JOURNAL OF PROTEOMICS 75 (2012) 2721–2735

Available online at www.sciencedirect.com

SciVerse ScienceDirect

www.elsevier.com/locate/jprot

The transcriptome and proteome are altered in marine polychaetes (Annelida) exposed to elevated metal levels

Matthew J. Neave[a,*], *Claire Streten-Joyce*[b], *Amanda S. Nouwens*[c], *Chris J. Glasby*[d], *Keith A. McGuinness*[a], *David L. Parry*[b], *Karen S. Gibb*[a]

[a]*Charles Darwin University, Darwin, N.T. 0909, Australia*
[b]*Australian Institute of Marine Science, Darwin, N.T. 0810, Australia*
[c]*The University of Queensland, St Lucia, QLD 4072, Australia*
[d]*Museum & Art Gallery of the Northern Territory, Darwin, N.T. 0801, Australia*

ARTICLE INFO

Article history:
Received 2 November 2011
Accepted 20 March 2012
Available online 29 March 2012

Keywords:
Polychaeta
Biomarkers
Copper
Atox1
Transcriptome

ABSTRACT

Polychaetes are often used in toxicological studies to understand mechanisms of resistance and for biomarker detection, however, we know of only a few genetic pathways involved in resistance. We found the marine polychaete *Ophelina* sp.1 (Opheliidae) in sediment containing high copper levels and investigated this phenomenon by measuring metal accumulation in the worms and changes in gene and protein expression. We sequenced the transcriptome of *Ophelina* sp.1 from both the impacted and reference sediments using 454-sequencing and analysed their proteomes using differential in gel electrophoresis (DIGE). We used the sequenced transcriptome to guide protein identification. Transcripts coding for the copper chaperone, Atox1, were up-regulated in the worms inhabiting the high copper sediment. In addition, genes coding for respiratory proteins, detoxification proteins and cytoskeletal proteins were significantly altered in metal-exposed worms; many of these changes were also detected in the proteome. This dual approach has provided a better understanding of heavy metal resistance in polychaetes and we now have a wider range of suitable indicator genes and proteins for future biomarker development.

1. Introduction

Polychaetes (Annelida) are widespread in the marine environment; they are known to inhabit all the oceanic habitats of the world, from shallow estuaries to deep-sea hydrothermal vents [1]. Consequently, polychaetes often come into contact with wastes generated by human activity. Among the most toxic of these wastes are metals, which can be detrimental to polychaetes even at low concentrations [2,3]. However, certain polychaetes have mechanisms to resist the toxic effects of the metals. This may be achieved through changes in gene expression, using symbiotic bacteria or increasing mucous secretion [4–6]. Moreover, individuals of some species living in metal-polluted environments adapt to the conditions and develop greater metal tolerances than individuals living in clean sediments [7]. The resistance mechanisms used by polychaetes in contaminated environments may lead to the development of useful sub-lethal biomarkers of metal pollution.

One poorly understood aspect of metal resistance in polychaetes is the genetic component [8]. Despite polychaetes often being used in toxicological studies, the molecular mechanisms of polychaete resistance remain largely unknown. Genetic resistance mechanisms have traditionally been difficult to study in polychaetes because of a lack of baseline genetic

* *Corresponding author at:* Charles Darwin University, Darwin, 0909, Australia. Tel.: +61 8 8946 6555, +61 0408002582(Mobile); fax: +61 8 8946 6847.
E-mail address: matthew.neave@cdu.edu.au (M.J. Neave).

1874-3919/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.jprot.2012.03.031

information, forcing researchers to focus only on characterised pathways. These studies have found the alteration of genes involved in oxidative stress responses, including the glutamate cysteine ligase gene, glutathione S-transferase genes and superoxide dismutase genes [8–11]. While the techniques used in these studies provide valuable information about those specific processes, they do not necessarily provide enough information for obtaining a general picture of transcriptome changes. Techniques such as 'next generation' sequencing, or massively parallel sequencing, have great potential to expand sequence databases for non-model organisms [12]. These techniques can sequence hundreds of thousands of messenger RNA transcripts and provide an accurate count of messenger RNA abundance [13]. This means that animals, previously studied because of their ecological relevance, can now also be analysed in functional genomic studies, even if rich genetic resources are unavailable [14].

Since the expression of genes in the transcriptome does not always reflect changes in protein concentration [15,16], complementary protein studies may provide a more comprehensive understanding of expression. Protein expression in polychaetes under metal stress is largely studied using biochemical measurement methods [3]. These studies target and measure a single characterised protein, often finding that polychaetes exposed to metals contain greater concentrations of metallothionein-like proteins and oxidative stress proteins, such as glutathione S-transferase, superoxide dismutase and catalase [17–19]. Few researchers have examined the changes in large regions of the polychaete proteome when under metal stress. One technique, which can be used to examine a large number of proteins, is differential in gel electrophoresis [DIGE; 20]. This technique can be used to accurately and quantitatively examine differences in protein expression associated with stresses and discover novel protein biomarkers [21]. The DIGE technique has previously been limited by the difficulty of identifying proteins from organisms without a well-characterised genome; however, many more proteins can be identified when this technique is used in conjunction with 'next generation' sequencing data.

In this study, we coupled 'next generation' pyrosequencing of polychaete transcriptomes with DIGE analysis of polychaete proteins to obtain a comprehensive insight into the changes associated with metal stress. Our approach was to analyse polychaetes collected from the field in impacted and reference sediment, thus providing a complex but rigorous test of the consequences of metal stress in the real world. We predicted that the transcriptome and proteome would be altered in polychaetes exposed to elevated levels of metals. We also wanted to determine whether transcript abundance was correlated with protein abundance and whether any genes could potentially be developed as biomarkers of metal pollution.

2. Materials and methods

2.1. Study sites

Cullen Bay Marina, hereafter referred to as Cullen Bay, is a constructed lock in Darwin Harbour, Australia (Fig. 1). During the wet season of 1998–1999, an introduced species of mussel, *Mytilopsis sallei* (Récluz, 1849), infiltrated Cullen Bay [22]. To eradicate the mussel, copper sulphate and sodium hypochlorite were used to sterilise the harbour [22]. Cullen Bay still contains high levels of copper resulting from this treatment [23], and is a useful site to test for the effects of metals on gene and protein expression in polychaetes. A reference site for this experiment was located to the east of Darwin, at Dinah Beach. Despite this site being close to the city, previous chemical studies have shown the region to be relatively unpolluted [24].

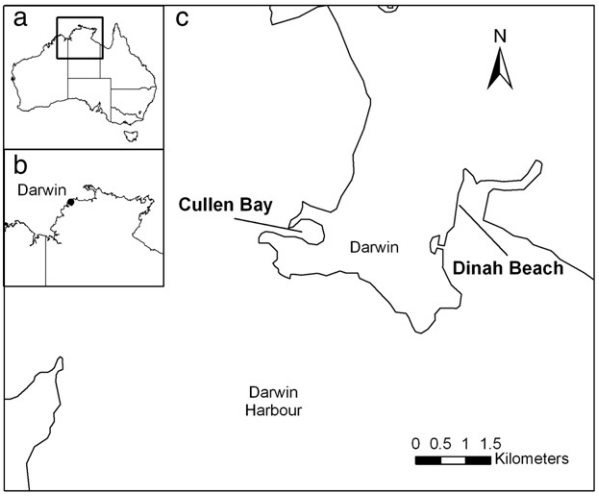

Fig. 1 – Map of the study sites, Darwin, Australia.

Cullen Bay has been recolonised by many marine invertebrates since the copper treatment, including opheliid polychaetes [25]. Opheliids are sub-surface burrowers in sandy or muddy sediments [26] and are generally considered to be non-selective deposit-feeders [27]. These polychaetes are likely to be exposed to metals, as they indiscriminately ingest sediment, and they are constantly in contact with the sediment, which is primarily where metals accumulate [28]. An opheliid polychaete, *Ophelina* sp.1, was chosen as an appropriate species to test for changes in gene and protein expression in the presence of elevated metal levels.

2.2. Sample collection

Samples were collected over approximately 3 weeks, from April 30 to May 21, 2010. Both sites were sampled using a stainless steel Van Veen sediment grab, which recovered approximately 5 L of sediment from the sea floor. The sediment was hand sorted and the *Ophelina* sp.1 specimens removed and placed in RNAlater (Qiagen, Hilden, Germany) for transcriptome analysis, or taken to the laboratory for protein analysis. An additional eight specimens were collected from Cullen Bay and two were collected from Dinah Beach to determine the metal levels in their tissues.

If a grab contained an *Ophelina* sp.1 specimen, the surrounding sediment and porewater were then sampled for chemical analysis. This chemical sampling was conducted on the first two grabs that contained an *Ophelina* sp.1 specimen for each sampling week; therefore, six chemical samples were taken over the three-week sampling period at each site.

For porewater analysis, the sediment was placed into 50 mL acid-washed falcon tubes. For sediment analysis, 0.5 L portions of sediment were placed into zip-lock bags. The samples were then transported to the laboratory on ice.

2.3. Chemical analysis

The concentrations of Al, P, V, Cr, Fe, Mn, Co, Ni, Cu, Zn, Ga, As, Mo, Cd, Pb and U were analysed in both the porewater and sediment samples. The porewater samples were first centrifuged in falcon tubes for 15 min at 3000×*g*. The supernatant was then removed and passed through a 0.45 μm syringe filter, before being analysed for the element concentrations by inductively coupled plasma mass spectrometry (ICP-MS, Agilent 7500ce). The sediment samples were separated into two grain-size fractions, using a 63 μm sieve, which were then dried and weighed. The element concentrations were analysed from the ≤63 μm fraction after a nitric acid digestion at 100 °C for 30 min, 130 °C for 30 min and 200 °C for 30 min, by ICP-MS (Agilent 7500ce). The concentration of total organic carbon (TOC) was also analysed from the ≤63 μm sediment fractions. The sediments were first reacted with concentrated hydrochloric acid to remove inorganic carbonates, then combusted in a LECO furnace at 1400 °C in the presence of strongly oxidising iron/tungsten chips. The evolved carbon was then measured using infrared detection. Quality control for each of the analyses included 4 blanks, 2 spikes and 5 duplicates in every 50 samples analysed. In addition, certified reference materials were added to ensure reliable results. For the sediment digestions, PACS-1 and MESS-3 were included and the porewater analysis included CASS-4 (Institute for National Measurement Standards, National Research Council of Canada). The total organic carbon analysis included Quasimeme reference materials.

To determine the metal concentration in the tissues of *Ophelina* sp.1, entire polychaetes were digested in 2 mL of nitric acid and 2 mL of high pure water using a microwave digestion technique. The microwave was run at 400 W and 200 °C for 15 min, before a 30 min cool down. The resulting digests were analysed for the concentrations of Fe, Mn, Co, Ni, Cu, Zn, As, Cd, Pb and U by ICP-MS (Agilent 7500ce). Quality control included duplicates of the following certified reference materials: DORM-2, AGAL-3 and 1566b oyster (Institute for National Measurement Standards, National Research Council of Canada).

2.4. RNA extraction and 454 sample preparation

At both Cullen Bay and Dinah Beach, 10 *Ophelina* sp.1 specimens were collected and immediately submersed in RNAlater (Qiagen, Hilden, Germany). The samples were placed on ice and taken back to the laboratory for messenger RNA (mRNA) extraction. The mRNA was extracted from each worm according to the methods of McGrath et al. [29] with some modifications. Briefly, the total RNA was extracted using the Qiagen RNeasy Mini Kit (Hilden, Germany) according to the manufacturer's instructions. To reduce contaminating DNA, the samples were then precipitated using lithium chloride (Ambion, Austin, TX, USA). The pellets were resuspended in 15 μL of RNase-free water and separated on a 1.5% agarose gel ran at 100 V for 45 min in 1× TBE buffer (45 mM Tris, 45 mM Boric acid, 1 mM EDTA). The areas between the 28S and 18S and between the 18S and 5S ribosomal RNA bands were expected to contain mRNA and were excised from the gel under UV illumination. The mRNA was then extracted from the gel using the Promega SV Gel and PCR Clean-up System (Madison, WI, USA). The mRNA extracted from the 10 worms at each site was combined to create a single mRNA library for both Cullen Bay and Dinah Beach. The two samples were sequenced using a Roche GS FLX (454) sequencer at the Australian Genome Research Facility in Brisbane, Queensland, using the standard 454 sequencing chemistry. All mRNA sequence data are available in the National Centre for Biotechnology Information (NCBI), Sequence Read Archive under the accession number SRA043945.

2.5. 454 data analyses

Raw 454 reads were assembled into contigs and filtered for quality using the default parameters in the programme GS De Novo Assembler, version 2.3 (Roche). Open reading frames (ORFs) were extracted from the contigs using 'GetORF' in the European Molecular Biology Software Suite [EMBOSS; 30]. ORFs were defined as the sequence between a start and stop codon consisting of at least 50 nucleotides. The ORFs were annotated using BlastX within the standalone BLAST package (www.ncbi.nlm.nih.gov/blast). The BlastX searches were conducted against the NCBI non-redundant (nr) protein database (www.ncbi.nlm.nih.gov, date accessed: 08.02.2011) with an e-value cut-off of 10^{-3}. To determine the functional group of each ORF, they were queried against the eukaryotic orthologous groups database [KOG; 31] with an e-value cut-off of 10^{-3}. In addition, the *Ophelina* sp.1 ORFs were compared to the protein coding gene models of *Capitella teleta* Blake, Grassle and Eckelbarger, 2009, which is the only polychaete with a sequenced genome (http://genome.jgi-psf.org/Capca1/Capca1.home.html). The *Capitella teleta* gene models were converted to a blast database using 'makeblastdb' within the standalone blast package.

To estimate differences in gene expression between the *Ophelina* sp.1 specimens in Cullen Bay and Dinah Beach, the bioconductor package 'DESeq' was used [32]. DESeq uses a negative binomial distribution, which may be more appropriate than other methods for modelling biological variation [33]. Uniquely expressed genes were defined as having at least 200 reads and only being detected at one of the sites. In addition, genes with greater than 10 fold expression differences, and a p-value<0.05 were considered differentially expressed. Only these large differences in expression were considered 'significant' because the samples were combined into one mRNA library at each site, reducing our ability to examine intra-site variability.

2.6. Sequence alignments

Transcripts with interesting expression changes across the two sites were further analysed to check for the presence of key amino acid residues and metal-binding sites. The *Ophelina* sp.1 transcripts coding for Atox1 and globin proteins were compared to homologues in various other polychaete and non-polychaete species (Table 1). These sequences were aligned with characterised NCBI nr database sequences using clustalW in MEGA (Molecular Evolutionary Genetics Analysis) software

Table 1 – Atox1 and globin homologues shown in the multiple alignments (Figs. 4 and 7).

Species	Protein	GenBank Acc.
Homo sapiens	Atox1	NP_004036.1
Oncorhynchus mykiss	Atox1	NP_001154173
Drosophila melanogaster	Atox1	NP_730672
Caenorhabditis elegans	cuc-1	NP_498707
Ophelina bicornis	Myoglobin	AAX73248.1
Alvinella pompejana	Hemoglobin I	CAI56311
Arenicola marina	Hemoglobin A2	CAJ32741.1
	Hemoglobin B2	CAI56309.1
	Hemoglobin B1	AM109951.1
Lumbricus terrestris	Hemoglobin A2	P02218.2
	Hemoglobin A1	P08924.1
	Hemoglobin B2	P13579.1
	Hemoglobin B1	P11069.3
Riftia pachyptila	Hemoglobin A2	P80592.1
	Hemoglobin A1	CAD29154.1
	Hemoglobin B2	CAD29159.1
	Hemoglobin B1	CAD29156.1
Sabella spallanzanii	Hemoglobin B2a	CAC37410.1
	Hemoglobin B2b	CAC37411.1

[34]. Protein structure was added and important amino acid residues were highlighted using GIMP (GNU Image Manipulation Program; www.gimp.org).

2.7. Protein extraction

At both Cullen Bay and Dinah Beach, 12 *Ophelina* sp.1 specimens were collected and taken to the laboratory alive for immediate protein extraction. Individual specimens were homogenised in 200 μL of cold sample buffer (7 M urea, 2 M thiourea, 4% (w/v) CHAPS, 30 mM Tris, 40 mM DTT, pH 8.3) using plastic micropestles. The lysate was centrifuged at 12,000×*g* for 10 min, and the supernatant was removed. The solubilised proteins were then probe sonicated to shear any DNA, and purified using the GE Healthcare protein purification kit (Amersham Biosciences, Piscataway, NJ, USA). The samples were then resuspended in DIGE labelling buffer (7 M urea, 2 M thiourea, 4% (w/v) CHAPS, 30 mM tris, pH 8.3). The protein concentration of each sample was obtained using the 2-D Quant Kit (Amersham Biosciences, Piscataway, NJ, USA). The concentration of each sample was obtained in duplicate and in 2 different dilutions to ensure accurate protein quantification.

2.8. CyDye labelling and differential in gel electrophoresis

A 24 μg sample of each protein extract was labelled with 100 pmol of CyDye DIGE Fluor minimal dye Cy3 or Cy5 (GE Healthcare, Piscataway, NJ, USA) resuspended in fresh dimethylformamide (DMF). Half of the samples from each site were labelled with Cy3 and the other half with Cy5 to ensure site comparisons were not compromised by different dye binding rates. An internal control of pooled protein extracts was labelled with CyDye DIGE Fluor minimal dye Cy2 (GE Healthcare, Piscataway, NJ, USA). The labelling reactions were carried out at 4 °C for 30 min and then quenched with 10 mmol lysine.

Since the three dyes (Cy2, Cy3 and Cy5) fluoresce at different wavelengths, proteins, which are labelled with different dyes, can be run on the same gel. Differential in gel electrophoresis was performed 12 times, with each sample containing a labelled protein extract from Cullen Bay, Dinah Beach and the internal control. The volume of each of the combined samples was increased to 450 μL with rehydration buffer (7 M urea, 2 M thiourea, 4% (w/v) CHAPS, 30 mM Tris, 40 mM DTT, 1% (v/v) pH 3-11NL pharmalytes (GE Healthcare, Piscataway, NJ, USA), pH 8.3) and actively rehydrated at 30 V for 12 h on 24 cm pH 3–11 non-linear IPG strips (GE Healthcare, Piscataway, NJ, USA). Proteins were focused for 1 h at 100 V, 1 h at 500 V, 1 h for 1000 V, and finally 7 h at 5000 V for a total of 30 kVh. The IPG strips were then equilibrated in two steps using equilibration base buffer (6 M urea, 0.075 M Tris.HCl, 2% (w/v) SDS, 20% (v/v) glycerol). The strips were first placed in equilibration base buffer with added 1% (w/v) DTT for 15 min, then placed in equilibration base buffer with added 2.5% (w/v) iodoacetamide for a further 15 min. The strips were embedded on top of 14% SDS-PAGE gels using 1% agarose in standard Laemmli running buffer (192 mM glycine, 25 mM Tris, pH 8.3, 0.1% (w/v) SDS) containing a small amount of bromophenol blue. SDS-PAGE was performed in the standard Laemmli running buffer at 2 W/gel for 45 min, 5 W/gel for 30 min, then 15 W/gel until the dye front reached the bottom of the gels (approximately 4 h). The gels were imaged using a Typhoon 9400 Variable Mode Imager (Amersham Biosciences, Piscataway, NJ) for each of the fluorescent dyes. The different gels and dyes were compared to each other using Progenesis SameSpots, version 4.1 (Nonlinear Dynamics). The internal control, which was identical across the 12 gels, was used to precisely match the gels. Differentially expressed proteins were recorded if the p-value was <0.05 and the protein concentration was more than 1.5 times altered.

2.9. Protein identification

Several replicate gels were run as previously, except that the gels contained 500 μg of unlabelled pooled protein extract. After electrophoresis the gels were stained with 0.1% (w/v) Coomassie Brilliant Blue G-250 in 17% (w/v) ammonium sulphate, 34% (v/v) methanol and 3.6% (v/v) *ortho*-phosphoric acid for 3 days. Gels were destained in 1% (v/v) acetic acid for 1 h and imaged with a LI-COR Odyssey Imaging System. Proteins were targeted for identification if their expression was significantly altered at Cullen Bay, and they were sufficiently abundant, therefore, 21 protein spots with altered expression were selected for identification. We also targeted highly abundant proteins for identification to generate a protein 'reference' map for this potentially important polychaete species. Consequently, a further 25 highly expressed proteins were selected for identification, increasing the total number to 46. The target proteins were excised and destained in 50 mM ammonium bicarbonate and 50% (v/v) acetonitrile. The gel pieces were then dehydrated in acetonitrile and rehydrated in 10 ng/μL trypsin in 50 mM ammonium bicarbonate and incubated at 37 °C for 16 h. Peptides were extracted from the gel pieces using 50% (v/v) acetonitrile and 0.1% TFA by sonicating the samples for 10 min followed by concentration to 2 μL in a speed-vac. A 0.5 μL aliquot was taken from each sample and combined with 0.5 μL of matrix (10 mg/mL α-cyano-4-hydroxycinnamic acid/25 mM diammonium citrate/50% acetonitrile/0.1% TFA) and allowed to dry. Tryptic peptide mass fingerprints (PMFs) and MS/MS spectra were obtained

using an Autoflex II MALDI-TOF/TOF (Bruker Daltonics, GmbH, Bremen, Germany). PMFs were queried against the National Centre for Biotechnology Information NCBI nr protein database and the Swissprot database using MASCOT accessed via the Australian Proteomics Computational Facility. Searches were performed using all species, trypsin as enzyme, 2 miscleavages, and mass tolerance of 0.5 Da, and modification of Cys with carbamidomethylation (fixed) and Met with oxidation (variable) also included. Similarly MS/MS data were also searched. For MS/MS spectra that were not matched in MASCOT, spectra were de novo sequenced. Using this technique, the amino acid sequences of several peptides from each protein were obtained. The amino acid sequences were then queried against the translated *Ophelina* sp.1 mRNA transcripts obtained from the transcriptome analysis using blastp in the standalone BLAST package. Only translated transcripts with significant homology to a database protein ($e<10^{-3}$) were used as a database for the amino acid sequences. This ensured that the transcripts were protein-coding mRNA and that they were translated into the correct reading frame. Identification was assigned to a protein when the amino acid sequences were both identical and unique to a translated transcript.

2.10. *Atox1 western blots*

Western blots were used to validate the up-regulation in Cullen Bay polychaetes of an mRNA transcript with homology to the copper chaperone, Atox1. The same protein extracts from the DIGE experiments were used for western blot analysis. After the DIGE experiments, 5 Cullen Bay samples and 7 Dinah Beach samples contained enough protein for use in the western blot experiments. A 100 μg protein extract from each sample was separated using Mini-Protean TGX—Any kD precast gels (BioRad, USA). The proteins were then transferred to 0.2 μM nitrocellulose membranes (BioRad, USA) using the Trans-Blot Turbo Transfer System (BioRad, USA). The Atox1 protein was detected using a mouse monoclonal antibody to Atox1 amino acids 1–69 (Abcam, Cambridge, UK) and the WesternDot™ 625 Western Blot Kit (Invitrogen, Paisley, UK), according to the manufacturer's instructions with some minor modifications. The primary antibody was used at a concentration of 0.1 μg/mL and the incubation times for the primary and secondary antibodies were reduced to 30 min. A positive control (Atox1 human, Abcam, Cambridge, UK) was run with the polychaete samples. The western blots were imaged under UV illumination and the relative intensities were obtained using ImageJ [35].

3. Results

3.1. *Chemical composition of the sites and polychaetes*

In the sediments at Cullen Bay, the concentration of two metals, copper and zinc, exceeded ISQG-Low sediment quality guidelines [Table 2; 36]. The concentration of copper was approximately 7 times higher, and the concentration of zinc was over 2 times higher, in Cullen Bay sediments compared to Dinah Beach. All other metal concentrations in the sediments were below the ISQG-Low sediment quality guidelines. In the Dinah Beach sediments, the metal concentrations were all below the ISQG-Low sediment quality guidelines [36], indicating that this site was an appropriate reference site.

In the porewaters at Cullen Bay, copper and zinc concentrations were above marine water guideline values for the protection of 95% of species but below the values for the protection of 90% of species [36]. In the porewaters at Dinah Beach, all metal concentrations were lower than the guidelines for the protection of 99% of species [36].

The *Ophelina* sp.1 specimens at Cullen Bay contained over 4 times more copper, and approximately 1.4 times more zinc, than *Ophelina* sp.1 specimens inhabiting Dinah Beach (Table 2). These results may be confounded, however, because the tissue digests also included sediment-bound metals from the gut of the worms. Ideally the animals would be kept alive in the laboratory until the digestive tract is empty. Unfortunately we

Table 2 – Metal concentrations in the sediments (dry weight), porewaters and *Ophelina* sp.1 tissue (wet weight) at the two study sites, Cullen Bay and Dinah Beach. The standard error is included.

	Cullen Bay sediment	Dinah Beach sediment	Cullen Bay porewater	Dinah Beach porewater	Cullen Bay *Ophelina* sp.1	Dinah Beach *Ophelina* sp.1
	ppm	ppm	ppb	ppb	ppm	ppm
Al	55,800±467	56,600±458	5.61±3.9	2.24±0.238		
P	471±9.6	418±35.8	429±45.5	34.5±6.48		
V	58.7±0.703	60.1±0.828	1.63±0.15	0.816±0.0347		
Cr	26.1±0.87	27.1±1.71	0.1±0.00491	0.0689±0.00713		
Fe	30,000±455	33,600±930	474±37.3	2400±426	2050±161	3390±169
Mn	227±0.917	138±6.79	755±82.6	1130±100	18.7±1.59	19.1±1.98
Co	9.58±0.0485	6.88±0.587	0.252±0.0191	0.404±0.0754	0.802±0.0647	1.05±0.117
Ni	20.5±0.116	17.8±0.892	0.599±0.0558	1.32±0.172	2.05±0.172	2.68±0.676
Cu	140±2.73	19.9±2.47	1.82±0.569	0.656±0.293	46±7.65	9.05±4.71
Zn	235±10.6	89.1±16.1	12.5±0.836	5.3±0.771	43.1±5.99	31.4±5.29
Ga	16.6±0.116	16.1±0.124	0.018±0.00803	0.0473±0.00778		
As	18.5±0.359	12.3±0.379	3.71±1.18	8.09±0.552	30.5±8.19	31.4±2.21
Mo	1.84±0.0471	1.75±0.0664	6.83±1.13	5.69±1.54		
Cd	0.0858±0.00124	0.148±0.0265	0.0137±0.00373	0.0396±0.0296	0.0478±0.0273	0.0569±0.0223
Pb	20.7±0.186	22.7±2.14	0.0731±0.0216	0.0734±0.023	1.3±0.0921	2.34±0.585
U	2.49±0.0446	1.78±0.045	2.42±0.167	4.68±0.988	0.143±0.00857	0.188±0.07

could not do this, so the samples were a mixture of metals in the tissue and metals from the sediment in the gut of the worms. To overcome this problem, the sediment metal concentrations were plotted against the tissue metal concentrations (Fig. 2). Any metals which fall markedly above the line of best fit are at a higher concentration in the tissues than would be expected if the tissue digest results were only a reflection of the sediment in their gut. In Fig. 2, As, Cu and Zn appear to be accumulating in the polychaete tissues, while U, Co, Pb and Ni may be actively excreted.

3.2. *Transcriptome sequences*

Of 647,439 total RNA fragments sequenced, 421,105 reads fully assembled into 6477 contigs at Cullen Bay and 5960 contigs at Dinah Beach (Table 3). In these contigs, over 10,000 open reading frames (ORFs) were detected (Table 3). The ORFs were then translated into amino acid sequences and queried against protein databases using BlastX. At both Cullen Bay and Dinah Beach, approximately 20% of the ORFs (corresponding to 40% of the contigs) were significantly matched to a protein on the NCBI non-redundant database. A similar number of ORFs were matched to a protein group in the eukaryotic orthologous groups database (KOG). When the ORFs were queried against the sequenced polychaete genome of *Capitella teleta*, the significant matches increased to approximately 33% (corresponding to over 60% of the contigs).

The 25 functional groups in the KOG database provided general functions for the *Ophelina* sp.1 ORFs (Fig. 3). The percentage of genes in each functional group was similar for both the Cullen Bay and Dinah Beach *Ophelina* sp.1 specimens (Fig. 3). The functional group that contained the most genes was 'cytoskeleton'. This group was dominated by genes coding for myosin, tubulin and actin. Some functional groups showed slight differences: Cullen Bay worms expressed more genes coding for extracellular structures and cell motility, and more genes involved in replication, recombination and repair. The *Ophelina* sp.1 specimens from Cullen Bay expressed 6 genes for DNA repair and nucleotide excision repair, while in Dinah Beach worms only 2 DNA repair genes were detected.

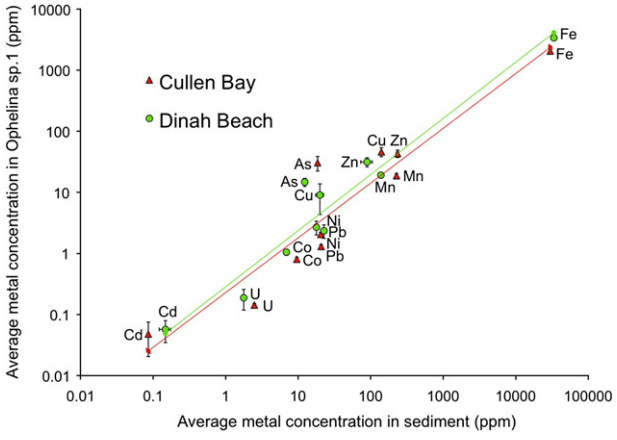

Fig. 2 – Regression analysis of metal concentrations in *Ophelina* sp.1 tissues plotted against metal concentrations in sediments. Both axes were log-transformed and a line of best fit for both sites was plotted. Standard error bars are included.

Table 3 – Sequencing, assembly and annotation results for *Ophelina* sp.1 transcriptomes from Cullen Bay and Dinah Beach.

	Cullen Bay	Dinah Beach
Sequencing results		
Total reads	350,369	297,070
Assembly results		
Reads assembled to contigs	223,856	197,249
Singleton reads	62,333	56,803
Contigs	6477	5960
Average contig size (bp)	370	467
Gene annotation		
Predicted open reading frames (ORFs)	11,936	14,533
ORFs with NCBI nr match (e<10-3)	2462	2731
ORFs with KOG match (e<10-3)	2380	2814
Capitella teleta gene match (e<10-3)	3911	4851

3.3. *Gene expression*

Uniquely expressed genes were defined as having at least 200 reads and only being detected at one of the sites: 14 characterised genes meet these criteria (Table 4). Two genes similar to the oxygen transport proteins, hemoglobin B1 and cytochrome c oxidase VIb-like, were only detected at Cullen Bay. In addition, the copper transport and antioxidant protein Antioxidant 1 (Atox1) was only detected in Cullen Bay polychaetes. At Dinah Beach, the uniquely expressed genes were involved in reproduction, digestion and antimicrobial activity. Transcripts coding for the peptide regulating enzyme, alanine aminopeptidase, were also much more abundant at Dinah Beach.

Differential expression of genes at the two sites was recorded when there were at least 10 times more copies of a gene at one of the sites and the p-value was <0.05. Most of the genes meeting these requirements were up-regulated at Cullen Bay and again coded for many cytoskeletal proteins (Table 5). Sulfotransferase 1C-like and β-thymosin were also up-regulated at Cullen Bay.

3.4. *Atox1 alignment*

Transcripts coding for the copper chaperone, antioxidant 1 (Atox1), were only found in Cullen Bay polychaetes (Table 4). This protein might be a useful biomarker, as its up-regulation suggests that it may play a role in copper homeostasis, so it was analysed further. The translated amino acid sequence of the *Ophelina* sp.1 Atox1 homologue was aligned with Atox1 homologues in human (*Homo sapiens*), rainbow trout (*Oncorhynchus mykiss*), fruit fly (*Drosophila melanogaster*) and nematode (*Caenorhabditis elegans*) (Fig. 4). The Atox1 homologue in *Ophelina* sp.1 contained 15 conserved amino acids across the region, including the two cysteine residues that bind copper [37].

3.5. *Protein expression*

Across all of the samples and sites, a total of 576 proteins were seen on the 2 DE gels (Fig. 5). Of these proteins, the expression of 77 was significantly ($P<0.05$) and more than 1.5 times altered at

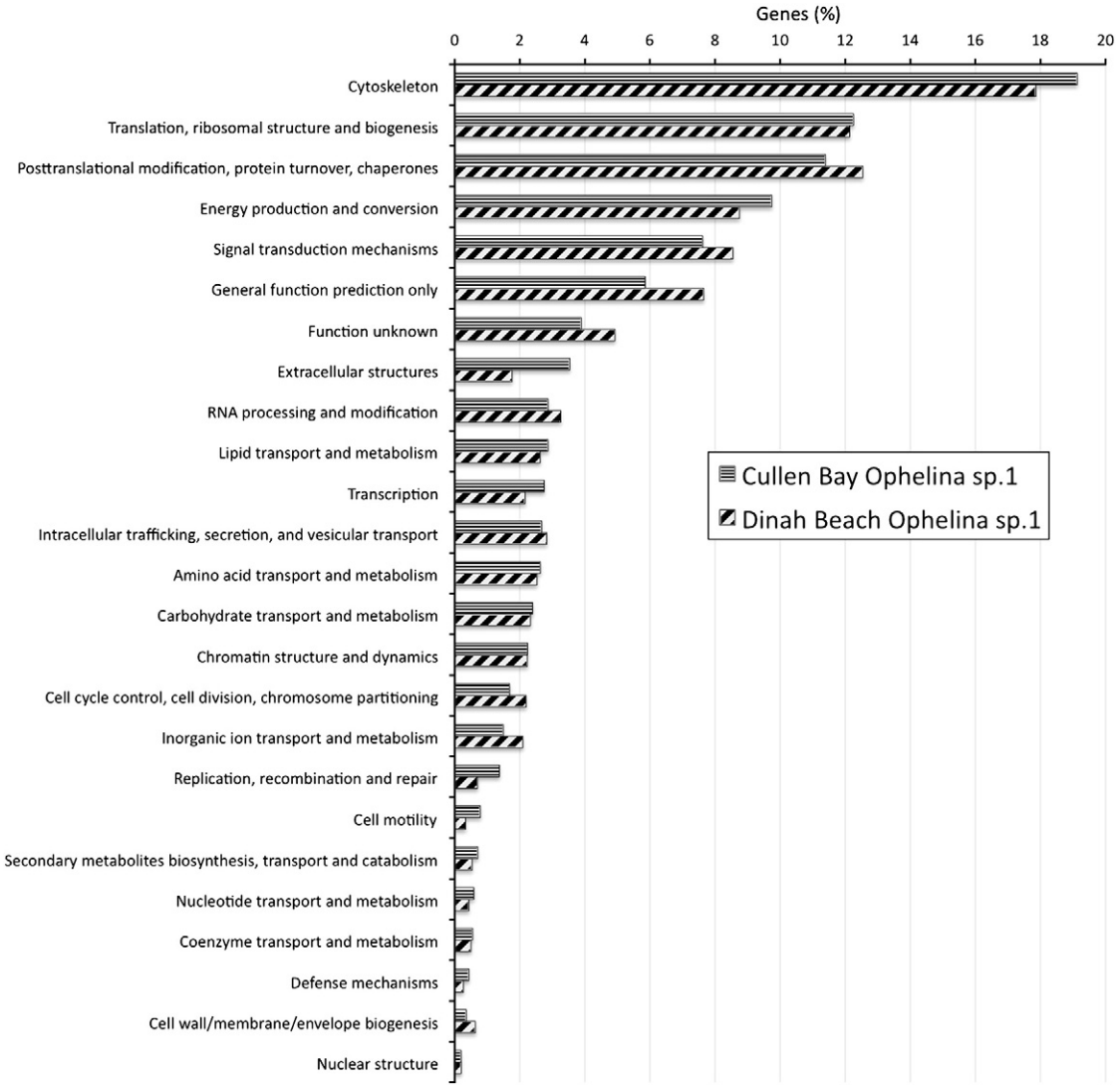

Fig. 3 – The percentage of genes from Cullen Bay and Dinah Beach *Ophelina* sp.1 specimens in each of the 25 functional groups of the KOG database.

Cullen Bay: 50 were up-regulated and 27 were down-regulated. Of 46 protein spots excised from the gels, only 24 could be confidently identified (Fig. 5 and Table 6). Many of these were cytoskeletal proteins, including actin, myosin, paramyosin and intermediate filament proteins. Other proteins were peptidases, globin molecules and proteins containing LIM domains. Many of these proteins were not differentially expressed across the sites but were abundant and important inclusions for building a protein 'reference' map.

Only 9 of the proteins that were identified (Table 6) were differentially expressed at the sites. The expressions of the differentially expressed proteins were compared to the mRNA transcript expression levels and many similarities were observed (Table 7). All of the protein spots that were significantly up-regulated at Cullen Bay had corresponding mRNA transcripts that were also up-regulated. There was less agreement between the protein spots that were down-regulated: two of the down-regulated proteins were significantly up-regulated in the transcriptome. The hemoglobin chain B1 was less abundant in the proteome of worms at Cullen Bay; however, mRNA transcripts coding for this hemoglobin chain were more abundant at Cullen Bay. A protein spot identified as paramyosin (spot #20) was down-regulated at Cullen Bay, although a spot similarly identified as paramyosin (spot #14) was up-regulated at Cullen Bay. The transcript coding for paramyosin was also up-regulated at Cullen Bay. This suggests truncation or degradation of the large paramyosin protein.

3.6. *Atox1 western blots*

mRNA transcripts that coded for a protein homologous with Atox1 were up-regulated in Cullen Bay polychaetes, but this protein was not detected in the DIGE protein experiments (Table 6). One reason might be that DIGE is unlikely to resolve small proteins such as Atox1 [38]. Western blots were used to validate the up-regulation of this protein in Cullen Bay worms. A single band of the correct weight for the Atox1 protein was detected in 3 Cullen Bay worms from 5 tested and 2

Table 4 – Gene homologues that were only detected at one of the sites with a minimum of 200 transcript copies.

Gene homologue	Amino acid identity (%)	Organism	GenBank Acc.	Cullen reads	Dinah reads
Slow myosin heavy chain	63	*Dicentrarchus labrax*	CBN81811.1	1703	0
Myosin II heavy chain	67	*Ilyanassa obsoleta*	AAD13782.1	1527	0
Hemoglobin B1	40	*Riftia pachyptila*	CAD29156.1	538	0
Intermediate filament protein	68	*Phascolion strombus*	CAB38180.1	398	0
Cytochrome c oxidase subunit VIb-like	57	*Drosophila mojavensis*	XP_002008311.1	384	0
acidic ribosomal protein	65	*Culex quinquefasciatus*	XP_001857537.1	302	0
Atox1	59	*Aedes aegypti*	XP_001656348.1	210	0
Ovulatory protein-2 precursor-like	47	*Strongylocentrotus purpuratus*	XP_001202083.1	0	2482
Alanine aminopeptidase-like	34	*Saccoglossus kowalevskii*	XP_002738530.1	0	1096
Lysozyme	67	*Eisenia andrei*	ABC68610.1	0	697
Ribosomal protein rpl21ap	76	*Eurythoe complanata*	ABW23224.1	0	609
Microspherule protein-like	32	*Strongylocentrotus purpuratus*	XP_797067.2	0	565
ATP synthase subunit beta	85	*Acyrthosiphon pisum*	NP_001119645	0	412
Peritrophic matrix protein-like	36	*Tribolium castaneum*	NP_001161929.1	0	239

Dinah Beach worms from 7 tested. The Atox1 protein in Cullen Bay worms was generally at higher concentrations than in Dinah Beach worms (Fig. 6).

3.7. *Polychaete globin alignment*

Four different sequences coding for globin proteins were detected in the *Ophelina* sp.1 transcriptomes: one coded for intracellular myoglobin and three for extracellular hemoglobins (Fig. 7). The amino acid sequences were identical in Cullen Bay and Dinah Beach *Ophelina* sp.1 specimens. One of the globin chains was differentially expressed at the sites in both the transcriptome and proteome of the worms, warranting further analysis. The globin amino acid sequences were aligned with the globin sequences of several different annelid species (Fig. 7). Each of the *Ophelina* sp.1 extracellular hemoglobin chains was then categorised based on the presence or absence of several key amino acid residues and the overall sequence similarity. One *Ophelina* sp.1 globin chain was categorised into the A family of globin chains. This was confirmed by the presence of a lysine residue (K) at position A9, a tryptophan residue (W) at position B10 and residue deletions between the A an B helices [Fig. 7; 39]. Another two globin chains were categorised as part of the B family of globins based on the presence of an aspartic acid residue (D) at position A4, a tryptophan residue (W) at position A16, a phenylalanine residue (F) at position B10 and insertions between the A and B helices and between the F and G helices [Fig. 7; 39]. The further division of the globin chains into A and B subtypes was completed with the help of a phylogenetic tree (not shown). The *Ophelina* sp.1 hemoglobins were best categorised as chains A2, B1a and B1b. No hemoglobin chains matching to A1 or B2 were detected. The globin chain that was up-regulated in the transcriptome was identified as extracellular hemoglobin chain B1a. This transcript was confirmed as coding for the same protein that was down-regulated in the proteome by identical amino acid matches from two peptide fragments (Fig. 7).

4. Discussion

Several genes and proteins in specimens of *Ophelina* sp.1 collected from Cullen Bay were differentially expressed compared to the reference site. These different expression levels may be caused by contaminants in the surrounding sediment, as *Ophelina* sp.1 is a subsurface deposit-feeding polychaete

Table 5 – Gene homologues that were differentially expressed at one of the sites by at least 10 fold. Fold changes are relative to Cullen Bay. Positive numbers indicate up-regulation and negative numbers indicate down-regulation.

Gene homologue	Amino acid identity (%)	Organism	GenBank Acc.	Fold change relative to Cullen Bay
Extracellular serine proteinase	58	*Thermus* sp. Rt41A	P80146	97.81
Myosin regulatory light chain	74	*Riftia pachyptila*	AAG16892.1	51.70
Probable protein brick1	90	*Sus scrofa*	XP_003132409.1	−33.00
Sodium-dependent phosphate transport protein	53	*Sus scrofa*	XP_003128941.1	−29.50
Antistasin-like	43	*Haliotis discus discus*	ACJ12609.1	27.39
Phosphatidylinositol glycan anchor-like	68	*Ornithorhynchus anatinus*	XP_001508401.1	−24.00
β-thymosin	63	*Triatoma infestans*	ABR27867.1	23.87
Sulfotransferase 1C-like	64	*Canis familiaris*	XP_531771.1	22.88
Protein phosphatase regulatory subunit 3B-B	43	*Xenopus laevis*	NP_001085518.1	22.47
TEPP protein	73	*Lottia gigantea*	DAA06552.1	22.42
Placenta-specific gene 8 protein-like	44	*Oncorhynchus mykiss*	ACO08545.1	12.88
Paramyosin	57	*Haliotis discus discus*	BAJ61596.1	12.39
Alcohol dehydrogenase	76	*Camponotus floridanus*	EFN64102.1	11.28
Ribonuclease h1-like	33	*Caligus rogercresseyi*	ACO11369.1	10.68

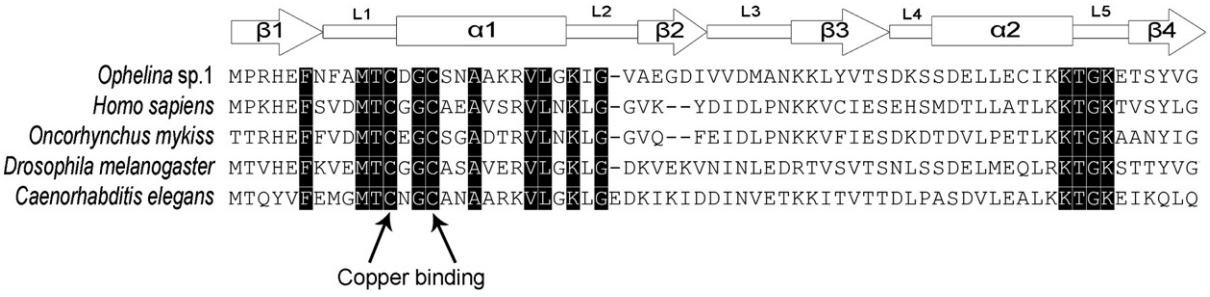

Fig. 4 – Amino acid alignments of Atox1 homologues in diverse species (GenBank accessions are given in Table 1). Atox1 structure is based on Rosenzweig et al. [37], β refers to beta sheets, α refers to alpha helices and L refers to loops.

and is constantly in contact with the sediment [26,27]. In the sediment at Cullen Bay, the concentrations of copper and zinc were significantly higher than at the reference site and copper appeared to be accumulating in the tissues of *Ophelina* sp.1. Arsenic also accumulated in polychaete tissues but, as the levels were similar at the two sites, it was not likely to cause differential expression between the sites. Based on the chemical analysis, the metal most likely to cause different expression at Cullen Bay was copper. An excess of copper in cells can result in the production of reactive oxygen species (ROS), which cause oxidative stress and DNA damage [41,42]. Copper can also damage the iron–sulphur dehydratase protein family when in high concentrations [43]. Copper levels in Cullen Bay were above the ISQG-low guidelines for sediment quality [36], and it is likely that changes in the expression of certain genes and proteins in *Ophelina* sp.1 were related to dealing with the excess copper.

In the *Ophelina* sp.1 specimens in Cullen Bay, a gene coding for a protein similar to the copper chaperone, antioxidant 1 (Atox1), was up-regulated. This protein was also detected using western blots and was more abundant in Cullen Bay worms. Atox1 is a protein that has been linked to maintaining copper homeostasis within cells and may also have antioxidant properties [41,44]. The Atox1 homologue in the *Ophelina* sp.1 inhabiting Cullen Bay may be up-regulated in order to deal with the excess copper levels in their cells. One strategy may be to bind excess intracellular copper and transport it to the secretory pathway [41]. In support of this suggestion there is evidence from other systems that Atox1 binds copper and transports it to ATPases where it is then transported into the bile for incorporation into secreted copper-dependent proteins [45]. Copper binding in Atox1 occurs via the thiol groups on free cysteine residues between its first loop and first alpha helix [37]. The homologue in *Ophelina* sp.1 also contained these free cysteine residues, suggesting that it is also capable of binding copper. In addition, the Atox1 homologue may be involved in mitigating oxidative stress, which is a likely consequence of excess intracellular copper [41]. Kelner et al. [44] reported that the highest concentration of Atox1 was in cells that contained high levels of copper and that Atox1 protected these cells against oxidative stress. Moreover, when Atox1 is over-expressed it may be able to substitute for superoxide dismutase and prevent oxidative damage [46].

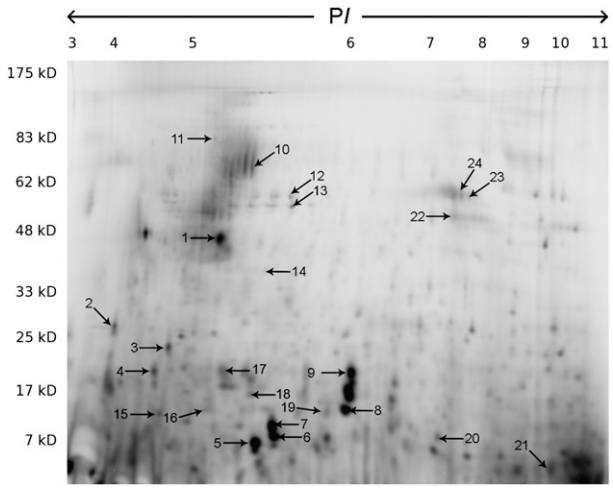

Fig. 5 – Representative DIGE gel image showing the 24 identified protein spots and their approximate isoelectric points (pI) and molecular weights (kD).

Atox1 homologues have been detected in many organisms, from worms to vertebrates, providing the potential for this protein to be used as a biomarker in many different species and environments [41,46,47]. Moreover, the Atox1 protein in *Ophelina* sp.1 was detected using western blots with commercially available antibodies designed for use in human experiments, removing the need to develop specific antibodies. In addition, Atox1 may be specifically induced by copper contamination. This is important, as many of the proteins regularly identified as differentially expressed in toxicological studies are part of a general cellular stress response [38] and are, therefore, not indicative of specific contamination types. Biomarkers that are specific for contamination types are important in understanding cause–effect relationships and are required for the targeted mitigation of contaminants. There is, however, some evidence to suggest that Atox1 is inducted by oxidative stress, which could be caused by a range of factors. As a consequence, more work is required to determine if copper specifically induces Atox1.

Several genes coding for intracellular and extracellular globins were detected in the *Ophelina* sp.1 transcriptome, including myoglobin, and the extracellular hemoglobin chain A2 and two B1 chains. The *Ophelina* sp.1 hemoglobins contained all five of the conserved amino acid residues present across the region and the cysteine residues involved in the formation of the hemoglobin disulphide bridge [39]. *Ophelina* sp.1 hemoglobins did not, however, contain the free cysteine residues thought to bind sulphide in polychaetes inhabiting sulphide-rich

Table 6 – *Ophelina* sp.1 proteins identified by MALDI-TOF/TOF and peptide mass fingerprints (PMF).

Spot #	Peptide sequence	Transcript match	Gene homologue	Organism	GenBank Acc.	MW (kD)		pI	
						Theor.	Observ.	Theor.	Observ.
1	QPKQLSDYEAEINLLR	Isogroup 00090	Cytoplasmic intermediate filament protein	*Phascolion strombus*	CAB38180.1	68.9	48	5.65	5.1
2	VNTNELR EFIDGEGTGFVQR	Isogroup 00184	Troponin-C	*Perinereis vancaurica tetradentata*	BAB18897.1	17.4	26	4.33	4.0
3	*Identified by PMF* (MOWSE score 90)		Beta actin	*Gerarcinus lateralis*	AAL40077.1	42.0	24	5.41	4.9
4	TYQVYR ELEDALEMER DYQEISFQF	Isogroup 00681	Paramyosin	*Haliotis discus discus*	BAJ61596.1	99.5	20	5.36	4.5
5	*Identified by PMF* (MOWSE score 96)		Actin	*Luciola cruciata*	BAH79108.1	8.3	7	5.74	5.4
6	*Identified by PMF* (MOWSE score 104)		Beta actin	*Acipenser schrenckii*	AAV65947.1	14.6	8	5.45	5.5
7	*Identified by PMF* (MOWSE score 82)		Beta actin	*Acipenser schrenckii*	AAV65947.1	14.6	10	5.45	5.5
8	*Identified by PMF* (MOWSE score 83)		Skeletal muscle actin	*Homarus americanus*	ACI23565.1	41.8	15	5.23	6.0
9	*Identified by PMF* (MOWSE score 76)		Actin	*Apis mellifera*	BAA74592.1	14.6	18	5.37	6.0
10	VEIETR	Isogroup 00681	Paramyosin	*Haliotis discus discus*	BAJ61596.1	99.5	65	5.36	5.4
11	QFASFIDKVR	Isogroup 00090	Cytoplasmic intermediate filament protein	*Phascolion strombus*	CAB38180.1	68.9	80	5.65	5.2
12	MSLGGGGS GAAGSGSFAAITIQDAAR	Isogroup 00115	Peptidase s8 and s53	*Stigmatella aurantiaca*	YP_003954556.1	39.6	55	5.45	5.5
13	ATVGAIQGLR	Isogroup 02258	Subtilisin-type proteinase	*Paracoccidioides brasiliensis*	EEH21742.1	53.3	50	5.79	5.8
14	HTYQVYR ELEDALEMER DYQEISFQ	Isogroup 00681	Paramyosin	*Haliotis discus discus*	BAJ61596.1	99.5	35	5.36	5.4
15	INNLGDIR IWNSP	Isogroup 01892	Hemoglobin B1a	*Riftia pachyptila*	CAD29156.1	12.8	12	5.41	4.8
16	QWQSAER ELEVALDGANR	Isogroup 00016	Myosin heavy chain	*Doryteuthis pealeii*	AAC24207.1	221.5	13	5.47	5.0
17	LGLELEIAAYR	Isogroup 00090	Cytoplasmic intermediate filament protein	*Phascolion strombus*	CAB38180.1	67.8	20	5.59	5.1
18	DILCPDCG	Isogroup 00485	LIM-9-like	*Saccoglossus kowalevskii*	XP_002731816.1	54.8	16	4.95	5.3
19	SYELPDGQVITIGNER VWHHTFYNELR	Isogroup 00336	Actin	*Haliotis iris*	AAX19286.1	41.7	15	5.29	5.8
20	HTYQVYR	Isogroup 00681	Paramyosin	*Haliotis discus discus*	BAJ61596.1	99.5	7	5.36	7.0
21	DLYANTVLSGGTTMYPGIADR	Isogroup 00037	Actin	*Hypsibius klebelsbergi*	ADK66838.1	41.7	5	5.30	9.5
22	CHDNNFAAR	Isogroup 00485	LIM-9-like	*Saccoglossus kowalevskii*	XP_002731816.1	54.8	50	4.95	7.4
23	KDILCPDCG	Isogroup 00485	LIM-9-like	*Saccoglossus kowalevskii*	XP_002731816.1	54.8	55	4.95	7.6
24	DILCPDCG HDNNFAAR FISFEER	Isogroup 00485	LIM-9-like	*Saccoglossus kowalevskii*	XP_002731816.1	54.8	55	4.95	7.5

Table 7 – Differentially expressed proteins at Cullen Bay compared to their corresponding transcript regulation.

Spot #	Protein identification	Protein regulation at Cullen Bay	Transcript expression at Cullen Bay
11	Intermediate filament protein	2.3× up	>300× up
14	Paramyosin	2.2× up	12× up
18	LIM-9-like	1.9× up	>100× up
1	Intermediate filament protein	1.6× up	>300× up
22	LIM-9-like	1.5× up	>100× up
2	Troponin-C	2.6× down	1.3× down (not sig)
15	Hemoglobin B1	1.5× down	>500× up
19	Actin	1.5× down	1.1× down (not sig)
20	Paramyosin	1.5× down	12× up

environments [48]. This may not be surprising given the relatively low-sulphide, coastal habitat of *Ophelina* sp.1. One of the extracellular hemoglobin chains, B1a, was significantly over-expressed in the transcriptome of *Ophelina* sp.1 inhabiting Cullen Bay compared to Dinah Beach. This same hemoglobin chain was, however, down-regulated in the proteome of *Ophelina* sp.1 in Cullen Bay. It is possible that the hemoglobin protein had been damaged by copper-induced ROS, which resulted in reduced expression in the proteome. In fact, ROS has long been known to degrade heme proteins [49]. To compensate for this reduction in the proteome, transcripts coding for the hemoglobin chain may then be up-regulated. Kim et al. [50] reported an increase in the expression of a hemoglobin gene in daphnids under oxidative stress. Although several hemoglobin chains were detected in the *Ophelina* sp.1 transcriptome, only one showed altered expression at both the gene and protein level. This suggests that this particular hemoglobin chain, B1a, is more susceptible to ROS damage than the other hemoglobin chains. Fedeli et al. [51] also reported differences in the susceptibility of different trout hemoglobin chains to ROS damage, attributing the differences to their structural and functional properties.

High levels of cellular copper can catalyse the production of ROS through Fenton reactions [52], which in turn can cause DNA damage [42,53]. We found more genes coding for DNA repair proteins expressed in the Cullen Bay worms, despite the Dinah Beach library containing more genes overall. This suggests that the Cullen Bay worms were under stress from ROS and were experiencing greater levels of DNA damage. Another transcript that was up-regulated at Cullen Bay, and may be related to reducing oxidative stress, was β-thymosin. These small proteins are actin regulating and may also be anti-oxidants and play a role in inflammation control and wound healing [54,55]. In addition, several up-regulated protein spots from Cullen Bay samples were identified as homologous to

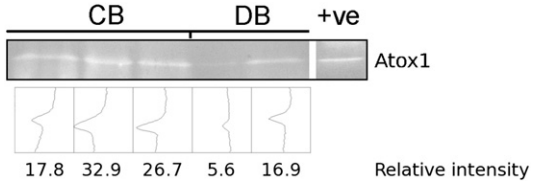

Fig. 6 – Western blot detection of Atox1 protein of approximately 8 kD from Cullen Bay (CB) and Dinah Beach (DB) polychaetes. A positive control (+ve; human Atox1) is included and the relative intensities of each band are given.

Lim-9-like proteins—transcripts coding for this protein were also significantly up-regulated at Cullen Bay. The Lim-9-like homologue was closely related to DRAL/FHL2 proteins, which have many functions and may be up-regulated in response to DNA damage [56,57]. Although the function of the Lim-9-like homologue in *Ophelina* sp.1 is difficult to ascertain, the fact that two Lim-9-like proteins and the Lim-9-like transcript were all up-regulated at Cullen Bay suggests it may have a role in mitigating the effects of excess copper. Another up-regulated gene in Cullen Bay polychaetes was homologous with sulfotransferase 1C-like (SULT1C-like). Sulfotransferases detoxify various drugs and xenobiotics and this particular subfamily, SULT1, exhibits activity towards phenolic compounds [58,59]. Although these enzymes are not known to respond to metal contamination, Ling et al. [60] reported that a member of the sulfotransferase family was up-regulated in fish exposed to cadmium. In addition, Hattori et al. [58] reported that a sulfotransferase was up-regulated in nematodes during their sensitive 'dauer' larval-stage, possibly functioning as a defence against xenobiotics. It is likely that the sulfotransferase in *Ophelina* sp.1 was up-regulated as part of a general cellular stress response, rather than a specific response to copper contamination. Transcripts coding for a homologue of alcohol dehydrogenase were also up-regulated at Cullen Bay. This short-chain dehydrogenase may play a detoxification role against reactive compounds. Oliveira et al. [61] found that alcohol dehydrogenase was up-regulated in nematodes exposed to arsenic. Another transcript up-regulated at Cullen Bay was similar to cytochrome-c oxidase subunit VIb (COI VIb-like), which may be directly related to copper as COI is a copper-dependant enzyme and may help store excess intracellular copper in the mitochondria [41].

We also detected genes and proteins that were down-regulated in Cullen Bay *Ophelina* sp.1 specimens, including the digestive proteins ATP synthase and a peritrophic matrix protein, homologous of alanine aminopeptidase (AP), and the calcium-binding protein troponin-C. The down-regulation of digestive proteins may be a reflection of reduced food availability and, consequently, lower metabolic rates in the Cullen Bay population. Alternatively the higher copper levels may compromise certain energy pathways, as seen in abalone exposed to endocrine disrupting chemicals [55]. Our results for AP support the view that this enzyme may have potential as a general toxin indicator [62]. Another study found that it was inhibited by a Shiff base compound that contained copper [63]. The AP present in *Ophelina* sp.1 may be inhibited by xenobiotic compounds at Cullen Bay, however, AP is also

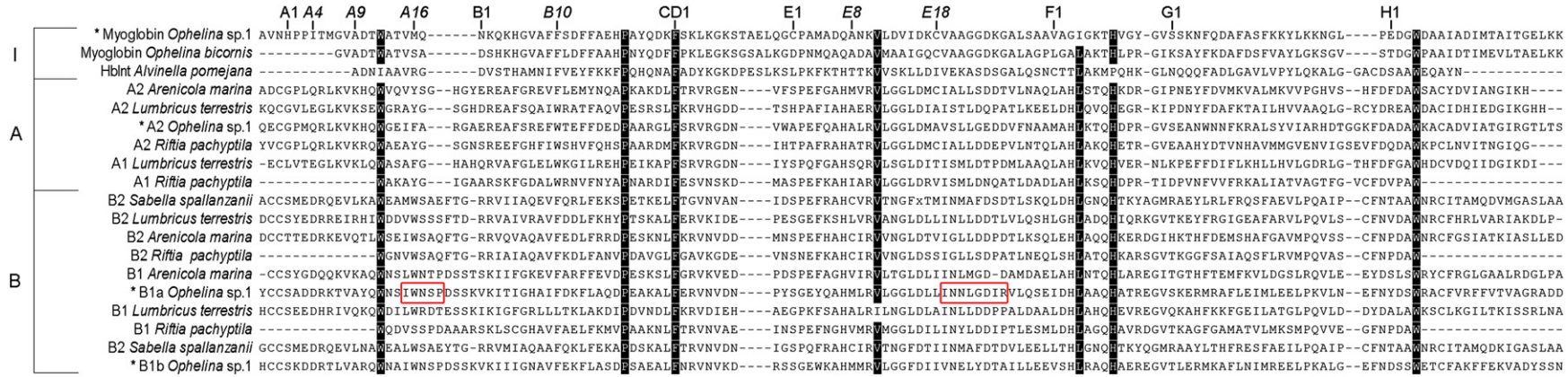

Fig. 7 – Amino acid sequence alignment of the translated transcripts of *Ophelina* sp.1 globins with other annelid intracellular globins (I) and annelid extracellular hemoglobin chains A and B (GenBank accessions are given in Table 1). Letters above the sequence denote the key residues based on the globin structure of *Lumbricus erythrocruorin* [40]. * denotes the new *Ophelina* sp.1 sequences generated in this study and the boxes indicate the amino acid sequence derived from MALDI-TOF/TOF sequencing of peptides from protein #15 (Table 6).

involved in metabolic pathways and its down-regulation may be again related to lower metabolic rates at Cullen Bay. Troponin-C, may have been down-regulated at Cullen Bay due to competition for its calcium binding sites by exogenous ions, a phenomenon that has been reported in fish exposed to cadmium [60].

Many of the transcripts and proteins that were differentially expressed belonged to cytoskeletal families, including: actin, myosin, paramyosin, intermediate filament proteins and troponin-C. These cytoskeletal proteins often have altered expression levels in organisms under metal stress, and it is likely that they are involved in protecting cells using an unknown mechanism. Miura [64] suggested that cytoskeletal proteins may be the target of oxidative stress, so the reorganisation of the cytoskeleton may, therefore, be involved in protecting the cells against damage induced by oxidative stress [55,65]. Wang et al. [38] suggested that while cytoskeletal proteins may be legitimately involved in stress responses, the nature of these proteins makes them more likely to be identified by DIGE and mass spectrometry techniques, introducing bias. When DIGE and mass spectrometry techniques are employed, the identified proteins are likely to be hydrophilic, medium sized (10–100 kDa) and abundant [38], precisely what we can expect from most cytoskeletal proteins. The many cytoskeletal proteins identified in the *Ophelina* sp.1 proteome may be an artefact of the techniques used. On the other hand, many differentially expressed transcripts were also identified as coding for cytoskeletal proteins, even though RNA sequencing is not generally biassed towards any gene family [13]. The proportion of genes in each of the functional KOG groups, the largest of which was cytoskeletal, further emphasised this. The fact that cytoskeletal proteins were over represented in this study is perhaps not surprising given the muscular physiology of polychaetes.

The up- or down- regulation of a gene does not always translate into an up- or down- regulated protein [16], nevertheless, we found strong correlations between the gene and protein regulations, especially for the up-regulated proteins. However, many of the genes with significantly altered expressions were not identified in the proteome, probably because far fewer proteins could be studied and identified using current proteomic methods when compared to RNA sequencing.

The number of protein spots on the DIGE gels reported in this study (576) was similar to the number of polychaete proteins in other DIGE studies [66]. From these proteins, 46 were excised for mass spectrometric analysis and 24 were confidently assigned an identity. The identification procedure used in this study was unusual in that it relied on generation of baseline transcriptome data. The amino acid sequences from the *Ophelina* sp.1 proteins were first matched to the *Ophelina* sp.1 transcripts and, if the amino acid sequences matched a single mRNA transcript sequence with 100% similarity, the entire translated mRNA transcript was then queried against protein databases. This approach resulted in far more protein identifications than would have been possible if the amino acid sequences were directly queried against the databases. Using new sequencing technologies to generate genetic baseline data for non-model organisms in proteomic studies is likely to be the future for proteomics, especially with the decreasing cost of next generation sequencing platforms.

In general, the predicted molecular weights (MW) and isoelectric points (pI) for the identified proteins were in approximate agreement with the observed MW and pI, although there were some notable exceptions. The observed MW for many of the cytoskeletal proteins, such as myosin, paramyosin, intermediate filament proteins and actins, were lower than the predicted value. This discrepancy is likely the result of degradation or truncation of these large muscular proteins. The observed pI for most proteins was within 1 pH unit of the predicted pI, however, this was not the case for the Lim-9-like proteins. This may be because these proteins were identified based on matches to conserved regions that may be present in many different proteins and, therefore, the predicted values may not apply to these specific polychaete proteins. Alternatively, post translational modifications may have altered their pI.

The 454-sequencing method recovered approximately 6000 *Ophelina* sp.1 contigs from each site, of which around 40% were significantly matched to an NCBI nr protein and over 60% were matched to a *Capitella teleta* gene model. These significant matches are high considering the simple and rapid mRNA extraction method used [29]. We expected that this gel-extraction method would co-purify more contaminating ribosomal RNA than other mRNA enrichment methods and result in fewer contigs with significant database matches. However, the number of significant matches in this study is comparable to other RNA sequencing studies on non-model organisms [14,67,68].

This study has generated valuable genetic and proteomic data for future toxicological polychaete experiments, in particular for toxicological studies in the Australian tropics. The opheliid polychaete, *Ophelina* sp.1, was a useful organism to test for the effects of copper and, coupled with the data generated in this study, may be valuable in future toxicological studies in the tropics. Candidates for biomarker development include: Atox1, hemoglobin chain B1a, β-thymosin, sulfotransferase 1C-like, alcohol dehydrogenase, cytochrome-c oxidase subunit VIb-like and alanine aminopeptidase.

Acknowledgements

We thank the Charles Darwin University for funding this project. We thank the Cullen Bay Management Corporation for allowing us to sample in Cullen Bay Marina.

REFERENCES

[1] Glasby CJ. Polychaete distribution patterns revisited: an historical explanation. Mar Ecol 2005;26:235–45.
[2] Bonnard M, Roméo M, Amiard-Triquet C. Effects of copper on the burrowing behaviour of estuarine and coastal invertebrates, the polychaete *Nereis diversicolor* and the bivalve *Scrobicularia plana*. Hum Ecol Risk Assess 2009;15:1549–7860.
[3] Reish DJ, Gerlinger TV. A review of the toxicological studies with polychaetous annelids. Bull Mar Sci 1997;60:584–607.
[4] Alain K, Olagnon M, Desbruyères D, Pagé A, Barbier G, Juniper SK, et al. Phylogenetic characterization of the bacterial assemblage associated with mucous secretions of the hydrothermal vent polychaete *Paralvinella palmiformis*. FEMS Microbiol Ecol 2002;42:463–76.

[5] Amiard-Triquet J-C, 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;76:160–202.
[6] Desbruyères D, Chevaldonné P, Alayse AM, Jollivet D, Lallier FH, Jouin-Toulmond C, et al. Biology and ecology of the "Pompeii worm" (*Alvinella pompejana* Desbruyères and Laubier), a normal dweller of an extreme deep-sea environment: a synthesis of current knowledge and recent developments. Deep Sea Res (II Top Stud Oceanogr) 1998;45:383–422.
[7] Mouneyrac C, Mastain O, Amiard JC, Amiard-Triquet C, Beaunier P, Jeantet A-Y, et al. Trace-metal detoxification and tolerance of the estuarine worm *Hediste diversicolor* chronically exposed in their environment. Mar Biol 2003;143:731–44.
[8] Rhee JS, Lee YM, Hwang DS, Won EJ, Raisuddin S, Shin KH, et al. Molecular cloning, expression, biochemical characteristics, and biomarker potential of theta class glutathione S-transferase (GST-T) from the polychaete *Neanthes succinea*. Aquat Toxicol 2007;83:104–15.
[9] Rhee JS, Won EJ, Kim RO, Lee J, Shin KH, Lee JS. Expression of superoxide dismutase (*SOD*) genes from the copper-exposed polychaete, *Neanthes succinea*. Mar Pollut Bull 2011;63:277–86.
[10] Sandrini JZ, Laurino J, Hatanaka T, Monserrat JM, Marins LF. cDNA cloning and expression analysis of the catalytic subunit of glutamate cysteine ligase gene in an annelid polychaete after cadmium exposure: a potential tool for pollution biomonitoring. Comp Biochem Physiol C Toxicol Pharmacol 2006;143:410–5.
[11] Won EJ, Kim RO, Rhee JS, Park GS, Lee J, Shin KH, et al. Response of glutathione S-transferase (*GST*) genes to cadmium exposure in the marine pollution indicator worm, *Perinereis nuntia*. Comp Biochem Physiol C Toxicol Pharmacol 2011;154:82–92.
[12] Collins LJ, Biggs PJ, Voelckel C, Joly S. An approach to transcriptome analysis of non-model organisms using short-read sequences. Genome Inform 2008;21:3–14.
[13] Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 2009;10:57–63.
[14] Vera JC, Wheat CW, Fescemyer HW, Frilander MJ, Crawford DL, Hanski I, et al. Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. Mol Ecol 2008;17:1636–47.
[15] Ghazalpour A, Bennett B, Petyuk VA, Orozco L, Hagopian R, Mungrue IN, et al. Comparative analysis of proteome and transcriptome variation in mouse. PLoS Genet 2011;7: e1001393.
[16] Hegde PS, White IR, Debouck C. Interplay of transcriptomics and proteomics. Curr Opin Biotechnol 2003;14:647–51.
[17] García-Alonso J, Greenway GM, Munshi A, Gómez JC, Mazik K, Knight AW, et al. Biological responses to contaminants in the Humber Estuary: disentangling complex relationships. Mar Environ Res 2011;71:295–303.
[18] Geracitano LA, Bocchetti R, Monserrat JM, Regoli F, Bianchini A. Oxidative stress responses in two populations of *Laeonereis acuta* (Polychaeta, Nereididae) after acute and chronic exposure to copper. Mar Environ Res 2004;58:1–17.
[19] Won EJ, Raisuddin S, Shin KH. Evaluation of induction of metallothionein-like proteins (MTLPs) in the polychaetes for biomonitoring of heavy metal pollution in marine sediments. Mar Pollut Bull 2008;57:544–51.
[20] Marouga R, David S, Hawkins E. The development of the DIGE system: 2D fluorescence difference gel analysis technology. Anal Bioanal Chem 2005;382:669–78.
[21] Dowling VA, Sheehan D. Proteomics as a route to identification of toxicity targets in environmental toxicology. Proteomics 2006;6:5597–604.
[22] Ferguson R. The effectiveness of Australia's response to the black striped mussel incursion in Darwin, Australia. A report of the marine pest incursion management workshop; August 27–28 1999.
[23] Munksgaard NC, Parry DL. Metal levels in Darwin marina sediments 2000–2006. Department of Primary Industry, Fisheries and Mines, Northern Territory Government; 2006.
[24] Munksgaard NC, Parry DL. Metals, arsenic and lead isotopes in near-pristine estuarine and marine coastal sediments from northern Australia. Mar Freshw Res 2002;53:719–29.
[25] Russell BC, Hewitt CL. Baseline survey of the port of Darwin for introduced marine species. A report to the Northern Territory Department of Transport and Works. Northern Territory Government, Department of Transport and Works; 2000. p. 98.
[26] Seike K. Burrowing behaviour inferred from feeding traces of the opheliid polychaete *Euzonus* sp. as response to beach morphodynamics. Mar Biol 1998;153:1199–206.
[27] Fauchald K, Jumars PA. The diet of worms: a study of polychaete feeding guilds. Oceanogr Mar Biol Annu Rev 1979;17:193–284.
[28] Silva MAL, Rezende CE. Behaviour of selected micro and trace elements and organic matter in sediments of a freshwater system in south-east Brazil. Sci Total Environ 2002;292:121–8.
[29] McGrath KC, Thomas-Hall SR, Cheng CT, Leo L, Alexa A, Schmidt S, et al. Isolation and analysis of mRNA from environmental microbial communities. J Microbiol Methods 2008;75:172–6.
[30] Rice P, Longden I, Bleasby A. EMBOSS: the European molecular biology open software suite. Trends Genet 2000;16:276–7.
[31] Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, et al. The COG database: an updated version includes eukaryotes. BMC Bioinformatics 2003;4:41.
[32] Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol 2010;11:R106.
[33] Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol 2010:R25.
[34] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using likelihood, distance, and parsimony methods. Mol Biol Evol 2011;28:2731–9.
[35] Abramoff MD, Magalhaes PJ, Ram SJ. Image processing with ImageJ. Biophotonics Int 2004;11:36–42.
[36] ANZECC. Australian and New Zealand guidelines for fresh and marine water quality, vol. 1. Canberra, Australia: Australian and New Zealand Environment and Conservation Council of Australia and New Zealand; 2000.
[37] Rosenzweig AC, Huffman DL, Hou MY, Wernimont AK, Pufahl RA, O'Halloran TV. Crystal structure of the Atx1 metallochaperone protein at 1.02 Å resolution. Structure 1999;7:605–17.
[38] Wang P, Bouwman FG, Mariman ECM. Generally detected proteins in comparative proteomics—a matter of cellular stress response? Proteomics 2009;9:2955–66.
[39] Chabasse C, Bailly X, Rousselot M, Zal F. The multigenic family of the extracellular hemoglobin from the annelid polychaete *Arenicola marina*. Comp Biochem Physiol B Biochem Mol Biol 2006;144:319–25.
[40] Strand K, Knapp JE, Bhyravbhatla V, Royer Jr WE. Crystal structure of the hemoglobin dodecamer from *Lumbricus* Erythrocruorin: allosteric core of giant annelid respiratory complexes. J Mol Biol 2004;344:119–34.
[41] Kim B-E, Nevitt T, Thiele DJ. Mechanisms for copper acquisition, distribution and regulation. Nat Chem Biol 2008;4:176–85.
[42] Nishino Y, Ando M, Makino R, Ueda K, Okamoto Y, Kojima N. Different mechanisms between copper and iron in catecholamines-mediated oxidative DNA damage and disruption of gene expression in vitro. Neurotox Res 2010;20: 84–92.
[43] Macomber L, Imlay JA. The iron–sulfur clusters of dehydratases are primary intracellular targets of copper toxicity. Proc Natl Acad Sci U S A 2008;106:8344–9.

[44] Kelner GS, Lee M-H, Clark ME, Maciejewski D, McGrath D, Rabizadeh S, et al. The copper transport protein Atox1 promotes neuronal survival. J Biol Chem 2000;275:580–4.
[45] Jeney V, Itoh S, Wendt M, Gradek Q, Ushio-Fukai M, Harrison DG, et al. Role of antioxidant-1 in extracellular superoxide dismutase function and expression. Circ Res 2005;96:723–9.
[46] Portnoy ME, Rosenzweig AC, Rae T, Huffman DL, O'Halloran TV, Culotta VC. Structure–function analyses of the ATX1 metallochaperone. J Biol Chem 1999;274:15041–5.
[47] Wakabayashi T, Nakamura N, Sambongi Y, Wada Y, Oka T, Futai M. Identification of the copper chaperone, CUC-1, in *Caenorhabditis elegans*: tissue specific co-expression with the copper transporting ATPase, CUA-1. FEBS Lett 1998;441:141–6.
[48] Bailly X, Leroy R, Carney S, Collin O, Zal F, Toulmond A, et al. The loss of the hemoglobin H_2S-binding function in annelids from sulfide-free habitats reveals molecular adaptation driven by Darwinian positive selection. Proc Natl Acad Sci U S A 2003;100:5885–90.
[49] Gutteridge JMC. Iron promoters of the Fenton reaction and lipid peroxidation can be released from haemoglobin by peroxides. FEBS Lett 1986;201:291–5.
[50] Kim J, Park Y, Choi K. Phototoxicity and oxidative stress responses in *Daphnia magna* under exposure to sulfathiazole and environmental level ultraviolet B irradiation. Aquat Toxicol 2009;91:87–94.
[51] Fedeli D, Carloni M, Falcioni G. Oxidative damage in trout erythrocyte in response to "in vitro" copper exposure. Mar Environ Res 2010;69:172–7.
[52] Folk DS, Franz KJ. A prochelator activated by β-secretase inhibits Aβ aggregation and suppresses copper-induced reactive oxygen species formation. J Am Chem Soc 2010;132:4994–5.
[53] Caldwell GS, Lewis C, Pickavance G, Taylor RL, Bentley MG. Exposure to copper and a cytotoxic polyunsaturated aldehyde induces reproductive failure in the marine polychaete *Nereis virens* (Sars). Aquat Toxicol 2011;104:126–34.
[54] Huff T, Müller CSG, Otto AM, Netzker R, Hannappel E. β-thymosins, small acidic peptides with multiple functions. Int J Biochem Cell Biol 2001;33:205–20.
[55] Zhou J, Cai Z-H, Li L, Gao Y-F, Hutchinson TH. A proteomics based approach to assessing the toxicity of bisphenol A and diallyl phthalate to the abalone (*Haliotis diversicolor* supertexta). Chemosphere 2010;79:595–604.
[56] Johannessen M, Møller S, Hansen T, Moens U, van Ghulue M. The multifunctional roles of the four-and-a-half-LIM only protein FHL2. Cell Mol Life Sci 2006;63:268–84.
[57] Mori M, Benotmane MA, Vanhove D, van Hummelen P, Hooghe-Peters EL, Desaintes C. Effect of ionizing radiation on gene expression in CD4+ T lymphocytes and in Jurkat cells: unravelling novel pathways in radiation response. Cell Mol Life Sci 2004;61:1955–64.
[58] Hattori K, Inoue M, Inoue T, Arai H, Tamura H-O. A novel sulfotransferase abundantly expressed in the dauer larvae of *Caenorhabditis elegans*. J Biochem 2006;139:355–62.
[59] Hehonah N, Zhu X, Brix L, Bolton-Grob R, Barnett A, Windmill K, et al. Molecular cloning, expression, localisation and functional characterisation of a rabbit SULT1C2 sulfotransferase. Int J Biochem Cell Biol 1999;31:869–82.
[60] Ling X-P, Zhu J-Y, Huang L, Huang H-Q. Proteomic changes in response to acute cadmium toxicity in gill tissue of *Paralichthys olivaceus*. Environ Toxicol Pharmacol 2009;27:212–8.
[61] Oliveira RP, Abate JP, Dilks K, Landis J, Ashraf J, Murphy CT, et al. Condition-adapted stress and longevity gene regulation by *Caenorhabditis elegans* SKN-1/Nrf. Aging Cell 2009;8:524–41.
[62] Chen S-H, Cao M-J, Huang J-Z, Wu G-P. Identification of a puromycin-sensitive aminopeptidase from zebrafish (*Danio rerio*). Comp Biochem Physiol B Biochem Mol Biol 2011;159:10–7.
[63] Hueso-Ureña F, Illán-Cabeza NA, Moreno-Carretero MN, Martínez-Martos JM, Ramírez-Expósito MJ. Synthesis and spectroscopic studies on the new Shiff base derived from the 1:2 condensation of 2,6-diformyl-4-methylphenol with 5-aminouracil (BDF5AU) and its transition metal complexes influence on biologically active peptides-regulating aminopeptidases. J Inorg Biochem 2003;94:326–34.
[64] Miura Y, Kano M, Abe K, Urano S, Suzuki S, Toda T. Age-dependent variations of cell response to oxidative stress: proteomic approach to protein expression and phosphorylation. Electrophoresis 2005;26:2786–96.
[65] Apraiz I, Mi J, Cristobal S. Identification of proteomic signatures of exposure to marine pollutants in mussels (*Mytilus edulis*). Mol Cell Proteomics 2006;5:1274–85.
[66] Mok FSY, Thiyagarajan V, Qian P-Y. Proteomic analysis during larval development and metamorphosis of the spionid polychaete *Pseudopolydora vexillosa*. Proteome Sci 2009;7:44.
[67] Frias-Lopez J, Shi Y, Tyson GW, Coleman ML, Schuster SC, Chisholm SW, et al. Microbial community gene expression in ocean surface waters. Proc Natl Acad Sci U S A 2008;105:3805–10.
[68] Meyer E, Aglyamova GV, Wang S, Buchanan-Carter J, Abrego D, Colbourne JK, et al. Sequencing and de novo analysis of a coral larval transcriptome using 454 GSFLx. BMC Genomics 2009;10:219.