Environ. Sci. Technol. **2008,** *42,* 6257–6263

Gene Expression Profiling in *Daphnia magna*, Part II: Validation of a Copper Specific Gene Expression Signature with Effluent from Two Copper Mines in California

HELEN C. POYNTON,[‡,†] RICK ZUZOW,[†] ALEXANDRE V. LOGUINOV,[†] EDWARD J. PERKINS,[§] AND CHRIS D. VULPE*[,†]

Nutritional Sciences and Toxicology, University of California, Berkeley, California 94720, and U.S. Army Engineer Research and Development Center, Vicksburg, Mississippi 39180

Received January 25, 2008. Revised manuscript received April 18, 2008. Accepted April 23, 2008.

Genomic technologies show great potential for classifying disease states and toxicological impacts from exposure to chemicals into functional categories. In environmental monitoring, the ability to classify field samples and predict the pollutants present in these samples could contribute to monitoring efforts and the diagnosis of contaminated sites. Using gene expression analysis, we challenged our custom *Daphnia magna* cDNA microarray to determine the presence of a specific metal toxicant in blinded field samples collected from two copper mines in California. We compared the gene expression profiles from our field samples to previously established expression profiles for Cu, Cd, and Zn. The expression profiles from the Cu-containing field samples clustered with the laboratory-exposed Cu-specific gene expression profiles and included genes previously identified as copper biomarkers, verifying that gene expression analysis can predict environmental exposure to a specific pollutant. In addition, our study revealed that upstream field samples containing undetectable levels of Cu caused the differential expression of only a few genes, lending support for the concept of a no observed transcriptional effect level (NOTEL). If confirmed by further studies, the NOTEL may play an important role in discriminating polluted and nonpolluted sites in future monitoring efforts.

Introduction

Since the inception of toxicogenomics, ecotoxicologists have proposed ways to integrate genomics into environmental monitoring, chemical screening, and risk assessment (*1*). These applications depend on the ability to classify the biological responses caused by chemicals into groups based on the gene expression, protein expression, or metabolomic signature. In medicine, initial studies proved that classification of cancer types was possible based on their gene expression profile (*2*), leading to the development of microarray based tests that are now used clinically to predict cancer prognosis (*3*). Toxicological studies have also shown the ability to classify chemicals and predict their mode of toxicity on the basis of their gene expression profiles (*4*).

One of the first obstacles facing the integration of genomics into ecotoxicology was establishing that ecologically relevant organisms respond to different chemicals by producing chemical specific expression profiles. Recent studies demonstrate that model ecotoxicology organisms exposed to pollutants in the laboratory produce a distinctive expression pattern for each chemical related to a similar mode of action. For example, gene expression patterns are discernible in *Chironomus tentans* exposed to DDT, phenanthrene, fluoranthene, Cd, Cu, and Zn (*5*); in *Daphnia magna* exposed to Cu, Cd, and Zn and organic pollutants (*6, 7*) and in rainbow trout exposed to a variety of toxicants (*8*). In another study examining the potential applications of genomics in field monitoring, clean and contaminated field sites were successfully discriminated through gene expression analysis in European flounder (*9*).

Although the previous examples illustrate the feasibility of applying genomics to ecotoxicology, the bridge connecting a genomic response to a specific environmental exposure has not yet been built. This study is the continuation of our work to establish and validate gene expression profiles for Cu, Cd, and Zn in *Daphnia magna*. In a companion paper, we reported gene expression profiles for these metals over a range of nontoxic and toxic concentrations and showed that a concentration exists corresponding to a no observed transcriptional effect level (NOTEL) (*10*). We next challenged our *D. magna* microarray to discriminate between polluted and nonpolluted field samples using the theory of the NOTEL and predict the contaminant present in the polluted samples. Using clustering analysis and class prediction algorithms we successfully predicted Cu as the primary pollutant responsible for toxicity in blinded field samples and found very few genes differentially expressed in the nonpolluted sites.

Materials and Methods

Site Selection and Characterization. To determine the potential of gene expression analysis as a biomarker of metal pollution, we chose two locations in California where Cu was suspected to be the primary pollutant in the surface waters. Our first location, Walker Mine, located on the eastern side of the Sierra-Nevada Mountains, is an abandoned copper mine where Cu contamination has been a long standing environmental problem for two streams in proximity to the mine, Dolly Creek, and Little Grizzly Creek (see map in the Supporting Information, Figure S1A). Dolly Creek receives mine drainage directly from the mine portal and flows through the mine tailings where a legacy of ore processing has left behind Cu contamination. Dolly Creek flows into Little Grizzly Creek, bringing its Cu load with it. The second location, Greenhorn Mine, located in Northern California west of Lake Shasta, is an abandoned mine with no remediation history (see map in the Supporting Information, Figure S1B). Acid mine drainage seeps from the mine at several locations and enters Willow Creek below. The acidic pH is diluted and neutralized by Willow Creek (see the Supporting Information, Table S2); however, the Cu in the mine drainage remains mobilized, contaminating the creek.

Field Sample Collection. Field samples were collected on August 24, 2006, from Walker Mine and on August 25,

* Corresponding author phone: (510) 642-1834; fax: (510) 642-0535; address: 119 Morgan Hall UC Berkeley Berkeley, CA 94720; e-mail: vulpe@berkeley.edu.

† University of California.

‡ Present address: Molecular Indicators Research Branch, U.S. EPA, 26 W. Martin Luther King Dr., Cincinnati, OH 45213.

§ U.S. Army Engineer Research and Development Center.

10.1021/es800262k CCC: $40.75
Published on Web 07/10/2008

2006, from Greenhorn Mine. At each location, four sites were chosen for sampling on the basis of their accessibility and the predicted copper concentrations. One site was chosen upstream from the mine portal. Figure S1 in the Supporting Information shows the sampling sites for Walker Mine (A) and Greenhorn Mine (B) and Tables S1 and S2 in the Supporting Information give the GPS coordinates for each sampling site, taken using a Garmin eTrex Vista GPS unit. At each site, two four liter grab samples of water were collected in 4 L fluorinated HDPE (high-density polyethylene) bottles. At the time of sampling, pH, conductivity, and temperature were measured in the stream using a Hanna HI99130 multiparameter meter. Dissolved oxygen (DO) was measured using a Fisher Scientific traceable digital oxygen meter. Results of the water quality parameters are given in Tables S1 and S2 in the Supporting Information. All samples were stored on ice until returning to UC Berkeley, where they were transferred to 4 °C until use.

Determination of Dissolved Metal Concentrations. After collecting grab samples, three aliquots were removed from the 4 L grab samples and filtered through a 0.45 μm filter on site for dissolved metal analysis. Trace metal grade HNO_3 (Fisher Scientific, Pittsburgh, PA) was added to a final concentration of 1.5%. Metal concentrations were determined by inductively coupled plasma mass spectrometry (ICP-MS) using EPA method 200.8 (*11*) at California Laboratory Services, Rancho Cordova, CA. The dissolved metal concentrations averaged for the three replicates are given in Table S3 in the Supporting Information with standard deviation.

Determination of Dissolved Organic Carbon (DOC) Concentrations. One hundred milliliter samples were removed from the field sample bottles and filtered through a 0.45 μm nitrocellulose filter (Millipore, Billerica, MA) to remove particulate carbon. Three 25 mL aliquots were collected from the filtrate for DOC analysis. DOC analysis was performed on a 1010 total organic carbon analyzer (O.I. Analytical, College Station, TX). Potassium hydrogen phthalate (Acros Organics, Geel, Belgium) was used as a standard, producing a standard curve with an R^2 value of 0.9991. Each sample was analyzed twice with three replicates from each site. DOC concentrations were averaged for the three replicates and reported in Tables S1 and S2 in the Supporting Information. DOC analysis of filtered Milli-Q water confirmed no DOC contamination was introduced during filtering.

Acute Toxicity Assays. *Daphnia magna* culture maintenance was described previously (*6*) and is also available in the Supporting Information. Acute toxicity assays were conducted according to the U.S. EPA whole effluent toxicity (WET) protocol (*12*). First, the toxicity of the undiluted field samples was assessed. First instar *D. magna* were placed in 25 mL of either culture media or undiluted field sample. Following a 24 h exposure, the number of survivors was determined as those remaining mobile. For the two samples causing acute lethality, GM_DE and GM_DS1, three dilutions (50, 25, and 12.5%), an undiluted sample (100%) and a media control were used to determine the toxicity of the field samples. The dilutions were made with modified COMBO media. Following a 24 h exposure, the number of survivors was counted to determine the dilution required to remove acute lethality.

Chronic Toxicity Assay. Static renewal chronic toxicity assays were conducted using protocols similar to the U.S. EPA chronic toxicity WET protocol (*13*). Ten first instar *D. magna* were each placed in 25 mL of media containing concentrations of copper sulfate (Fisher Scientific, Hampton, NH) equal to the 1/10 LC_{50} (6 μg/L), the $NOEC_{acute}$ (no observable effect concentration) (30 μg/L), or a zero concentration control (toxicity end points were reported previously (*10*)). The chronic toxicity of the Walker Mine field samples was measured using 25 mL of undiluted field sample. Over the next 21 days, shed exoskeletons (as evidence of molting) and the number of offspring were counted and removed each day. At the conclusion of the test (22 days) the length of the daphnids was measured from the top of the head capsule to the bottom of body carapace and the sex was determined using 40X magnification (*14*). The following end points were calculated: total number of offspring/female daphnid, total number of molting cycles/daphnid, total number of broods, average number of offspring/brood (first 2 broods only), and length. The results are shown in Figure S2 in the Supporting Information. Student's *t* test was used to determine statistical difference from control or upstream sample.

Field Sample Exposures. Exposures were performed using ~20 adult (16−18 day old) *D. magna* randomly distributed into either 1 L of modified COMBO media (used as the controls in the microarray hybridization experiments) or 1 L of the appropriate dilution of field sample for 24 h. For all the Walker Mine samples and the Greenhorn Mine samples GM_UP and GM_DS2, no dilution was necessary. For the Greenhorn Mine samples GM_DE and GM_DS1, the field samples were diluted to 50% with modified COMBO media to prevent acute toxicity. A COMBO media control was performed alongside each field sample exposure, so that each exposure had a complimentary unexposed control for the microarray hybridizations. To obtain a gene expression profile for the upstream reference field sites and compare these expression profiles to the expression profiles of contaminated sites as well as single metal exposures done previously, we used laboratory media and not upstream field water for the unexposed controls in all the microarray hybridizations. For the microarray experiments, three biological replicates were performed for each field sample exposure on separate dates. A fourth exposure was performed for q-RT-PCR analysis. Following exposure studies, pH, dissolved oxygen (DO), water hardness, and alkalinity were measured, and recorded. pH was measured using a basic pH electrode (Denver Instruments, Denver, CO). DO was measured with a traceable digital oxygen meter (Fisher scientific, Hampton, NH). Water hardness and alkalinity were determined using standard methods (*15*). In unexposed controls, pH varied between 7.9 and 8.2; DO ranged from 7.3 to 8.0 mg/L; water hardness was between 120 and 135 mg $CaCO_3$/L; and alkalinity was maintained between 65 and 80 mg $CaCO_3$/L. In the field sample exposures, the water quality measurements were similar to those recorded in Table 1 with the exception of DO. The DO in these experiments resembled the control values, ranging between 6.8 and 8.2 mg/L.

RNA Isolation and Microarray Hybridization. *D. magna* were harvested as described previously (*6*), and RNA was isolated using Trizol according to standard methods (Invitrogen, Carlsbad, CA). Before we proceeded to reverse transcription, RNA from both the unexposed and exposed *D. magna* was split into two pools, to provide two replicate hybridizations for each metal exposure. Because three exposures were performed for each metal, and RNA from each exposure was hybridized to two different microarrays, there were six hybridizations for each exposure condition. Details related to the construction of the *D. magna* microarray and the microarray hybridization procedure have been described previously (*6*) and are available in the Supporting Information. Information about experimental design, raw signal intensity values, and other MAIME compliant data are available at the Gene Expression Omnibus (GEO) (located at http://www.ncbi.nlm.nih.gov/geo) with the accession numbers GSE7666 (Walker Mine data) and GSE7667 (Greenhorn Mine data).

Identification of Candidate Differentially Expressed Genes. The statistical methods used to normalize the data

TABLE 1. Quantitative Reverse Transcription PCR (q-RT-PCR) Confirmation of Microarray Gene Expression Ratios[a]

			1/10LC50 Cu	NOEC Cu	WM_UP	WM_LG	GM_UP	GM_DS2
Monooxygenase	DV437798	μarray	1.23	2.20	0.38	3.25	0.37	2.90
		q-RT-PCR	2.32	2.45	0.10	2.91	-0.18	2.33
IMPase	DV437806	μarray	0.60	0.85	0.18	1.74	-0.28	1.48
		q-RT-PCR	0.49	-0.04	-0.37	1.76	0.07	1.08
cellulase	DV437797	μarray	-1.42	-2.47	-0.49	-3.10	0.16	-2.86
		q-RT-PCR	-1.31	-2.80	-0.65	-2.43	0.09	-2.55
slit homolog	DV437805	μarray	0.44	1.16	0.15	1.37	0.19	1.34
		q-RT-PCR	3.24	1.26	-0.22	1.34	0.46	1.86
MT (a)	DV437799	μarray	0.64	1.25	-0.55	1.67	0.12	1.59
		q-RT-PCR	2.37	1.53	-1.50	1.85	0.10	1.75
lectin	DV437813	μarray	-0.35	-0.44	-0.21	-1.23	-0.05	-1.22
		q-RT-PCR	-0.81	-3.94	-0.92	-3.57	-0.77	-4.75
unknown	ES408276	μarray	-0.28	-1.06	-0.54	-0.90	-0.02	-1.18
		q-RT-PCR	0.01	-0.92	-0.52	-4.07	-0.60	-3.38
preamylase	ES408209	μarray	-0.65	-1.18	-0.30	-1.83	-0.02	-0.80
		q-RT-PCR	0.19	-1.29	-0.20	-1.48	-0.67	-1.35

[a] *D. magna* were exposed to Cu at the 1/10 LC50 (6 μg/L) or the NOEC (30 μg/L) or a water sample collected from Walker Mine or Greenhorn mine for 24 h. Following the exposure, RNA was isolated and reverse transcribed. q-RT-PCR was carried out on the cDNA using SYBR Green as described in the methods section. Log_2 ratios from the q-RT-PCR analysis are compared to the averaged log_2 ratios for the microarray experiments. Genes significantly differentially expressed for a given condition are shown in red for upregulated genes or green for downregulated genes.

and identify differentially expressed genes are described in Loguinov et al. (*16*). We applied an average false positive cutoff of 1 to identify candidates for differential gene expression. cDNAs differentially expressed in both technical replicates and in two of the three biological replicates were chosen as candidate differentially expressed cDNAs. The methods used to identify differentially expressed genes is available in the companion paper (*10*) and in the Supporting Information.

Quantitative Reverse Transcription PCR. To confirm the differential expression, we chose several genes for quantitative reverse transcription PCR (q-RT-PCR) analysis. Experimental details pertaining to q-RT-PCR is available in the companion paper (*10*) and in the Supporting Information. Primer sequences for all genes assayed can be found in Table S5 in the Supporting Information.

Clustering and Class Prediction Analysis. To determine the relationship between the gene expression profiles from the upstream or reference sites, the mine field samples, and the metal exposures we conducted in the laboratory, we performed clustering analysis. Clustering was performed using the tools available from Expression Profile data CLUSTering and analysis (EPCLUST) (http://ep.ebi.ac.uk/EP/EPCLUST/). We clustered only the differentially expressed cDNAs with the averages given in Table S4 in the Supporting Information. The data was first transposed following gene clustering using the average linkage (average distance, UPGMA) clustering based on correlation measure-based distance (uncentered). The different exposures were then clustered using the same algorithm. Other clustering algorithms available through EPCLUST produced similar clusters. Following the clustering analysis, the samples were unblinded to reveal the identity of the mine samples and are shown in Figure 2. Additional clustering was performed with three biological replicates for each condition and are shown in Figure S3 in the Supporting Information. We used K-nearest neighbor (KNN) class prediction method through Gene Pattern program to predict the class of each unknown mine sample (*17, 18*). Technical replicates were averaged to obtain an averaged log_2 ratio for each biological replicate. The cross-validation algorithm was run on the gene expression values from laboratory exposures to obtain parameters that best-predicted the known classes of the metal exposures. A model was created for the metal exposure data using the following parameters: no. of features, 10; feature selection statistic, *t* test; no. of neighbors, 10; weighting type, distance; distance measure, cosine distance. This model was used to predict the classes of the "unknown" mine samples and results are shown in Table S6 in the Supporting Information.

Results and Discussion

The purpose of our study was to illustrate the feasibility of using gene expression profiling in environmental monitoring by confronting two major challenges. Can gene expression profiling predict the contaminant present in field samples? Is it possible to differentiate between toxic samples that will likely cause effects and nontoxic samples? To address these questions, we chose two abandoned Cu mines to conduct our field research. These locations were ideally suited for this study because there were uncontaminated upstream sites close to the sampling sites, and the surface waters were primarily contaminated with only one pollutant, Cu.

Metal Concentrations in Mine Effluent. As predicted from past metal analysis done by the California Regional Water Quality Control Boards, Cu is the primary pollutant in the creeks of Walker Mine and Greenhorn Mine. Table S3 in the Supporting Information presents the results of the dissolved metal concentrations of Cu, Cd, and Zn at four sampling sites at Walker Mine and Greenhorn Mine. Upstream samples contain undetectable levels of Cu, providing evidence that the Cu contamination originated from the mines, and suggesting that the upstream sites are suitable reference sites. The downstream samples contain high levels of dissolved Cu in most cases greater than the $NOEC_{acute}$ concentration of 30 μg/L (see Poynton et al. (*10*)). At the Walker Mine location, the mine portal has been sealed during recent remediation and Cu levels coming directly out of the mine

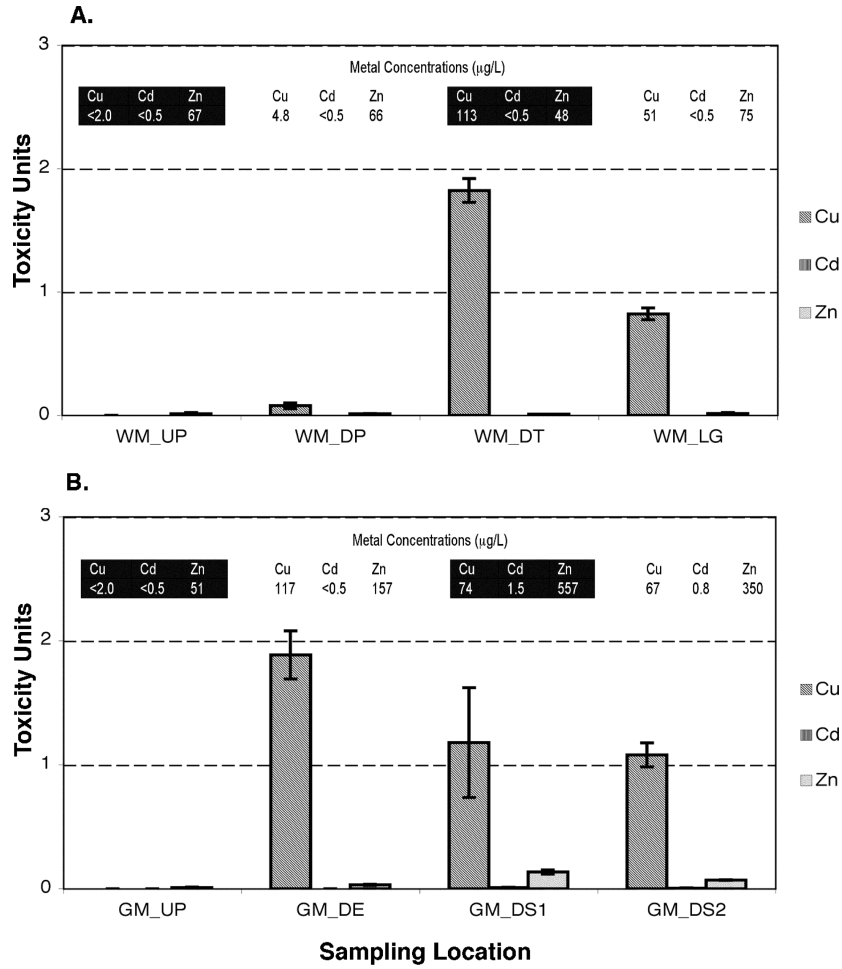

FIGURE 1. Copper, cadmium, and zinc toxicity units at Walker Mine and Greenhorn Mine. Dissolved metal concentrations were determined for each sampling site at (A) Walker Mine and (B) Greenhorn Mine using ICP-MS. The dissolved metal concentration (shown above each graph) was divided by the LC_{50} for each metal (Cu, 62 μg/L; Cd, 180 μg/L; Zn, 5000 μg/L) to determine the toxicity units (TUs) for each sample. TUs are shown with standard deviation.

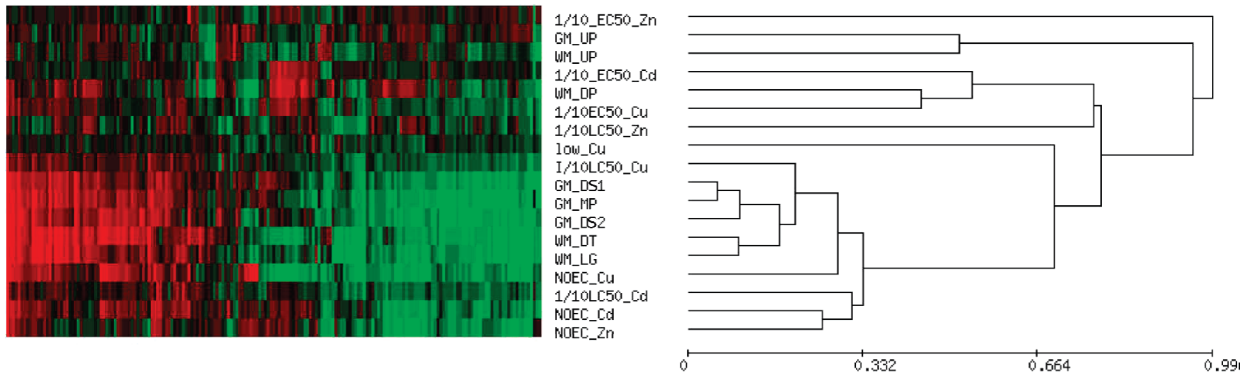

FIGURE 2. Clustering of differentially expressed genes following exposures to metals or field samples. cDNA collected from daphnids exposed to either a specific metal concentration or field sample were hybridized with cDNA from daphnids exposed to untreated laboratory media. cDNAs determined to be differentially expressed in the exposures compared to unexposed controls were used for the clustering analysis. Clustering was performed using the tools available from EPCLUST (http://ep.ebi.ac.uk/EP/EPCLUST/) including the average linkage (average distance, UPGMA) clustering based on correlation measure-based distance (uncentered).

portal are low. However, Dolly Creek flows directly through the mine tailings and accumulates the highest dissolved Cu levels. Zinc also appears to be a major contaminant especially at the Greenhorn Mine site. Because Zn is less toxic to many aquatic organisms as compared with Cu, we normalized the dissolved metal concentrations by the LC_{50} and presented their concentrations as Toxicity Units (TUs) in Figure 1. As shown in Figure 1 B, although Zn concentrations are greater than Cu at the Greenhorn mine sites, they contribute very little to the overall toxicity of these samples. Cu levels at the Greenhorn mine site surpass the LC_{50} at all three downstream locations.

Acute Toxicity of Mine Effluent. Because the concentration of Cu in the field samples was in many cases above the LC_{50}, we performed acute toxicity assays to determine appropriate dilutions to use for the *D. magna* exposures. All samples showed little or no acute toxicity except for two Greenhorn mine samples closest to the mine drainage entry

site, GM_DE and GM_DS1. Fifty percent dilution of these samples with standard laboratory media prevented all acute toxicity.

Prediction of Cu Toxicity by Gene Expression Analysis. To illustrate that expression profiling can be used to predict the presence of specific pollutants in environmental samples, we exposed daphnids to field samples collected from Walker Mine and Greenhorn Mine. Following RNA Isolation, we double-blinded samples to mask the identity of the samples during microarray hybridizations and data analysis.

After performing microarray hybridizations, we identified differentially expressed genes associated with each exposure (*16*) and calculated an average $\log_2$ ratio for each transcript. Ratios for the 189 differentially expressed genes were uploaded into EPCLUST to perform hierarchical clustering. The results shown in Figure 2 revealed that five of the mine samples clustered together and that they were most closely related to the expression profiles for Cu 1/10 LC_{50} and Cu NOEC. These samples are separated on a node with the Cu exposures and are distinct from the expression profiles of the other metals. Uncovering their identity revealed that these five field samples were the samples highly contaminated with Cu. Therefore, the gene expression analysis correctly predicted that these samples contained Cu and not toxic concentrations of Cd or Zn. We also performed clustering analysis on the individual exposures to assess the similarity of individual replicates (see Figure S3 in the Supporting Information). Although one of the Cu exposures failed to cluster with the other replicates, the only metal exposures which clustered with the contaminated mine samples were the other two 1/10 LC_{50} Cu exposures. As a final test to determine how well gene expression profiling predicted the toxicant present in the field samples, we performed the class prediction algorithm K-nearest neighbors (KNN). This specific class prediction method was chosen because it allows the user to cross-validate the training model on the training set and is also capable of multiple class predictions. Using KNN, the two contaminated Walker Mine sites (WM_DT) and (WM_LG) were predicted to belong to the 1/10 LC_{50} Cu class across all biological replicates (see Table S6 in the Supporting Information). The contaminated Greenhorn Mine sites (GM_MP, GM_DS1 and GM_DS2) were each predicted to belong to the 1/10 LC_{50} class in 2 of the 3 biological replicates. Interestingly, the other replicate in GM_MP and GM_DS1 was predicted to belong to the NOEC_Zn class, which may result from the presence of Zn in these samples. The upstream samples and WM_DP were not consistently predicted to belong to any class, illustrating that these samples are not similar to any one metal exposure. Overall, the clustering analysis and class prediction strongly suggested that Cu is the primary toxicant in the contaminated field samples. This is the first study to our knowledge that has used a genomic response to successfully predict the presence of a specific toxicant in a field sample.

This study has also shown that standard laboratory media is a suitable control for performing microarray hybridizations with field samples. We chose standard laboratory media for our unexposed controls because we were able to directly compare gene expression profiles from the field samples and single chemical exposures presented in a companion paper (*10*) and also construct gene expression profiles for the upstream samples (WM_UP and GM_UP), or reference sites. In the expression profiles for the upstream samples few genes are differentially expressed; however, Table S1 and S2 in the Supporting Information show that several general water quality measurements differ between the field samples and modified COMBO media. The water hardness and alkalinity in the field samples is slightly lower than our laboratory media. Also, the field samples contained dissolved organic carbon (DOC), which is not present in our laboratory media. Despite these differences, the expression profiles of the upstream samples caused very few gene expression changes compared with the laboratory media, making it a suitable control in our microarray studies. This confirms that gene expression responds to chemical stress and environmental challenges, but is not greatly affected by slight alterations in the environment. However, at other field sites, different water quality parameters may alter gene expression, affecting the expression profiles of the contaminated samples. Reference sites should be tested prior to using laboratory media for unexposed controls to ensure that gene expression is not affected by uncontaminated water samples at the site.

Biomarkers of Cu and Metal Exposure. After revealing the identity of the mine samples, we grouped the differentially expressed genes into functional categories. Table S4 in the Supporting Information shows a complete list of all the differentially expressed genes in the eight mine samples and also the ten metal exposures described previously in a companion paper (*10*).

Because we were also interested in identifying novel biomarkers of exposure and developing multibiomarker tests for environmental monitoring, we identified a set of genes that may be suitable biomarkers of metal exposure and specifically copper exposure. We selected candidate biomarkers that were both associated with Cu or general metal exposure from our previous microarray studies (*6, 10*) and differentially expressed in the five mine samples containing high levels of Cu. These include biomarkers predicative of general metal exposure, metallothionein (MT) (a) DV437799, monoxygenase DV437798, slit homologue DV437805, cellulase DV437797, and preamylase ES408209, and those specific for Cu exposure, inositol monophosphatase (IMPase) DV437806, lectin DV437813, and ES408276. Differential expression of these genes was confirmed by q-RT-PCR analysis and correlates well with the microarray data (see Table 1). The candidate biomarkers selected include genes possibly involved in the mode of toxicity of Cu, and are further discussed in our companion paper (*10*). Because these candidate biomarkers are novel, it is not known what other factors may influence their expression; therefore, further studies with other chemicals may be needed to validate their specificity before they can be adopted as biomarkers of copper exposure.

Correlation between Chronic Toxicity and Gene Expression Profiles. Because other factors in the field samples including DOC and water hardness influence the toxicity of Cu, we compared the toxicity of the mine samples to the toxicity of defined laboratory concentrations of Cu. Figure 1 shows that five field samples contain Cu concentrations close to or exceeding the LC_{50} for Cu; however, in 24 h acute toxicity assays, three of these samples did not result in appreciable mortality. Therefore, we could not directly compare the concentration of Cu in the laboratory exposures to the concentrations of Cu in the field samples. We performed a 21 day chronic toxicity assay with the Walker Mine samples to compare them to equitoxic levels of Cu. There was no difference in reproduction in the upstream or reference site, but we were able to see significant differences in daphnids exposed to contaminated mine samples compared to the control and upstream exposed daphnids (see Figure S2 in the Supporting Information). One important result was that the WM_DP sample, which contained 4.8 μg/L, a concentration approaching the Cu 1/10 LC_{50} of 6 μg/L, did not cause the same level of chronic toxicity as the 6 μg/L exposure. There was no difference in chronic toxicity of this sample and upstream sample or the control. Additionally, we found that the two samples with the highest Cu concentrations had similar chronic toxicity to the 6 and 30 μg/L samples, although the WM_DT resulted in higher mortality. The gene expression profiles reflect the chronic

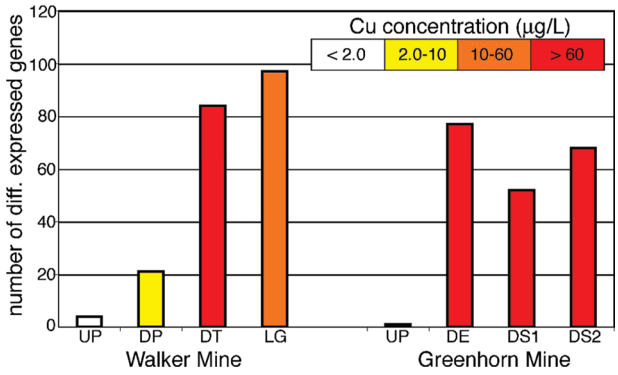

FIGURE 3. Number of genes differentially expressed after a 24 h exposure to each field sample collected from Walker Mine or Greenhorn Mine. Differentially expressed genes were determined based on the method of Loguinov et al. (*18*). The bars are colored to reflect the relative Cu levels in each of the samples.

toxicity of the field samples. As shown in Figure 2, the expression profiles from the field samples with a significant chronic response are clustered together. They are also most similar to the 1/10 LC_{50} and $NOEC_{acute}$. The WM_DP sample did not result in chronic toxicity and its gene expression profile clustered away from the samples that did. In fact, the most closely related gene expression profile was the 1/10 EC_{50} (2.4 μg/L) Cu profile, which is below chronically toxic levels and is considered a tolerated concentration. In conclusion, there appears to be a strong linkage between the gene expression profile and the level of toxicity observed in a sample.

Evidence for NOTEL in Field Samples. We examined the gene expression profiles from the upstream samples to determine how an uncontaminated environmental sample may affect gene expression in *D. magna*. Because Cu was undetectable in these samples, if gene expression changes occurred in these samples they would not be related to the presence of Cu, but to other constituents of the environmental sample. However, on the basis of the results presented in Table S4 in the Supporting Information, that few genes are differentially expressed the upstream samples, gene expression was not influenced by these other constituents. We next graphed the number of differentially expressed genes in each sample to visualize how Cu concentrations influenced the global changes in gene expression. As shown in Figure 3, few genes are differentially expressed in the samples with low Cu levels, but in samples with high Cu concentrations, the number of differentially expressed genes is much greater.

In a companion paper, we found that very low, nontoxic concentrations of Cu, Cd, and Zn caused the differential expression of only a few genes suggesting that a no observed transcriptional effect level (NOTEL) may exist for these metals (*10*). In addition, the expression profiles of the low concentration exposures were distinct from the profiles of the higher, toxic concentrations providing a means for differentiating between toxic and nontoxic metal concentrations. In this study, we found that samples with undetectable concentrations of copper caused few differentially expressed genes (see Figure 3). A similar result was reported by Roling et al., who examined gene expression changes in mummichogs (*Fundulus heteroclitus)* following remediation of Chromium contamination at a Superfund site. They showed that when tissue Cr concentrations were low in the mummichog, very few genes were differentially expressed (*19*). These findings are promising for a potential role of the NOTEL in environmental monitoring. It appears that when only a few genes are differentially expressed, there is no exposure to toxic pollutants and no observable toxicity. This suggests that a field sample resulting in no changes in gene expression, a NOTEL, could be assumed to be nontoxic.

However, the presence of differentially expressed genes does not equate to toxicity. In the WM_DP sample, 21 genes were differentially expressed, but this sample did not result in any chronic effects to reproduction or growth when compared to the upstream sample (see Figure S2 in the Supporting Information). Because Cu is an essential mineral, the organism may be responding to increased physiological Cu levels. Alternatively, the differentially expressed genes may include compensatory responses to low but tolerated levels of Cu. As suggested by Ankley et al., the identity and proposed function of the differentially expressed genes should be taken into consideration when applying the NOTEL to a regulatory setting (*1*). In our study, of the eight proposed biomarkers of Cu exposure, only one gene was differentially expressed in the WM_DP sample (See Table 1). Additionally, the gene expression profile from this sample did not cluster with the expression profiles of the toxic Cu concentrations. In a monitoring or regulatory setting, two techniques could be used to distinguish toxic and nontoxic concentrations. First, the NOTEL could be applied only to specific biomarkers correlated with toxicity such as the general metal and Cu specific biomarkers presented in this study. Second, clustering analysis of the gene expression profiles could help distinguish between toxic and nontoxic sites because nontoxic sites would not cluster with the expression profiles of the toxic sites. Proteomics and metabolic studies could also be employed to increase the weight of evidence and ensure that effects are not seen at the protein or metabolite levels. In conclusion, when no or few genes are differentially expressed in a field sample, it is very likely that this sample is not polluted. However, when gene expression changes are observed, additional criteria, such as differential expression of specific genes, will be needed to distinguish between toxic and nontoxic samples.

Using a *Daphnia magna* cDNA microarray, we illustrated the potential of gene expression profiling to predict the presence of a specific contaminant in a polluted field sample. Clustering analysis revealed that field samples containing high concentrations of Cu clustered most closely with the expression profiles of laboratory Cu exposures. This study also established the existence of a NOTEL in environmental samples suggesting a potential role for the NOTEL in environmental monitoring. Although further studies with other chemicals and mixtures are needed to validate the use of gene expression profiling in environmental monitoring, this study has shown that applying genomic tools to a field situation is feasible.

Acknowledgments

Special thanks to Steven Rosenbaum, Jeffery Huggins, and Philip Woodward of the California Regional Water Quality Control Board for advice on locations for this field study and help collecting field samples at Walker Mine and Greenhorn Mine. We also thank Dr. Garrison Sposito and Mr. Andrew Yang for their assistance and use of equipment for the TOC analysis. This research was supported by the U.S. EPA STAR fellowship (FP-91644201-0), and grants from the National Science Foundation (BES-0504603) and the Army Environmental Quality Program of the U.S. Army Corps of Engineers (BAA053799). Permission was granted by the Chief of Engineers to publish this information. This work benefits from and contributes to the *Daphnia* Genomics Consortium.

Supporting Information Available

Maps of the field sampling locations (Figure S1), chronic toxicity data for Walker Mine samples (Figure S2), and clustering analysis of individual exposures (Figure S3); water quality parameters for Walker Mine (Table S1), Greenhorn

Mine (Table S2), dissolved metal concentrations (Table S3), the list of all differentially expressed genes (Table S4), qPCR primer sequences (Table S5), and KNN class predication results (Table S6) (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

Literature Cited

(1) Ankley, G. T.; Daston, G. P.; Degitz, S. J.; Denslow, N. D.; Hoke, R. A.; Kennedy, S. W.; Miracle, A. L.; Perkins, E. J.; Snape, J.; Tillitt, D. E.; Tyler, C. R.; Versteeg, D. Toxicogenomics in regulatory ecotoxicology. *Environ. Sci. Technol.* **2006**, *40* (13), 4055–65.

(2) Ramaswamy, S.; Tamayo, P.; Rifkin, R.; Mukherjee, S.; Yeang, C. H.; Angelo, M.; Ladd, C.; Reich, M.; Latulippe, E.; Mesirov, J. P.; Poggio, T.; Gerald, W.; Loda, M.; Lander, E. S.; Golub, T. R. Multiclass cancer diagnosis using tumor gene expression signatures. *Ptoc. Natl. Acad. Sci. U.S.A.* **2001**, *98* (26), 15149–54.

(3) Pusztai, L. Chips to bedside: incorporation of microarray data into clinical practice. *Clin. Cancer Res.* **2006**, *12* (24), 7209–14.

(4) Waring, J. F.; Jolly, R. A.; Ciurlionis, R.; Lum, P. Y.; Praestgaard, J. T.; Morfitt, D. C.; Buratto, B.; Roberts, C.; Schadt, E.; Ulrich, R. G. Clustering of hepatotoxins based on mechanism of toxicity using gene expression profiles. *Toxicol. Appl. Pharmacol.* **2001**, *175* (1), 28–42.

(5) Perkins, E. J.; Furey, J.; Davis, E. The potential of screening for agents of toxicity using gene expression fingerprinting in Chironomus tentans. *Aquat. Ecosyst. Health* **2004**, *7* (3), 399–405.

(6) Poynton, H. C.; Varshavsky, J. R.; Chang, B.; Cavigiolio, G.; Chan, S.; Holman, P. S.; Loguinov, A. V.; Bauer, D. J.; Komachi, K.; Theil, J. E.; Perkins, E. J.; Hughes, O.; Vulpe, C. D. Daphnia magna ecotoxicogenomics provides mechanistic insights into metal toxicity. *Environ. Sci. Technol.* **2007**, *41* (3), 1044–50.

(7) Watanabe, H.; Takahashi, E.; Nakamura, Y.; Oda, S.; Tatarazako, N.; Iguchi, T. Development of a Daphnia magna DNA microarray for evaluating the toxicity of environmental chemicals. *Environ. Toxicol. Chem.* **2007**, *26* (4), 669–76.

(8) Koskinen, H.; Pehkonen, P.; Vehniainen, E.; Krasnov, A.; Rexroad, C.; Afanasyev, S.; Molsa, H.; Oikari, A. Response of rainbow trout transcriptome to model chemical contaminants. *Biochem. Biophys. Res. Co.* **2004**, *320* (3), 745–53.

(9) Williams, T. D.; Gensberg, K.; Minchin, S. D.; Chipman, J. K. A DNA expression array to detect toxic stress response in European flounder (Platichthys flesus). *Aquat. Toxicol.* **2003**, *65* (2), 141–157.

(10) Poynton, H. C.; Varshavsky, J. R.; Chan, S.; Loguinov, A. V.; Vulpe, C. D. Gene expression profiling in *Daphnia magna*, part I: concentration dependent profiles provide support for a no observed transcriptional effect level. *Environ. Sci. Technol.* **2008**, *42*, 6250–6256.

(11) Creed, J. T.; Brockhoff, C. A.; Martin, T. D. *Method 200.8: Determination of Trace Elements in Waters and Wastes by Inductively Coupled Plasma-Mass Spectrometry*; U.S. EPA: Cincinnati, OH, 1994.

(12) Weber, C. I. *Methods for Measuring the Acute Toxicity of Effluent and Receiving Waters to Freshwater and Marine Organisms*; Report EPA/600/4−90/027F; U.S. EPA: Cincinnati, OH, 1993.

(13) Lewis, P. A.; Klemm, D. J.; Lazorchak, J. M.; Norberg-King, T. J.; Peltier, W. H.; Heber, M. A. *Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms*; Report EPA 600/4−91/002; U.S. EPA: Cincinnati, OH, 1994.

(14) Olmstead, A.; LeBlanc, G. Effects of Endocrine-active chemicals in the development of sex characteristics of *Daphnia magna*. *Environ. Toxicol. Chem.* **2000**, *19*, 2107–2113.

(15) American Public Health Association, American Water Works Association, and Water Environment Federation. Standard Methods for the Examination of Water and Wastewater, 20th ed.; American Public Health Association: Washington, D.C., 1998.

(16) Loguinov, A. V.; Mian, I. S.; Vulpe, C. D. Exploratory differential gene expression analysis in microarray experiments with no or limited replication. *Genome Biol.* **2004**, *5* (3), R18.

(17) Reich, M.; Liefeld, T.; Gould, J.; Lerner, J.; Tamayo, P.; Mesirov, J. P. GenePattern 2.0. *Nat. Genet.* **2006**, *38* (5), 500–1.

(18) Golub, T. R.; Slonim, D. K.; Tamayo, P.; Huard, C.; Gaasenbeek, M.; Mesirov, J. P.; Coller, H.; Loh, M. L.; Downing, J. R.; Caligiuri, M. A.; Bloomfield, C. D.; Lander, E. S. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **1999**, *286* (5439), 531–7.

(19) Roling, J. A.; Bain, L. J.; Gardea-torresdey, J.; Key, P. B.; Baldwin, W. S. Using Mummichog (*Fundulus heteroclitus*) arrays to monitor the effectiveness of remediation at a Superfund site in Charleston, South Carolina, USA. *Environ. Toxicol. Chem.* **2007**, *26* (6), 1205–1213.

ES800262K