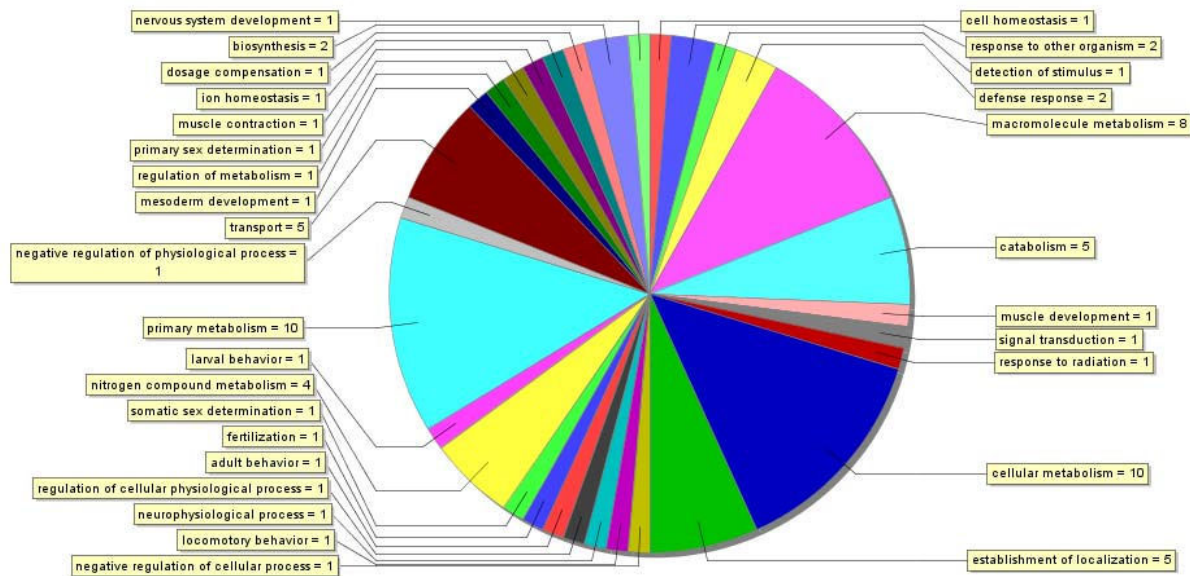


Figure 4

Annotations of genes responding to cadmium. The distribution of gene annotations for the list of 45 *Daphnia pulex* genes (EST clusters) responding to cadmium treatment on the microarray based on results from Blastx searches against the NCBI non-redundant protein database. (A) The assignment of 49 annotations of molecular function to 26 genes from level 3 of the Gene Ontology. (B) The assignment of 74 annotations of biological process to 17 genes from level 4 of the Gene Ontology. Blastx queries recorded the best 5 matches with an E-value threshold of 1×10^{-3} and a minimal value of 33 aligned amino acids. Gene Ontology (GO) terms were assigned to genes using Blast2GO [41] with the following configurations: Pre-eValue-hit filter 1×10^{-3} ; Pre-similarity-hit filter 2; Annotation cut-off 35; GO weight 5.

Table 3: Real-time PCR primer pairs

Contig	Gene	sense	antisense	TaqMan probe	RT primer
257	Cuticle protein	CGTCGCCGATGTGAAATAC	AAGAAGAACCTTTGTGATA GGAATC	GGATATGCCAAGTACCCCGAGT	GGCATCGTATTTTGG A
262	Hemoglobin	TTCAAAGCCAAACCCGAA GC	TTGGCAAATCCGTAATGGA CA	AGAAGCTCTTTTCGGAATTCGCC AACG	AGCGTTCAGGAAATCGT
221	Metallothionein	AAACTACCCAACGGAATCA ACAT	CAGTTGGGTCGCGATTTC	CCACACGAGCATTTCACCTGGC AAC	
274	STK ^a	TTTTTAACAGAACCATCTT TGCCAA	GACATAGTTTTTCAACATTC CTTCACAG	GTGTAAGTACGAGTTAAAGAAAT TATCAGCCATC	CTGATACACAAGGTACGATAA

^aSerine threonine kinase

Nevertheless, the translated amino acid sequence was still quite diverged from other crustacean MTs (~30% similarity), including differences in the common pattern observed at the N terminus, P- [GD]-P-C-C-x(3,4)-C-x-C [48]. Despite this lack of similarity, the number of cysteine residues was identical to other crustacean MTs

and these showed a high degree of conservation when aligned with sequences reported for other arthropods (Table 4). Based on amino-acid alignments with characterized MTs, the *D. pulex* MT formed two coordinative domains (α , C-terminus; β , N-terminus), hinging with the proline residue at position 20 (Fig. 6a, Dpu Mtn1). It

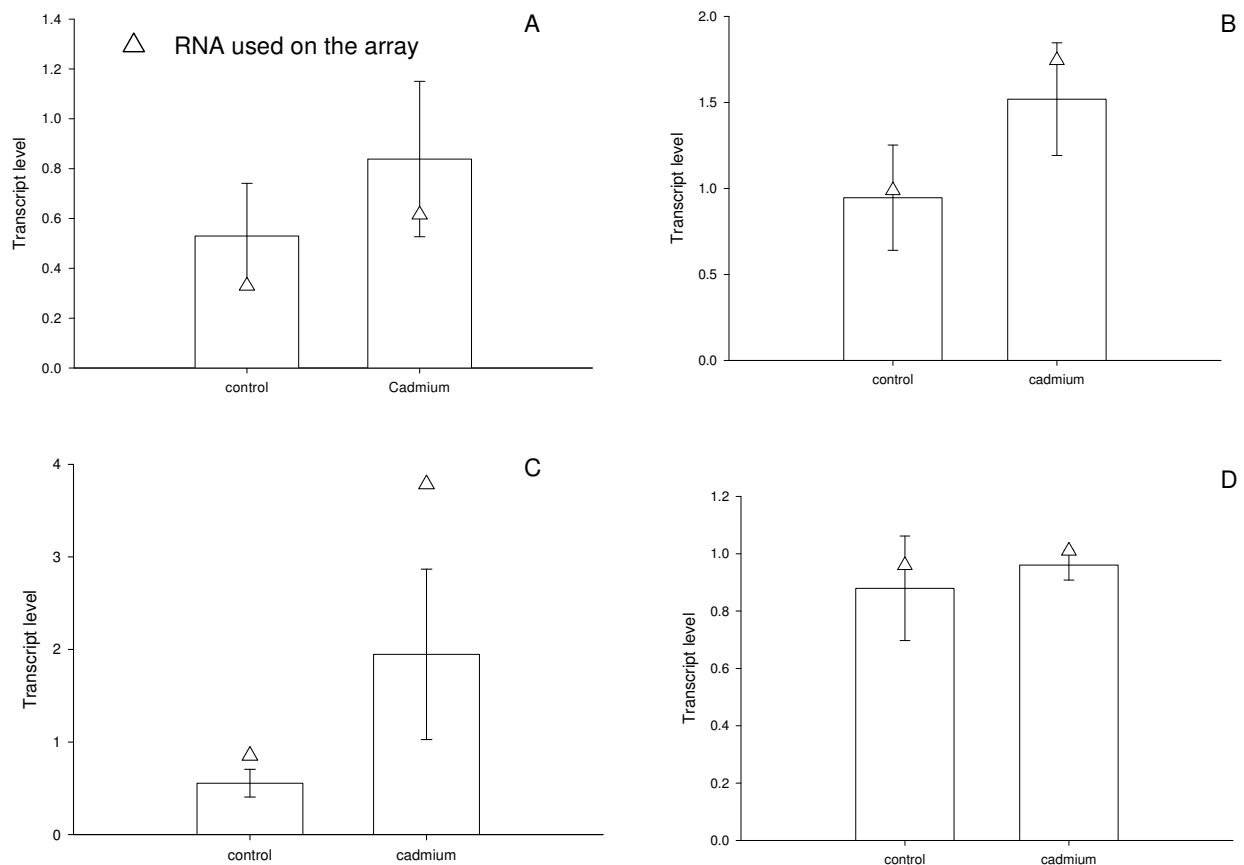


Figure 5

Confirmation of array results. Quantitative-real time (Q-RT)-PCR was used to confirm a subset of genes identified with the microarrays to be regulated in response to cadmium; A) cuticle protein-2, Contig 257, B) 2-domain haemoglobin protein subunit, Contig 262, C) metallothionein, Contig 221 and one gene for which expression was not altered D) serine-threonine kinase, Contig 274 (Table 2). Sequences for primer pairs and probes are provided in Table 3. Expression levels were validated using aliquots sub-sampled from pools of RNA that were used for microarray analysis (technical validation, open triangles) and RNA collected from repeated, independent exposures (biological validation, open bars, mean \pm SD, n = 5).

Table 4: Amino acid alignment of three *D. pulex* metallothionein genes against those from other crustaceans (decapods) and from selected insects^a.

Protein domains	----- Beta -----	----- Alpha -----
Mtn1 <i>Anopheles gambiae</i>	MPCKCCGN-DCKCTSG--CGSGQPCAT--DCKCACASGGCKEKS-- -----GGCCGK--	
Mtn2 <i>Anopheles gambiae</i>	MPCKTCVA-DCKCTSP--NCGAGCGCES--RCTCPCKDGAK----- -----EGCCK--	
MtnA <i>Drosophila melanogaster</i>	MPCP-CGS-GCKCASQ--ATKGSNCNCS--DCKCG----- -----GDKKSA-CGCSE--	
MtnB <i>Drosophila melanogaster</i>	MVCKGCGT-NCQCSAQ--KCGDNCACNK--DCQCVCKNGPK----- -----DQCCSNK-	
MtnC <i>Drosophila melanogaster</i>	MVCKGCGT-NCKCQDT--KCGDNCACNQ--DCKCVCKNGPK----- -----DQCCSK-	
MtnD <i>Drosophila melanogaster</i>	MGCKACGT-NCQCSAT--KCGDNCACSQ--QCQCSCKNGPK----- -----DKCCSTKN	
CuMtn2 <i>Callinectes sapidus</i> CRAB	MPCG-CGT-SCKCGSGKCCCGSTCNCTTTPSKQSCSNDGACGSAC-- -----QCKTSCCGADCK--CSPCPMK-	
Mtn1 <i>Callinectes sapidus</i> CRAB	MPGPCCND-KCVCEG--GCKAGCQCTS---CRCS-PCQKCTSGC-- -----KCATKEECSKTCTKP-CSCCPK--	
Mtn2 <i>Callinectes sapidus</i> CRAB	MPDPCCND-KCECKEG--ECKTGCKCKS---CRCP-PCDKCSSEC-- -----KCTSKEECSKTCSKP-CSCCP--	
Mtn1 <i>Scylla serrata</i> CRAB	-PGPCCND-KCVCEG--GCKAGCQCTS---CRCS-PCEKCSSGC-- -----KCANKEECSKTCSKA-CSCCPT--	
Mtn2 <i>Scylla serrata</i> CRAB	MPDPCCID-KDCKEG--ECKTGCKCKS---CRCP-PCEQCSSGC-- -----KCANKEECSKTCSKA-CSCCPT--	
Mtn1 <i>Potamon potamios</i> CRAB	-PDPCCAEGTCECEG--KCKAGCQCTS---CRCS-PCEKCTSEC-- -----ECKSKEECSKTCTKP-CSCCP--	
Mtn <i>Carcinus maenas</i> CRAB	MPDPCCID-KCECKEG--GCKAGCQCTS---CRCT-PCEKCSSGC-- -----KCTTKEDCCKTCTKP-CSCCP--	
Mtn <i>Eriocheir sinensis</i> CRAB	MPDPCCND-KCECKEG--KCEAGCQCTS---CRCP-PCEKCSSGC-- -----KCGSKEDCCKTCSKP-CSCCP--	
Mtn <i>Portunus pelagicus</i> CRAB	MPDPCCID-KCECKEG--KCEAGCQCTS---CRCP-PCEKCSSGC-- -----KCGSKEDCCKTCSKP-CSCCP--	
Mtn <i>Panulirus argus</i> LOBSTER	MPGPCCID-KCECAEG--KCKSGCQCKS---CTCSTPCDKCTTAC-- -----CCSTKEECSKTCTKP-CKCCP--	
Mtn <i>Homarus americanus</i> LOBSTER	MPGPCCKD-KCECAEG--GCKTGCKCKS---CRCA-PCEKCTSGC-- -----KCPKDECAKTCSKP-CKCCP--	
Mtn <i>Astacus astacus</i> CRAYFISH	MPGPCCND-VCECAAG--GCKTGCVCTS---CRCS-PCDKCTSGC-- -----KCPKDECAKTCSKP-CECCP--	
Mtn1 <i>Daphnia pulex</i> cDNA	MTKD-----CCQKCSGCD---NCKGPNCAQCQPPAAA-- --TCACATGGECKSGNCQCSSTSCPK-SACCK--	
Mtn2 <i>Daphnia pulex</i> cDNA	MPKECV-----RCQNGCTCGD--- DCKCAANCIKCPASSQGETCKGCTPGGCTCGTNCQCGASCVCKASSC CK---	
Mtn3 <i>Daphnia pulex</i> cDNA	MPNACCQN-KCSCGSGCNCQSKCTCGS--GCKGPNGAPQNSA-- --CICATGGG-GCGSDCRCPTSCGCK-TSCCK--	
Mtn3 <i>Daphnia magna</i> cDNA [¥]	MPNVCCQN-NCSCGNGCTCCQSKCTCGS--GCKGPNGAPQNSA-- --CICATGGG-GCGSDCRCPVSCGCK-TSCCK--	
Cysteine residues	0 00 0 0 0 0 0 0 1 1 1 1 1	1 1 1 1 2 2 2 2 2 2
	1 23 4 5 6 7 8 9 0 1 2 3 4	5 6 7 8 9 0 1 2 34
Metal binding	** * * * * * * * *	* * * * *

^aCysteine residues are **bold** for sites that are conserved in crustacean and insect genes (residues 3–10, 23, 24) and underlined when residues are clade-specific. Assuming a correct alignment, residues 2, 12, 13, 17, 19 and 20 are conserved in all (or most) of the crustacean genes, whereas residues 14 and 22 are conserved in the decapod genes. The residues 15, 16, 18 and 21 are exclusively found in the *Daphnia* genes (representing branchiopods) and based on the limited sampling of insect genes (from two dipterans), residues 1 and 11 are excluded in the crustacean genes. Domain partitions and metal bind residues (*) are based on experimental evidence from decapod proteins (Valls et al. [41]). ¥ Poynton et al. [19] provides full translation.

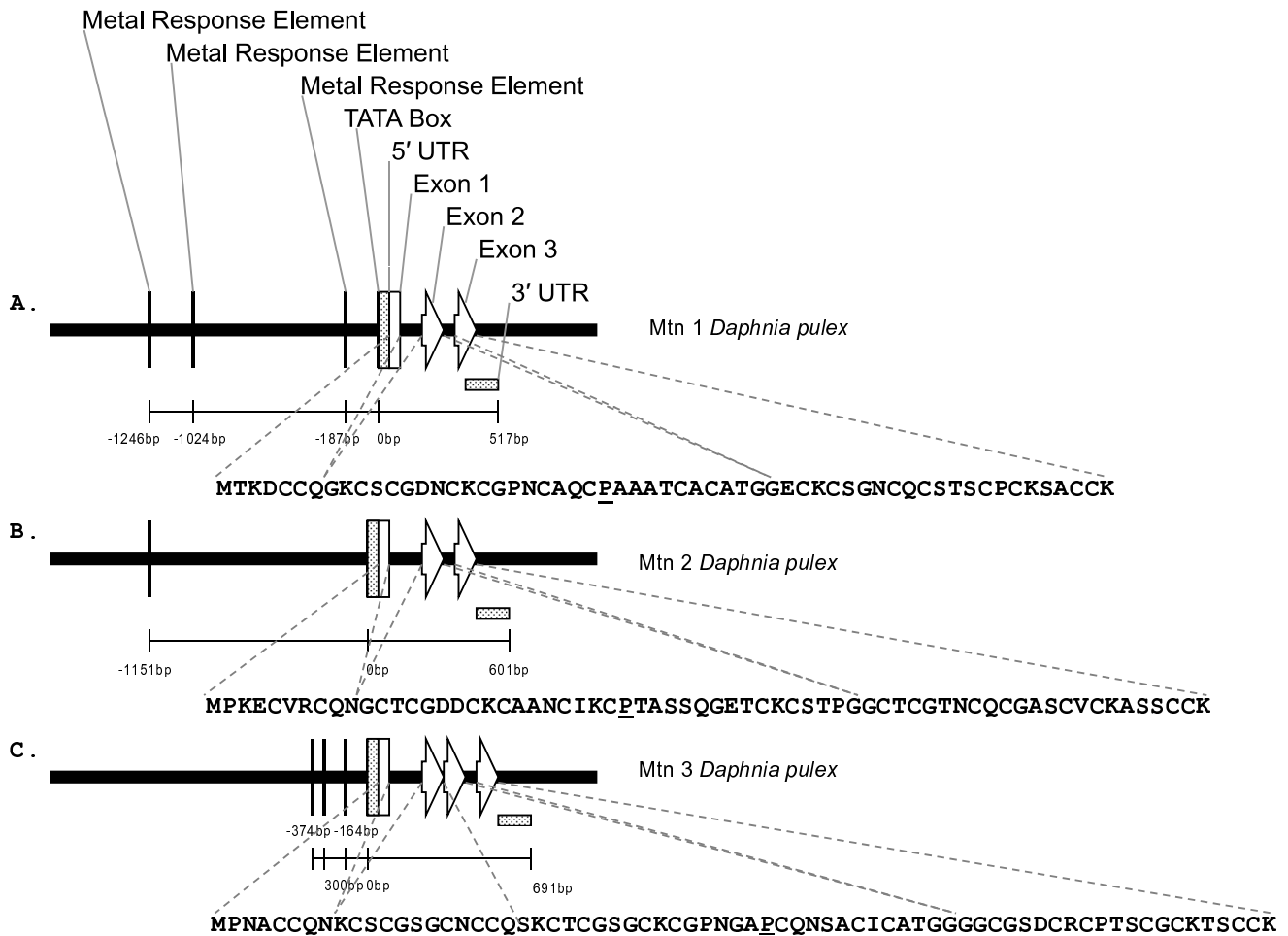


Figure 6
Metallothionein gene models. *Daphnia pulex* metallothionein gene models. A) Dpu Mtn1 [GenBank:EU307302], B) Dpu Mtn2 [GenBank:EU307303], and C) Dpu Mtn3 [GenBank:EU307304].

should be noted that during the preparation of this manuscript, an MT gene transcript and translated amino acid sequence was reported for another daphniid, *D. magna* [19].

Following this preliminary identification of the MT gene, further cDNA sequences were recognized from within a growing *D. pulex* cDNA sequence database [49], which was created as part of the annotation process for the *Daphnia* genome sequencing project (Colbourne et al., in prep). Using tBlastx, a total of 9 assembled EST clusters were discovered, representing transcripts from three genomic loci. Using Blastn, the regions within genomic sequence scaffolds were identified that matched the metallothionein cDNA sequences (Figure 6). The locus corresponding to the cDNA identified on the array, called Dpu Mtn1 [GenBank:EU307302] (Fig 6a), consists of three exons (86 bp, 90 bp, 168 bp) and two introns (107 bp, 76

bp). The 5' untranslated region (UTR, at least 64 bp) is fully contained within the first exon, while the 3' untranslated region consisted of 103 bases. The region upstream of the 5'UTR contains three putative regulatory elements showing exact sequence similarity with the core consensus sequence of metal responsive elements (TGCRNCS, where R is A or G and S is C or G; [50,51]). This flanking region also contains a TATA box 20 bases upstream of the putative transcriptional start site. Within the 5'UTR, the area immediately upstream of the ATG codon agrees with Kozack's rules of ribosome binding. The eukaryotic polyadenylation signal (AATAAA) is found in the 3'UTR. The second identified locus called Dpu Mtn2 [GenBank:EU307303] (Fig 6b) encodes a protein that consists of 66 amino acids. Like the first locus, this gene is composed of three exons (70 bp, 99 bp, 223 bp) and two introns (141 bp, 70 bp). Its 5'UTR is at least 42 bases and contained within the first exon, while the 3'UTR consists

of 152 bases. The region upstream of the 5'UTR contains a single putative metal responsive elements (-1151 bp). Finally, the third locus called Dpu Mtn3 [GenBank:EU307304] (Fig 6c) codes for a protein that consists of 70 amino acids that is almost identical to the amino acid sequence provided by Poynton et al. [19] for *D. magna* (Table 4). The 5'UTR of Dpu Mtn3 is at least 66 bases in length and the 3'UTR consists of 163 bases. Three putative metal response elements are found upstream of the 5'UTR (-164, -300, -374 bp). Unlike the other daphniid MTs, this gene is composed of four exons (88 bp, 39 bp, 84 bp, 232 bp) and three introns (131 bp, 59 bp 62 bp).

Phylogenetic analysis of crustacean metallothioneins

The amino acid sequences of *D. pulex* MTs were aligned and compared to 19 amino acid sequences from 13 other crustaceans, including the recently identified *D. magna* sequence, and two insects to determine the phylogeny of the protein within Crustacea (Table 4). These comparisons were made difficult by the lack of sequence conservation among the arthropod MTs (12.7% similarity). The greatest sequence divergence was found among the insect genes obtained from fully sequenced genomes, which averaged 72.3%. By contrast, the three *D. pulex* genes averaged 56.1% sequence divergence. This divergence was measured from genes that were also identified from a well characterized genome. Considering that the average divergence among the other entire crustacean MTs was only 27.4%, most of the characterized malacostracan genes were likely homologues or recent duplicates among their representative genomes. Clearly, further genome-wide investigations among the Crustacea will uncover paralogous loci that will broaden the phylogenetic account of this protein family.

While little similarity was observed between primary amino acid sequences, there was a great deal of conservation of the cysteine residues. The MTs contained an average of 30.4% cysteines; crustacean genes contained 3.5% more cysteines than the representative insects. This conservation translates into structural homologies that coordinate the disulfides within the protein domains, which are responsible for the molecules' unique metal binding properties. From the amino acid alignment, 10 cysteine residues were found to be largely preserved in both the insect and crustacean genes (Table 4). Some exceptions were observed. In particular, our alignment suggested that two of the three *Daphnia* genes lost the conserved residues 3–5. The two known copper-binding MTs also showed a single cysteine instead of two at the C-terminus of the proteins. A second class of conserved residues was composed of cysteines that are known to bind metals in malacostracans. Yet, these cysteines are generally absent in the insects. Here, most of the *Daphnia* genes shared six of the

eight residues. However, they all lacked residues 14 and 22. The *Daphnia* MTs contained a third class of conserved residues, whose four members are exclusive to this set of proteins (Table 4). Finally, the insects' α -domains are substantially reduced relative to the Crustacea. As a result, only two cysteine sites form the fourth class of residues, which are largely missing from the aligned proteins of the other taxa.

From this alignment, phylogenetic trees were constructed by neighbor-joining and maximum likelihood methods. The maximum likelihood tree was inferred from 77.6% fully resolved and 5.9% partly resolved quartets. Although this latter tree was not completely resolved (tree not shown), the results from both methods were congruent and showed three strongly supported monophyletic groupings with support values at nodes > 95% (Figure 7). The first group consisted of the cadmium-binding malacostracan MTs, including the presumed recent gene duplicates in the *Callinectes* and *Scylla* genomes. Both trees failed to place all the crab genes above a common node at the exclusion of the lobster and crayfish loci. In fact, a weakly supported node united the *Potamon* and *Panulirus* genes in the maximum likelihood tree (57%). The second group consisted of the three *Daphnia* MTs stemming from the base of the branch leading to the cadmium-binding malacostracan genes. Although this monophyletic grouping was strongly supported, 31% of the maximum likelihood bipartitions united Dp Mtn1 with Mtn2 and placed Dp Mtn3 at its base. The third group was composed of three *D. melanogaster* MTs, which lacked the *Drosophila* MtnA gene to form an intra-genomic clade, as seen in *Daphnia*. The excluded gene was instead placed at an earlier branch point of the insect clade, which included two intervening *Anopheles* genes (Figure 7). Finally, the copper-specific *Callinectes* MT was positioned at the root of the insect clade, along with the *Drosophila* copper thionein MtnA locus and the *Anopheles* Mtn1 gene. The relative positions of these genes were unresolved on the maximum likelihood tree, whereas 59% of the bipartitions supported the placement of the *Anopheles* Mtn2 gene at the base of the group 3 *Drosophila* genes.

Discussion

In this paper we describe studies with newly developed *D. pulex* cDNA microarrays investigating the expressed mRNA responses of daphniids to the environmental stressor cadmium. Although microarrays have been successfully used to identify patterns of genes responding to metal exposure in organisms for which there is an abundant amount of sequence information [52,53], it is often difficult to relate the identified genes to expressed phenotypes. These challenges exist in part because the environmental contributions to phenotype in most common laboratory species are poorly understood. However, as

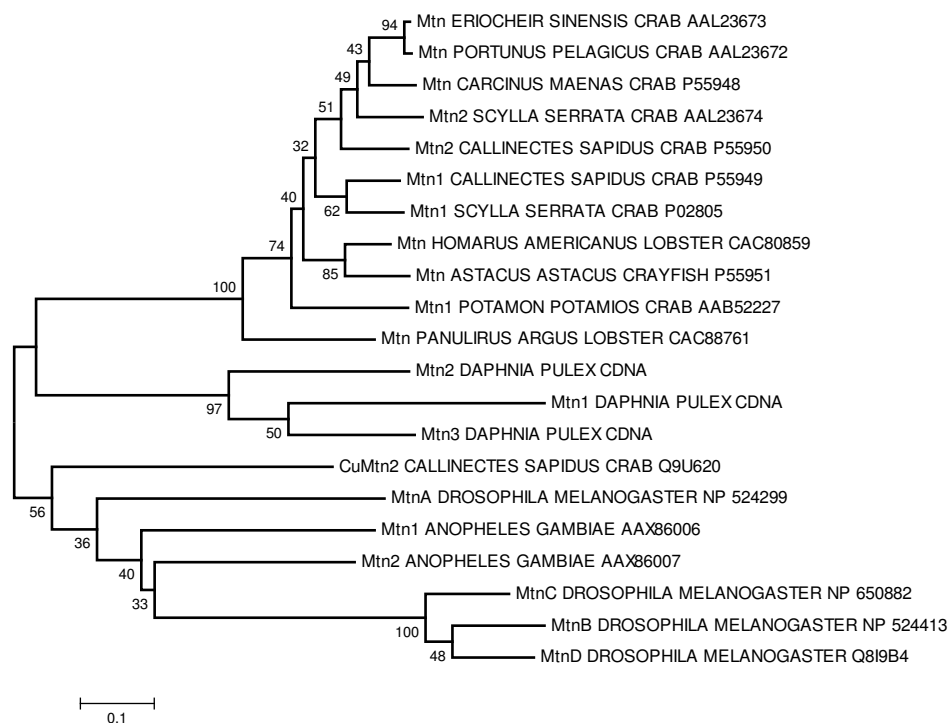


Figure 7

Phylogeny of crustacean metallothioneins. Neighbor-Joining tree constructed from an amino acid alignment of the three *Daphnia pulex* metallothionein proteins plus 18 sequences from 11 selected insect and crustacean taxa using the E-INS-I algorithm implemented by MAFFT [92]. Gaps within alignments were ignored in pairwise comparisons of the sequences and the genetic distances were corrected by the Poisson distribution model. The NCBI accession numbers are listed beside each taxonomic gene designation. Percent bootstrap support for nodes are shown, which are derived from 1000 pseudo-replication of the data. The *Daphnia* sequences were deposited at GenBank under the accession numbers Dpu Mtn1 [GenBank:EU307302]; Dpu Mtn2 [GenBank:EU307303]; and Dpu Mtn3 [GenBank:EU307304].

noted by Gracey et al., [12], abundant sequence data are not a pre-requisite for gene-expression profiling. Thus, one aim of the present study was to develop and apply microarray technologies to the ecologically tractable aquatic micro-crustacean, *D. pulex*. These studies revealed patterns of gene response that provide insight into the biological and potentially toxicological responses to this important environmental contaminant. The data emphasize the potential of this approach as a tool for discovering genes regulated by an environmental stressor, but in a species for which gene expression profiles can be interpreted in context of individual and population-level effects. Associations with cadmium induced phenotypes, identification of genes reported in the literature to be regulated following cadmium exposure, and independent validations via Q-RT-PCR provided support for their utility in this role. In addition, these studies led to the discovery *D. pulex* MT mRNAs and gene sequences, providing an important new biomarker for *Daphnia* studies, while also extending the phylogeny of this class of genes. The successful identification of the metal binding protein, MT, is highlighted, not only due to its critical biological functions, but also because the blind microarray approach

provided advantages over cloning methods that require significant sequence similarity to isolate genes.

We identified the acute (48-h) and chronic (21-d) responses of our *D. pulex* isolate to cadmium to better define sub-lethal exposure concentrations and cadmium-induced phenotypes. The present study is in agreement with those of others [16,21] that indicate *D. pulex* is one of the most acutely sensitive aquatic species to cadmium [54]. Reported LC50 values are comparable to data published by others (i.e., range from 46 to 90 $\mu\text{g Cd/L}$; [16]). As others have noted, chronic exposure of *D. pulex* to cadmium reduced individual fitness parameters, such as length and lipid-ovary indices, and inhibited population endpoints, such as number of clutches, cumulative reproduction, and per capita birth rate (Table 1; [55,56]). These endpoints have been quantitatively linked to population success [57].

For example, storage of lipids, vitellogenesis, and ecdysis (i.e., molt) are physiologically and chronologically associated with reproductive success and survivorship [57]. Adipocytes are incorporated into the ovary during egg

development providing energy for reproduction. Parthenogenetic embryos are then transferred into the brood chamber that is contained within the carapace for further development and the process ends with newborn *Daphnia* released into the water column during the molt. Also, molt directly relates to growth, as *Daphnia* size is constrained by the carapace. Therefore, decreased length, lowered lipid-ovary indices, and fewer clutches comprised of fewer offspring can be explained by impaired vitellogenesis and/or ecdysis. Cadmium has been reported by others to inhibit ecdysis [58] and vitellogenesis [59,60] in arthropods.

Cadmium may also decrease daphniid size by limiting calcium intake. Calcium content is directly related to the mass of the carapace and, subsequently size [61] and cadmium is known to impair calcium uptake and metabolism via substitution [62]. However, several studies have shown that while the period of increased calcium flux during molt can result in increased cadmium uptake [63,64], cadmium does not affect calcium accumulation or content [64]. Yet interestingly, our studies revealed two cDNAs that responded to cadmium that represent putative calcium binding proteins (singlet 97 and contig 220), suggesting that cadmium may affect calcium regulatory pathways in more subtle ways that contribute to its overall effects.

The present study is one of the first to apply microarray technologies to *D. pulex*. These "blind arrays" identified 99 elements for which gene-expression was positively regulated and 30 elements that were negatively regulated by cadmium. However, sequence analysis revealed that these probes included several redundant features. When they were assembled into EST contigs to isolate unique features, 30 up-regulated and 12 down-regulated genes were identified (Table 2). Given the redundancy that existed on the array and the 'blind' approach-employed, the observation of redundant, cadmium-regulated elements (e.g., cuticle protein) was reassuring. In addition, the number of unique elements regulated following exposure to cadmium only represents a very small percentage (~2%) of the estimated 1,550 unique elements on the array, which is similar to other microarray studies involving sub-toxic cadmium exposures (2%, [52]; 1%, [65]; 1–3%, [66] and as suggested by Andrew *et al.* [52] is consistent with expectations for low, non-toxic exposures that tend to induce specific pathways.

The *D. pulex* microarrays and their genomics database have proven to be insightful experimental tools for identifying genes regulated by cadmium. Yet, relating the gene sequence to putative gene function is made difficult by the excessively large phylogenetic distances between *Daphnia* and its closest relatives among the classical genomic

model systems. For example, the best model system to *Daphnia* with extensive functional genomic information is *Drosophila*, which last shared a common ancestor some 600 M years ago. There are also few ESTs that share sequence conservation with proteins in public databases; crustacean proteins represent only 0.1% of 6.9 million records in the NCBI taxonomic database. Of the 42 unique cadmium-regulated genes identified, only three were homologous to known *Daphnia* genes. Likewise, almost 36% of the elements identified were ESTs showing no sequence similarity with known proteins. Gracey *et al.* [12] reported similar successes with 'blind' microarrays, as approximately 40% of identified elements were unidentifiable ESTs. However, as these authors noted, novel ESTs may represent untranslated regions of previously identified genes or multiple distinct regions of unknown genes. These possibilities cannot be dismissed; we estimate that 1/4 of the orphan genes are unknown because of sequences not extending into recognizable functional domains of the gene [67]. The remaining 27 elements regulated in response to cadmium exposure on the microarray were putatively identified by similarity with known proteins.

The genes regulated in response to cadmium exposure identified in these experiments provided several measures of the utility of this approach for gene discovery and its reliability in identifying biologically sensible patterns of gene regulation. For example, many of the genes responding to cadmium exposure were part of common physiological pathways (i.e., ecdysis, metal detoxication) and few were indicative of general stress response (e.g., heat shock proteins, heme-oxidase) or overt cellular toxicity (e.g., housekeeping genes).

Ecdysteroid-responsive genes and other molt related genes that included chitinases 1 and 2, chitotriosidase, BcNCP 14.9, cuticle proteins 1 through 5, and singlet 251, comprised the majority of genes regulated in response to cadmium exposure (Figure 8, Table 2). This finding is critical given that altered expression of these genes, which are associated with the exoskeleton, provides a direct mechanistic link with demographic endpoints and individual fitness parameters that indicated ecdysis (i.e., molt) is impaired in *Daphnia* following cadmium exposure (Figure 1, Table 1). Deep sequencing of cDNA libraries constructed from RNA isolated from cadmium-exposed *D. pulex* provided support of these findings (Colbourne JK, personal communication). Cadmium derived libraries were enriched with genes that were identified as structural constituents of the cuticle. This observation suggests that ecdysis and molt related regulatory pathways are generally influenced by cadmium, which has been observed by others [63,64]. It is interesting that short-term exposures to cadmium of the *Daphnia* used in the array experiment

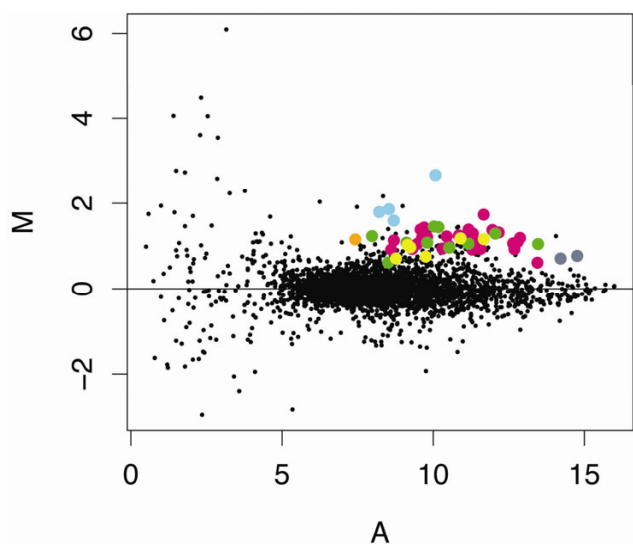


Figure 8

Cadmium effects on gene expression: literature reports. Gene expression data from control and cadmium (20 $\mu\text{g Cd/L}$ for 48-h) treated *Daphnia pulex* [GEO:GSE9746]. Microarray elements, which are known from the literature to be regulated by cadmium, are highlighted: Cuticle protein, pink; Hemoglobin, green; Metallothionein, blue; Ferritin, orange; Chitinase, yellow; Opsin, grey.

provided patterns of response that were meaningful in interpreting demographic experiments that involved longer-term exposures to lower cadmium concentrations. This observation raises the possibility that an approach such as this could provide biomarkers that serve as early indicators of chronic exposures, which is an area we are currently exploring, but beyond the scope of the current manuscript.

These experiments also identified altered expression-levels of genes known to be regulated by cadmium (Figure 8). These included the metal binding protein, MT, which plays a central role in cadmium detoxication and is known to be induced by cadmium [19,41,68]. In addition, the microarrays identified a gene responsible for oxygen transport and iron metabolism in *Daphnia*, hemoglobin [69]. Expression of hemoglobin is known to be induced by anoxia, limiting iron supplies, and cadmium [19,70,71; unpublished results). As discussed above, the arrays also identified structural components of the cuticle and associated regulatory components (i.e., chitinase) and chitinase activity has been shown to positively correlate with cadmium concentrations in previous studies [19,72]. Likewise, opsin was found to be regulated in response to cadmium on the microarrays, and cadmium has been shown to interfere with photo-behavior in daphniids, which is a complex genetically linked trait involving kairomone signalling, photoreceptor detection (e.g., opsin, [73]), and negative phototaxis [74,75]. At the

cellular-level cadmium is known to induce oxidative stress [28]. Glutathione S-transferase (GST), a phase II enzyme that catalyses the conjugation of reduced glutathione with electrophilic compounds, is an important antioxidant defence enzyme that works to detoxify the products of oxidative stress [76] and was up-regulated on the array. GST gene transcription [19] and enzyme activity [77] have been shown by others to increase following exposure to cadmium.

Metallothionein induction is the principal adaptive response associated with organism survival during exposure to elevated cadmium concentrations [41,78]. The discovery of *D. pulex* MTs highlight the advantages of this approach in identifying genes regulated by cadmium. While the conservation of cysteine residues ensures thiolate cluster formation, metal binding and provides both the structural and functional basis to identify these genes as MTs, cloning methods such as hybridization or amplification using degenerate oligonucleotides were ineffective because of its lack of similarity with other known crustacean MT sequences (Figure 6). The present study is the first to provide genomic DNA sequence information for a non-malacostracan crustacean, the first to report the promoter region of any crustacean MT, and one of the few to identify the metal-responsive element (MRE) consensus core sequence in any invertebrate. MREs were identified in the 5' flanking region of *D. pulex* MTs (Fig 6). Both *Dpu Mtn1* and *Dpu Mtn3* contained MREs within 200 bp of the transcriptional start site. The promoter region of the sea urchin MT gene was found to contain two MREs residing within 300 bp of the start of transcription [79,80]. There was only genomic (primary) sequence provided for one crustacean MT (i.e., green shore crab, *Carcinus maenas*) in the entire NCBI repository, but it lacked any flanking sequences (i.e., upstream and downstream regulatory regions). Green shore crab MT has a structure common to *Dpu Mtn1* and *Dpu Mtn2* with three exons separated by two introns [81]. In fact, similar architectures have been reported for the urchin, trout and several mammals. Also, *D. pulex* have much smaller intronic regions (Figure 6), likely a consequence of its compact genome size (~200 Mb). Although the architectures of these genes were similar to others, little sequence similarity was observed within their coding sequence for any known MT.

The translated daphniid sequences (i.e. *D. pulex*; *D. magna*, Poynton *et al.* [19]) also showed little similarity to other MT genes (Table 4, Fig 7), including other crustaceans. This includes the N-terminal crustacean motif (P-[GD]-P-C-C-x(3 or 4)-C-X-C; [41,48]. However, since daphniid MTs represent the first MTs isolated from a non-malacostracan crustacean, the lack of phylogenetic representation in the NCBI repository could account for some these differences. There was considerable similarity in

terms of number and distribution of cysteine residues. All crustacean MTs sequenced to date are comprised of 18 cysteine residues with the exceptions of the copper MT isolated from the green crab [41], *D. magna* MT [19] and Dpu Mtn3 identified in this investigation that each contain 19 cysteines. The conservation of cysteine residues is expected, given that they are responsible for a great deal of the tertiary structure and metal-binding functions of the protein [78]. In fact, Valls *et al.* [41] has suggested using the binding properties (e.g., stoichiometry of metal-MT species) to strengthen current classifications that are based on phylogeny.

The current classification procedure for MTs [82] is based on phylogenetic relationships rather than amino acid composition and cysteine content. This nomenclature protocol was established to better identify related functional properties in which MTs are divided into families based on evolutionary conservation. In this classification scheme, crustaceans constitute a single family, given the name number three, which is comprised of three sub-families (crustacean one, c1; crustacean two, c2; crustacean, c), where c1 and c2 genes each constitute monophyletic clades. The third family termed, c, is reserved for crustacean MTs that are different from these others. *Daphnia pulex* MTs do not align well with malacostracan type 1 or type 2 MTs and thus, according to this nomenclature are designated Dpu Mtn1, 3, c; Dpu Mtn2, 3, c; and Dpu Mtn3, 3, c.

Conclusion

In summary, the development of *D. pulex* cDNA microarrays and associated sequence information has provided a useful first generation functional genomic tool for examining biological responses of this key sentinel species to environmental agents and other stressors. Treatment of *D. pulex* with sub-toxic levels of cadmium revealed a specific pattern of gene expression changes that provide new insights into their biological and toxicological responses to this environmental contaminant. Moreover, microarray responses to cadmium led to the discovery of *D. pulex* MTs, whose gene structure and cysteine content clearly place it in this gene family, but whose sequence divergence reveals that classical cloning and sequencing techniques based on similarity were likely to fail. Further identification of *D. pulex* genes that are responsive to various experimental treatments through use of these genomics tools will provide new and important insights into their biology. Advances in *Daphnia* genomics will enable the further development of this species as a model organism for a wide variety of biological investigations.

Methods

Animals and cadmium exposure

Daphnia pulex (subclade arenata) used in this study were obtained from isoclonal laboratory cultures of an isolate collected from an ephemeral pond near the Pacific coast in Oregon. This pond is found north of Florence on the east side of highway 101 at milepost 201 in Douglas County. *Daphnia pulex* subclade arenata is a member of the *Daphnia pulex* complex [83] and our isolate is from the same population as the strain whose genome has been sequenced by the Joint Genome Institute as part of the *Daphnia* Genomics Consortium initiative [84]. The *Daphnia* were housed in 3L borosilicate glass beakers (20 per beaker) held inside an environmental chamber at a constant temperature ($20 \pm 1^\circ\text{C}$; [85]) and photoperiod (16:8 light-dark). Organisms were maintained in nanopure water reconstituted to moderate hardness [86] and renewed weekly. They were fed daily *Ankistrodesmus falcatulus* at a rate of 75,000 cells/mL. Our pre-experimental procedure, described in Folt *et al.* [87], controlled for maternal effects in acute toxicity tests, demographic experiments and batch exposures. For these experiments, neonates (< 24 h old) were isolated from maintenance cultures one generation prior to metal exposure. These organisms are referred to as 'brood females', which were synchronized with respect to time of maturity for producing neonates for metal experiments.

Acute toxicity tests

Acute (48-h) toxicity tests were conducted with cadmium according to recommendations given by the United States Environmental Protection Agency with slight modifications [10]. Test solutions were prepared immediately prior to use with culture media from stocks made with CdCl_2 (analytical grade, Sigma Chemical, St. Louis, MO, USA) dissolved in deionized water. Test concentrations nominally ranged from 1 to 150 $\mu\text{g Cd/L}$. Toxicity tests employed a completely random design consisting of five or six metal treatments and a control group arrayed in two-fold serial dilutions. Ten neonates (< 24 h old) were randomly placed into a 40 ml glass exposure chamber containing 30 ml of test solution. Four replicate exposure chambers were employed per treatment or control group. Daphniids were not fed during tests. Mortality was assessed for individuals in each container after 48-h exposure. An individual was labeled dead if it was unresponsive to gentle prodding with a pipette tip. These tests were repeated and results combined to determine lethal concentrations (LC x values, where x equals a given percent mortality) estimated from the probit transformed concentration-response curves.

Demographic experiments

Twenty-one day life-table experiments followed a completely randomized design as given in U.S. EPA [11]. A

single organism was randomly placed into a 120 ml exposure chamber. Ten replicate chambers were used per test group. Treatment groups consisted of 0.25, 0.5, 1, 1.75, 2.5 µg Cd/L and controls. Test waters were renewed every other day, at which point general water quality parameters of temperature, dissolved oxygen, conductivity, and pH were monitored. Water column metal concentrations were measured at the beginning and end of each test. Mortality and reproduction were observed over the duration of the test and endpoints included age to first reproduction, cumulative reproduction, neonates per adult, and percent survival as per Chen and Folt [88]. Length measurements [89] and lipid ovary indices [57] were taken at the end of the experiment using an Olympus BX40 microscope fitted with a digital camera (Hitachi KP-D50) driven by Scion Image software (V. 1.63).

Batch exposure

The *Daphnia* used for microarray experiments and for validation tests by RT-PCR were exposed in repeated experiments in batches of 50 adult organisms plus their offspring per 3.5-l exposure chamber. The offspring of these animals were not discarded to increase biological variability and increase opportunities for gene discovery. Batch number was optimized to provide adequate sample mass for molecular evaluation (e.g., 1 adult *Daphnia* equals 1 µg of total RNA). Animals were introduced as neonates (< 24-h old) and cultures were maintained until they produced their third clutch (15-d). These experiments, which provided source material for microarray experiments and validation tests, included short-term (48-h) independently replicated (n = 3) exposures to non-lethal concentrations of cadmium (LC01, 20 µg Cd/L) and control conditions.

Gene response profiles

Microarray construction

Working in the absence of abundant a priori DNA sequence data for *D. pulex*, we constructed a 3,842 element microarrays using 3,602 PCR-amplified cDNA and 240 control probes. A detailed description of this microarray platform is archived at the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus under the accession number [GEO:GPL6195], series [GEO:GSE9746]. Briefly, cDNA for seeding the amplifications were obtained by directly transferring into the reactions 5 µl of cDNA bacterial transformants, which were grown in 1.2 ml of 2X YT and 0.005% chloramphenicol in 96-deepwell plates for 24 hours at 37°C. The arrayed cDNA clones were randomly picked from two high-quality *D. pulex* cDNA libraries (Creator SMART, Clontech). For details about the libraries and results from quality assurance tests, see Colbourne *et al.* [67]. The amplifications were conducted in 100 µl reactions containing 1× Taq buffer (Eppendorf), 0.2 mM dNTPs, 0.2

µM primers (Fwd. 5'-GTGTAACGACGGCCAGTAG 3' and Rev. 5'-AAACAGCTATGACCATGTTAC 3'), 5 U Taq (Eppendorf). The PCR cycling conditions involved a 3 minute initial denaturation step at 94°C followed by 35 cycles of 94°C, 54°C and 72°C each for 1 minute. The products were purified using the Multiscreen-PCR 96-well system (Millipore) on a Biomek FX liquid handling robot (Beckman). Our pilot experiments for this section of our workflow indicated that this purification system recovered 80–90% of the original sample and introduced no impurities that interfered with immobilizing DNA onto glass. The quality of PCR amplifications was visually inspected by agarose-gel electrophoresis and their number and size were recorded using Kodak's 440cf scanner and 1D imaging software (v.3.6). The sample concentrations were determined by 96-well microplate spectrophotometer (Molecular Devices, SpectraMax 190) and adjusted to 50–200 ng/µl for printing. The PCR amplifications were classified as having produced a single high yield product (3,238; 91%), as failures (163 reactions; 5%), as having produced more than one amplicon (102 reactions; 3%) or as being weak (49 reactions; 1%). The DNA yields averaged 409.7 ng with a standard deviation of 210.8 ng. Printing was achieved using an Omnigridd 100 robot (GeneMachines). The cDNA were spotted in tandem on GapsII amino-silane slides (Corning) in 3× SSC and 1.5 M Betaine buffer using Stealth Micro-Spotting Pins (Tel-echem) at 20°C and 65% humidity. The cDNA was fixed to the microarray slides by baking at 85°C for 3 hours. To achieve minimal signal to background ratios averaging 40–50 fluorescence units, the slides were post-processed by washing in 5 × SSC buffer with 0.1% SDS at 55°C for 5 minutes, rinsing in water at room temperature for 2 minutes, denaturing the DNA in water at 95°C for 4 minutes, then rinsing in water at room temperature for 30 seconds. The slides were finally rinsed in isopropanol at 4°C and dried by centrifugation at 500 g for 5 minutes before being stored. Slides were printed in groups of 100 or 120, where 95% of the slides were free of defects. Negative controls were included, designed to detect potential problems. To test for the cross-contamination of probes, printing buffer containing no DNA was first deposited at the beginning of each subarray. Printing buffer was also printed following positive control DNA (coding cytochrome c, cytochrome b, actin and ferritin) at both the beginning and end of the subarrays. To test for the effect of template DNA during the hybridizations, the product of intentionally failed PCR reactions with template DNA but no primers were printed. Finally, amplified DNA from *Arabidopsis* and lambda phage plus bacterial spiking controls (Ambion) was also included. Based on random sequencing of 619 cDNAs probes, and post hoc sequencing of an additional 927 cDNA probes, gene-redundancy on the array was calculated to be roughly 57%. Thus, there were likely ~1,550 interrogated unique genes on the array.

Labeling and hybridization

Microarrays were used to discover genes that were differentially expressed following sub-lethal cadmium stress in *D. pulex*. They utilized RNA isolated from three independent and concurrently replicated exposures of *Daphnia* to cadmium and control conditions, applied to three replicate microarrays using a standard control vs. treated design that included dye swaps. This design provided three replicates of metal exposed and control *D. pulex*, with the technical variability of hybridization kinetics captured on each slide across duplicate probes.

Total RNA was isolated from *Daphnia* that were acutely (48-h) exposed to 20 µg Cd/L (treated) and from their genetic clones reared under standard (control) conditions as given above for batch cultures. Animals were directly placed in lysis buffer and RNA was extracted using Qiashredder columns and RNeasy kits (Qiagen). DNA contamination was removed by DNase treatment (Ambion) and RNA was quantified by spectrophotometry (Nanodrop Technologies); and quality determined with a Bioanalyzer 2100 (Agilent). For each sample, 10 µg of total RNA was reverse transcribed with random hexamer primers using SuperScript II (Invitrogen) and an overnight incubation, which included aminoallyl-dNTPs. Following reverse transcription and clean up (alkaline hydrolysis and Qiaquick columns, Qiagen), cDNA samples were coupled to Alexa Fluor dyes (555, 647), using amino-allyl labeling methods and alternating direction with replicate arrays (i.e., dye-swap; [90]). The amounts of dye incorporated cDNA were measured by spectrophotometer using the acceptability cutoffs of > 200 pmol of dye incorporation and a nucleotide to dye molecule ratio of > 50 [91]. The labeled samples were then pooled according to treatment comparisons (e.g., control vs. Cd treated), dried, and resuspended in hybridization buffer (50% formamide, 5× SSC, 0.1% SDS, 20 µg SSDNA, and 20 µg of poly(A)-DNA). The hybridization solution (containing the dye labeled samples) was placed on the microarray and hybridization was accomplished overnight at 42°C in a specially fitted hybridization chamber (Corning). Following hybridization, the glass slides were washed successively in a low stringency solution (1× SSC, 0.2% SDS) at 42°C for four minutes, high stringency solution (0.1× SSC, 0.2% SDS) at room temperature for four minutes, twice in 0.1× SSC for 2.5 minutes, and finally they were dipped in water, dried and stored in the dark until fluorescence was measured [91].

Gene Expression Analysis

The fluorescent signals were scanned and the array data were extracted using GeneChip Scanner 3000 software version 5.1 (Axon). The data were coupled to the array template using the GeneChip Operating System (GCOS) software, V. 1.4. For each microarray, LIMMA functions as

defined in [92-94] were used to create unique, background subtracted, LOESS (i.e., locally weighted least squares regression) normalized, log expression values, which accounted for duplicate spots using gene-wise linear models fit to expression data. The LIMMA derived log expression values were fit to a linear model based on treatment, which moderated the resulting t statistics using the empirical Bayes method. Probes with a p-value ≤ 0.05 were deemed significant and selected for sequencing. A 5' expressed sequence tag was generated for the majority of the differentially expressed clones. These sequences were analyzed and assigned putative gene function annotations based on sequence similarity searches against NCBI protein and insect genome databases [67]. Analyses of the functional grouping of genes based on Gene Ontology (GO) assignments were performed using Blast2GO [44]. For those probes sharing common putative annotations, we performed permutation tests to estimate the likelihood that random processes would place these highly represented annotations on our list of differentially expressed cDNA and to establish p-values on certain annotations (i.e., P values provided in Table 3).

Confirmation of genes regulated in response to cadmium exposure

Quantitative real-time (RT) PCR was used to validate expression levels of four genes identified on the microarray; three genes were regulated following cadmium exposure (i.e., cuticle protein-2, Contig 257; 2-domain haemoglobin protein subunit, Contig 262; metallothionein, Contig 221) and one gene for which expression was not altered (i.e., serine-threonine kinase, Contig 274). This validation test included a set of replicate RNA from the microarrays plus six independent biological replicates. Primers and *TaqMan* probes were designed using PRIMER EXPRESS, V (Applied Biosciences), which are listed in Table 3. Reverse transcription was performed with the Omniscript reverse transcription kit from Qiagen. Two micrograms of total RNA was reverse transcribed for each sample using random primers (final concentration of 5 µM) for metallothionein and gene-specific primers for the other three genes (0.5 µM final concentration, Table 3). Real-time PCR was performed following the Qiagen protocols on the Applied Biosystems 7700 machine and included a standard curve in each run for each gene amplified in that run. The standard curve consisted of serial dilutions of the cDNA being amplified. Applied Biosystems Master Mix was used in each amplification, which contained all PCR components necessary except the cDNA, primers (900 nM each, final concentration) and the *Taqman* probe (250 nM, final concentration). *Taqman* probes were FAM labeled and contained an MGB quencher. Controls to test for DNA contamination were always included, even though DNase digestion was performed on the RNA before reverse transcription. The amplification steps consisted of the standard 40 cycles

preceded by two minutes at 50°C and ten minutes at 95°C to activate the enzyme. Each cycle included 15 seconds at 95°C and 1 minute at 60°C. For each sample, the cycle at which amplification reached the exponential phase was recorded as the Ct value. Final level of transcript for metallothionein was normalized using the Nanodrop spectrophotometer to directly measure the cDNA level in each sample. The other genes were normalized to the levels of serine threonine kinase, which was not differentially regulated.

Metallothionein characterization

cDNA and genomic sequence determination

Quality EST sequences were obtained from 1,529 cDNA elements on the array. These were clustered, their open reading frames were determined and they were annotated based on sequence similarity searches against NCBI and custom protein databases. Details on our methods and of the results are presented elsewhere [67]. The sequences of two elements on the microarray that respond to cadmium stress revealed identical transcripts for a cysteine-rich protein resembling metallothionein. The translated gene sequence was used to identify additional metallothionein-like loci within a sequence database derived from 36 cDNA libraries used to support the ongoing *D. pulex* genome sequencing project (Colbourne et al., in prep). The program tBlastn [95] was used at a significance cut-off value of $e < 1 \times 10^{-10}$ to search 36,342 sequence assemblies (EST clusters). To determine whether the identified sequences were homologous or alternatively spliced loci, these were subsequently matched to the latest *D. pulex* genome sequence assembly by the Joint Genome Institute using the Blastn tool on wFleaBase [96]. The cDNA was aligned to genomic DNA sequences by the Clustal method using MegAlign (DNASTAR, Inc.).

Phylogenetic analysis

Metallothionein protein sequences for 18 selected insect and crustacean species were obtained from the NCBI database. The multiple cDNA alignment was produced by the MAFFT version 5 program using the E-INS-i strategy [97]. The scoring matrix for the protein sequences was Blosum62 and the alignment parameters included a gap opening penalty of 3 and a gap extension penalty (offset value) of 0.15. This protein alignment was then used to calculate a genetic similarity matrix including all loci by using MEGA version 3.1 [98] with a Poisson correction of the distances and pairwise deletion of the alignment gaps. A corresponding neighbor-joining tree was constructed that included 1,000 bootstrap pseudo-replicates of the data for assigning confidence to nodes. A maximum likelihood phylogeny was also constructed by quartet puzzling using the program Tree-Puzzle version 5.2 [99] with 10,000 puzzling steps and with the Dayhoff model of

amino acid substitution under uniform rates of molecular evolution.

Authors' contributions

JRS, JKC, JCD, CYC, CLF and JWH conceived the study and designed the experiments and subsequent analyses. JRS and SPG performed cadmium exposures and toxicity assays, JRS and JCD conducted the microarray experiments and THH performed the statistical analyses. JRS drafted the manuscript and all authors contributed to, improved upon, and read and approved the final version.

Acknowledgements

The authors thank Brandon Mayes, Noah Greenberg, and Heather Hudenko (Dartmouth College) for valuable help maintaining daphniid cultures, and Darren Bauer, Kelley Thomas, Jim Haney (University of New Hampshire) and Elizabeth Bohuski (CGB, Indiana University) for their contributions towards the construction of cDNA libraries. The ultra-quality microarrays are thanks to Elizabeth Bohuski, Justen Andrews and staff at The Center for Genomics and Bioinformatics (CGB). Brian Eads contributed valuable sequence information and analyses. Computer support was provided in part by Phillip Steinbachs (CGB) and Dick Repasky (Indiana University Information Technology Services). We referenced an early draft genome sequence for *Daphnia pulex*, which is the work of DOE Joint Genome Institute under the auspices of the U.S. Department of Energy's Office of Science, Biological and Environmental Research Program, and by the University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48, Lawrence Berkeley National Laboratory, under Contract No. DE-AC02-05CH11231, and Los Alamos National Laboratory, under Contract No. W-7405-ENG-36 and in collaboration with the Daphnia Genomics Consortium (DGC) [17]. This manuscript benefited from comments provided by two anonymous reviewers. This work was supported by grants from NSF (BE/GEN-EN DEB-0221837, JWH, CLF, CYC, JRS and JKC) and NIH-NIEHS (P42 ES07373, Dartmouth Superfund Basic Research Program on Toxic Metals, JWH; Project 2, JWH, and Project 7, CLF and CYC) and also supported in part by the METACyt Initiative of Indiana University, funded in part through a major grant from the Lilly Endowment, Inc. This research benefits from, and contributes to the Daphnia Genomic Consortium.

References

1. White KP: **Functional genomics and the study of development, variation and evolution.** *Nat Rev Genet* 2001, **2(7)**:528-537.
2. Stoughton RB, Friend SH: **How molecular profiling could revolutionize drug discovery.** *Nat Rev Drug Discov* 2005, **4(4)**:345-350.
3. Stranger BE, Dermitzakis ET: **The genetics of regulatory variation in the human genome.** *Hum Genomics* 2005, **2(2)**:126-131.
4. Lubbock J: **An account of the two methods of reproduction in *Daphnia* and of the structure of the ephippium.** *Phil Trans Roy Soc London* 1857, **147**:79-100.
5. Richard J: **Revision des Cladoceres.** *Ann Sci Nat Zool Paleon* 1896, **18**:279-389.
6. de Bernardi RPR: **Why *Daphnia*.** In "*Daphnia*" *Memorie dell'Istituto Italiano di Idrobiologia Dr Marco De Marchi Volume 45.* Edited by: de Bernardi RPaR. Verbania Pallanza: Consiglio Nazionale Delle Ricerche Istituto Italiano Di Idrobiologia; 1987.
7. Gunatilaka ADP, Puzicha H: **The evaluation of "Dynamic Daphnia Test" after a decade of use: Benefits and constraints.** In *Biomonitoring and Biomarkers as Indicators of Environmental Change 2* Edited by: Butterworth FM GA, Gonshebbatt ME. New York: Kluwer Academic/Plenum Publishers; 2000.
8. Lilius H, Hastbacka T, Isomaa B: **A Comparison of the Toxicity of 30 Reference Chemicals to *Daphnia-Magna* and *Daphnia-Pulex*.** *Environ Toxicol Chem* 1995, **14(12)**:2085-2088.

9. EPA US: **Quality Criteria for Water**. In Agency USEP Washington, DC: Office of Waste Regulation and Standards, Criteria and Standards Division; 1986.
10. EPA US: **Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms, 5th addition**. Washington, DC: U.S. EPA, Office of Water; 2002:275.
11. EPA US: **Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms, 4th addition**. Washington, DC: U.S. EPA, Office of Water; 2002:335.
12. Gracey AY, Troll JV, Somero GN: **Hypoxia-induced gene expression profiling in the euryoxic fish *Gillichthys mirabilis***. *Proc Natl Acad Sci USA* 2001, **98(4)**:1993-1998.
13. Gracey AY, Cossins AR: **Application of microarray technology in environmental and comparative physiology**. *Ann Rev Phys* 2003, **65**:231-259.
14. Lettieri T: **Recent applications of DNA microarray technology to toxicology and ecotoxicology**. *Environ Health Perspect* 2006, **114(1)**:4-9.
15. Hebert PDN: **The Daphnia of North America: an illustrated fauna [CD-ROM distributed by author]**. Guelph, ON: Cyber-Natural Software-Dept. of Zool, Univ. of Guelph; 1995.
16. Shaw JR, Dempsey TD, Chen CY, Hamilton JW, Folt CL: **Comparative toxicity of cadmium, zinc, and mixtures of cadmium and zinc to daphnids**. *Environ Toxicol Chem* 2006, **25(1)**:182-189. [<http://daphnia.cgb.indiana.edu>].
18. Soetaert A, Moens LN, Van der Ven K, Van Leemput K, Naudts B, Blust R, De Coen WM: **Molecular impact of propiconazole on *Daphnia magna* using a reproduction-related cDNA array**. *Comp Biochem Physiol C Toxicol Pharmacol* 2006, **142(1-2)**:66-76.
19. Poynton HC, Varshavsky JR, Chang B, Cavigliolo G, Chan S, Holman PS, Loguinov AV, Bauer DJ, Komachi K, Theil EC, Perkins EJ, Hughes O, Vulpe CD: ***Daphnia magna* ecotoxicogenomics provides mechanistic insights into metal toxicity**. *Environ Sci Technol* 2007, **41(3)**:1044-1050.
20. Soetaert A, van der Ven K, Moens LN, Vandenbrouck T, van Remortel P, De Coen WM: ***Daphnia magna* and ecotoxicogenomics: gene expression profiles of the anti-ecdysteroidal fungicide fenarimol using energy-, molting- and life stage-related cDNA libraries**. *Chemosphere* 2007, **67(1)**:60-71.
21. Eisler R: **Cadmium hazards to fish, wildlife, and invertebrates: a synoptic review**. In *Contaminant Hazard Reviews* Edited by: Eisler R. Laurel, MD: U.S. Fish and Wildlife Service; 1988:65. [<http://www.atsdr.cdc.gov/cercla/05list.html>].
23. [<http://www.epa.gov/superfund/>].
24. Guan R, Wang WX: **Multiphase biokinetic modeling of cadmium accumulation in *Daphnia magna* from dietary and aqueous sources**. *Environ Toxicol Chem* 2006, **25(11)**:2840-2846.
25. Verboost PM, Flik G, Lock RAC, Bonga SEW: **Cadmium Inhibition of Ca-2+ Uptake in Rainbow-Trout Gills**. *Am J Phys* 1987, **253(2)**:R216-R221.
26. Verboost PM, Flik G, Lock RAC, Bonga SEW: **Cadmium Inhibits Plasma-Membrane Calcium-Transport**. *J Mem Biol* 1988, **102(2)**:97-104.
27. Verboost PM, Flik G, Pang PKT, Lock RAC, Bonga SEW: **Cadmium Inhibition of the Erythrocyte Ca-2+ Pump – a Molecular Interpretation**. *J Biol Chem* 1989, **264(10)**:5613-5615.
28. Stohs SJ, Bagchi D: **Oxidative Mechanisms in the Toxicity of Metal-Ions**. *Free Radical Bio Med* 1995, **18(2)**:321-336.
29. Pearson RG: **Hard and Soft Acids and Bases**. *Chem Brit* 1967, **3(3)**:103-8.
30. Schwarzenbach G: **The General, Selective, and Specific Formation of Complexes by Metallic Cations**. *Adv In Chem* 1961, **3**:257-285.
31. Paquin PR, Gorsuch JW, Apte S, Batley GE, Bowles KC, Campbell PG, Delos CG, Di Toro DM, Dwyer RL, Galvez F, Gensemer RW, Goss GG, Hogstrand C, Janssen CR, Naddy RB, Playle RC, Santore RC, Schneider U, Stubblefield WA, Wood CM, Wu KB: **The biotic ligand model: a historical overview**. *Comp Biochem Physiol C Toxicol Pharmacol* 2002, **133(1-2)**:3-35.
32. Knapen D, Redeker ES, Inacio I, De Coen W, Verheyen E, Blust R: **New metallothionein mRNAs in *Gobio gobio* reveal at least three gene duplication events in cyprinid metallothionein evolution**. *Comp Biochem Physiol C Toxicol Pharmacol* 2005, **140(3-4)**:347-355.
33. Klaassen CD, Liu J, Choudhuri S: **Metallothionein: an intracellular protein to protect against cadmium toxicity**. *Annu Rev Pharmacol Toxicol* 1999, **39**:267-294.
34. Vallee BL: **Introduction to metallothionein**. *Methods Enzymol* 1991, **205**:3-7.
35. Palmiter RD: **The elusive function of metallothioneins**. *Proc Natl Acad Sci USA* 1998, **95(15)**:8428-8430.
36. Sato M, Bremner I: **Oxygen Free-Radicals and Metallothionein**. *Free Radical Bio Med* 1993, **14(3)**:325-337.
37. Dallinger R: **Invertebrate organisms as biological indicators of heavy metal pollution**. *Appl Biochem Biotechnol* 1994, **48(1)**:27-31.
38. Goering PL, Klaassen CD: **Resistance to cadmium-induced hepatotoxicity in immature rats**. *Toxicol Appl Pharmacol* 1984, **74(3)**:321-329.
39. Olsson PE, Kille P: **Functional comparison of the metal-regulated transcriptional control regions of metallothionein genes from cadmium-sensitive and tolerant fish species**. *Biochim Biophys Acta* 1997, **1350(3)**:325-334.
40. Amiard JC, Amiard-Triquet C, Barka S, Pellerin J, Rainbow PS: **Metallothioneins in aquatic invertebrates: their role in metal detoxification and their use as biomarkers**. *Aquat Toxicol* 2006, **76(2)**:160-202.
41. Valls M, Bofill R, Gonzalez-Duarte R, Gonzalez-Duarte P, Capdevila M, Atrian S: **A new insight into metallothionein (MT) classification and evolution. The in vivo and in vitro metal binding features of *Homarus americanus* recombinant MT**. *J Biol Chem* 2001, **276(35)**:32835-32843.
42. Subkhankulova T, Livesey FJ: **Comparative evaluation of linear and exponential amplification techniques for expression profiling at the single-cell level**. *Genome Biology* 2006, **7(3)**:.
43. Schaffer AA, Wolf YI, Ponting CP, Koonin EV, Aravind L, Altschul SF: **IMPALA: matching a protein sequence against a collection of PSI-BLAST-constructed position-specific score matrices**. *Bioinformatics* 1999, **15(12)**:1000-1011.
44. Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M: **Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research**. *Bioinformatics* 2005, **21(18)**:3674-3676.
45. Hentze MW, Muckenthaler MU, Andrews NC: **Balancing acts: Molecular control of mammalian iron metabolism**. *Cell* 2004, **117(3)**:285-297.
46. Rouault TA: **The role of iron regulatory proteins in mammalian iron homeostasis and disease**. *Nat Chem Biol* 2006, **2(8)**:406-414.
47. Fowler BA HC, Kojima Y, Webb M: **Nomenclature of metallothionein**. *Experientia Suppl* 1987, **52**:19-22.
48. Pedersen SN, Pedersen KL, Hojrup P, Depledge MH, Knudsen J: **Primary structures of decapod crustacean metallothioneins with special emphasis on freshwater and semi-terrestrial species**. *Biochem J* 1996, **319**:999-1003. [<http://dgc.cgb.indiana.edu>].
50. Searle PF: **Metallothionein gene regulation**. *Biochem Soc Trans* 1987, **15(4)**:584-586.
51. Culotta VC, Hamer DH: **Fine Mapping of a Mouse Metallothionein Gene Metal Response Element**. *Mol Cell Biol* 1989, **9(3)**:1376-1380.
52. Andrew AS, Warren AJ, Barchowsky A, Temple KA, Klei L, Soucy NV, O'Hara KA, Hamilton JW: **Genomic and proteomic profiling of responses to toxic metals in human lung cells**. *Environ Health Perspect* 2003, **111(6)**:825-835.
53. Tan YX, Shi LM, Hussain SM, Xu J, Tong WD, Frazier JM, Wang C: **Integrating time-course microarray gene expression profiles with cytotoxicity for identification of biomarkers in primary rat hepatocytes exposed to cadmium**. *Bioinformatics* 2006, **22(1)**:77-87.
54. EPA US: **2001 Update of Ambient Water Quality Criteria for Cadmium**. Agency USEP. Washington, DC: Office of Waste Regulation and Standards, Criteria and Standards Division; 2001.
55. Suedel BC, Rodgers JH, Deaver E: **Experimental factors that may affect toxicity of cadmium to freshwater organisms**. *Arch Environ Con Tox* 1997, **33(2)**:188-193.
56. De Coen WM, Janssen CR: **A multivariate biomarker-based model predicting population-level responses of *Daphnia magna***. *Environ Toxicol Chem* 2003, **22(9)**:2195-2201.

57. Tessier AJ, Goulden CE: **Estimating Food Limitation in Cladoceran Populations.** *Limnol Oceanogr* 1982, **27(4)**:707-717.
58. Abigail P, Moreno R, Medesani DA, Rodriguez EM: **Inhibition of molting by cadmium in the crab *Chasmagnathus granulata* (Decapoda Brachyura).** *Aquat Toxicol* 2003, **64(2)**:155-164.
59. Hook SE, Fisher NS: **Sublethal effects of silver in zooplankton: Importance of exposure pathways and implications for toxicity testing.** *Environ Toxicol Chem* 2001, **20(3)**:568-574.
60. Cervera A, Maymo AC, Martinez-Pardo R, Garcera MD: **Vitellogenesis inhibition in *Oncopeltus fasciatus* females (Heteroptera: Lygaeidae) exposed to cadmium.** *J Insect Physiol* 2005, **51(8)**:895-911.
61. Hessen DO, Rukke NA: **The costs of moulting in *Daphnia*; mineral regulation of carbon budgets.** *Freshwater Biol* 2000, **45(2)**:169-178.
62. Wang WX, Fisher NS: **Effects of calcium and metabolic inhibitors on trace element uptake in two marine bivalves.** *J Exp Mar Biol Ecol* 1999, **236(1)**:149-164.
63. Bondgaard M, Bjerregaard P: **Association between cadmium and calcium uptake and distribution during the moult cycle of female shore crabs, *Carcinus maenas*: an in vivo study.** *Aquat Toxicol* 2005, **72(1-2)**:17-28.
64. Norum U, Bondgaard M, Pedersen TV, Bjerregaard P: **In vivo and in vitro cadmium accumulation during the moult cycle of the male shore crab *Carcinus maenas* - interaction with calcium metabolism.** *Aquat Toxicol* 2005, **72(1-2)**:29-44.
65. Koizumi S, Yamada H: **DNA microarray analysis of altered gene expression in cadmium-exposed human cells.** *J Occup Health* 2003, **45(6)**:331-334.
66. Zhou T, Jia XD, Chapin RE, Maronpot RR, Harris MW, Liu J, Waalkes MP, Eddy EM: **Cadmium at a non-toxic dose alters gene expression in mouse testes.** *Toxicol Lett* 2004, **154(3)**:191-200.
67. Colbourne JK, Eads BD, Shaw J, Bohuski E, Bauer DJ, Andrews J: **Sampling *Daphnia*'s expressed genes: preservation, expansion and invention of crustacean genes with reference to insect genomes.** *BMC Genomics* 2007, **8**:217.
68. Schlenk D, Brouwer M: **Induction of Metallothionein Messenger-Rna in the Blue-Crab (*Callinectes Sapidus*) after Treatment with Cadmium.** *Comp Biochem Phys C* 1993, **104(2)**:317-321.
69. Berglind R: **The Effects of Cadmium on Ala-D Activity, Growth and Hemoglobin Content in the Water Flea, *Daphnia-Magna*.** *Comp Biochem Phys C* 1985, **80(2)**:407-410.
70. Zeis B, Becher B, Lamkemeyer T, Rolf S, Pirow R, Paul RJ: **The process of hypoxic induction of *Daphnia magna* hemoglobin: subunit composition and functional properties.** *Comp Biochem Phys B* 2003, **134(2)**:243-252.
71. Gorr TA, Cahn JD, Yamagata H, Bunn HF: **Hypoxia-induced synthesis of hemoglobin in the crustacean *Daphnia magna* is hypoxia-inducible factor-dependent.** *J Biol Chem* 2004, **279(34)**:36038-36047.
72. Metwally A, Safronova VI, Belimov AA, Dietz KJ: **Genotypic variation of the response to cadmium toxicity in *Pisum sativum*.** *J Exp Bot* 2005, **56(409)**:167-178.
73. Sakamoto K, Hisatomi O, Tokunaga F, Eguchi E: **Two opsins from the compound eye of the crab *Hemigrapsus sanguineus*.** *J Exp Biol* 1996, **199(2)**:441-450.
74. Michels E, Semsari S, Bin C, De Meester L: **Effect of sublethal doses of cadmium on the phototactic behavior of *Daphnia magna*.** *Ecotox Environ Safe* 2000, **47(3)**:261-265.
75. Yuan L, Michels E, De Meester L: **Changes in phototactic behavior of *Daphnia magna* clone C(1)242 in response to copper, cadmium and pentachlorophenol.** *J Environ Sci-China* 2003, **15(6)**:841-847.
76. Clark AG: **The Comparative Enzymology of the Glutathione S-Transferases from Non-Vertebrate Organisms.** *Comp Biochem Phys B* 1989, **92(3)**:419-446.
77. Barata C, Varo I, Navarro JC, Arun S, Porte C: **Antioxidant enzyme activities and lipid peroxidation in the freshwater cladoceran *Daphnia magna* exposed to redox cycling compounds.** *Comp Biochem Physiol C Toxicol Pharmacol* 2005, **140(2)**:175-186.
78. Palmiter RD: **Regulation of Metallothionein Genes by Heavy-Metals Appears to Be Mediated by a Zinc-Sensitive Inhibitor That Interacts with a Constitutively Active Transcription Factor, Mtf-1.** *P Natl Acad Sci USA* 1994, **91(4)**:1219-1223.
79. Harlow P, Watkins E, Thornton RD, Nemer M: **Structure of an Ectodermally Expressed Sea-Urchin Metallothionein Gene and Characterization of Its Metal-Responsive Region.** *Mol Cell Biol* 1989, **9(12)**:5445-5455.
80. Cserjesi P, Fairley P, Brandhorst BP: **Functional-Analysis of the Promoter of a Sea-Urchin Metallothionein Gene.** *Biochem Cell Biol* 1992, **70(10-11)**:1142-1150.
81. Li B, Savva D: **Cloning of metallothionein cDNAs and its gene in shore crab (*Carcinus maenas*).** *Acta Biochimica Et Biophysica Sinica* 2000, **32(6)**:640-644.
82. Binz PAKJ: **Metallothionein: Molecular evolution and classification.** In *Metallothionein IV* Edited by: Klaassen CD. Basel: Birkhauser Verlag; 1999:7-13.
83. Colbourne JK, Crease TJ, Weider LJ, Hebert PDN, Dufresne F, Hobaek A: **Phylogenetics and evolution of a circumpolar species complex (*Cladocera*: *Daphnia pulex*).** *Biol J Linn Soc* 1998, **65(3)**:347-365.
84. [<http://daphnia.cgb.indiana.edu/projects/genome/>].
85. Lewis PA, Horning VWB: **Differences In Acute Toxicity Test-Results Of 3 Reference Toxicants On *Daphnia* At 2 Temperatures.** *Environ Toxicol Chem* 1991, **10(10)**:1351-1357.
86. Kilham SS, Kreeger DA, Lynn SG, Goulden CE, Herrera L: **COMBO: a defined freshwater culture medium for algae and zooplankton.** *Hydrobiologia* 1998, **377**:147-159.
87. Folt CL, Chen CY, Moore MV, Burnaford J: **Synergism and antagonism among multiple stressors.** *Limnol Oceanogr* 1999, **44(3)**:864-877.
88. Chen CY, Folt CL: **Consequences of fall warming for zooplankton overwintering success.** *Limnol Oceanogr* 1996, **41(5)**:1077-1086.
89. Stemberger RS, Miller EK: **Cladoceran body length and Secchi disk transparency in northeastern US lakes.** *Can J Fish Aquat Sci* 2003, **60(12)**:1477-1486.
90. Jin W, Riley RM, Wolfinger RD, White KP, Passador-Gurgel G, Gibson G: **The contributions of sex, genotype and age to transcriptional variance in *Drosophila melanogaster*.** *Nat Genet* 2001, **29(4)**:389-395.
91. Hedge PQR, Abernathie K, Gay C, Dharap S, Gaspard R, Hughes JE, Snesrud E, Lee N, Quackenbush J: **A concise guide to cDNA microarray analysis.** *BioTechniques* 2000, **29(3)**:548-562.
92. Smyth GK, Speed T: **Normalization of cDNA microarray data.** *Methods* 2003, **31(4)**:265-273.
93. Wettenhall JM, Smyth GK: **limmaGUI: A graphical user interface for linear modeling of microarray data.** *Bioinformatics* 2004, **20(18)**:3705-3706.
94. Smyth GK, Michaud J, Scott HS: **Use of within-array replicate spots for assessing differential expression in microarray experiments.** *Bioinformatics* 2005, **21(9)**:2067-2075.
95. Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, **25(17)**:3389-3402.
96. Colbourne JK, Singan VR, Gilbert DG: **wFleaBase: The *Daphnia* genome database.** *BMC Bioinformatics* 2005, **6**.
97. Katoh K, Kuma K, Toh H, Miyata T: **MAFFT version 5: improvement in accuracy of multiple sequence alignment.** *Nucleic Acids Res* 2005, **33(2)**:511-518.
98. Kumar S, Tamura K, Nei M: **MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment.** *Brief Bioinform* 2004, **5(2)**:150-163.
99. Schmidt HA, Strimmer K, Vingron M, von Haeseler A: **TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing.** *Bioinformatics* 2002, **18(3)**:502-504.