The Regulation of Copper Stress Response Genes in the Polychaete *Nereis diversicolor* during prolonged Extreme Copper Contamination

Jonathan S. McQuillan[*,†], *Peter Kille*[‡], *Kate Powell*[‡] *and Tamara S. Galloway*[†].

[†]Biosciences, College of Life and Environmental Sciences, University of Exeter, Exeter, EX4 4QD, UK.

[‡]Cardiff School of Biosciences, Cardiff University, Park Place, Cardiff, CF10 3US, UK;

*Corresponding author: email. j.mcquillan@soton.ac.uk; tel. +44(0)2380591688

KEYWORDS: Polychaete, Copper, Transcriptomics, Nereis diversicolor, Pollution.

Abstract

Polychaetes are frequented in toxicological studies, one reason being that some members occupy shallow burrows in sediments and are maximally exposed to the contaminants that accumulate within them. We have been studying one population of the polychaete *Nereis* (*Hediste*) *diversicolor* exhibiting inheritable tolerance to extreme copper contamination in estuarine sediment. Using transcriptome sequencing data we have identified a suite of genes with putative roles in metal detoxification and tolerance, and measured their regulation. Copper tolerant individuals display significantly different gene expression profiles compared to animals from a nearby population living without remarkable copper levels. Gene transcripts encoding principle copper homeostasis proteins including membrane copper ion transporters, copper ion chaperones and putative metallothionein-like proteins were significantly more abundant in tolerant animals occupying contaminated sediment. In contrast, those encoding antioxidants and cellular repair pathways were unchanged. Non-tolerant animals living in contaminated sediment showed no difference in copper homeostasis-related gene expression but did have significantly elevated levels of mRNAs encoding Glutathione Peroxidase enzymes. This study represents the first use of functional genomics to investigate the copper tolerance trait in this species and provides insight into the mechanism used by these individuals to survive and flourish in conditions which are lethal to their conspecifics.

Introduction

It is widely accepted that the release of anthropogenic waste into natural waters can, through natural selection, effect fundamental changes on species whose habitat preferences leave them vulnerable to exposure. Waste metals released from mining and quarrying frequently enter aquatic systems through drainage and run-off, and are deposited into the sediment of rivers beds and estuaries during transit to the sea [1]. Species that occupy these sediments are, therefore, highly exposed and in many cases develop an evolutionary tolerance [2]. These ecological traits offer a unique opportunity to study the methods used by organisms to adapt to extreme conditions of metal excess when normal metal homeostasis is overwhelmed.

In some cases the magnitude of contamination events can instigate changes to a population on an evolutionarily short time-scale (e.g. a few hundred years). In this study we refer specifically to the infaunal Polychaete *Nereis (Hediste) diversicolor* (Harbour Ragworm). The life history traits of this species may enhance its ability to quickly adapt to inhospitable conditions, for example, its occupation of the intertidal zone make it naturally tolerant of fluctuating environmental conditions, and its distribution within relatively isolated populations with restricted gene flow may prolong the existence of specific adaptations even after the selective forces have been removed [3].

One population, found within Restronguet Creek, a branch of the Fal Estuary, Cornwall, UK occupies sediment that contains extraordinary concentrations of copper and to a lesser extent arsenic, tin and zinc [4-6]; a product of the local mining heritage dating back to the 1700's. Copper (Cu), at 5073 parts per million (ppm) [4], is an order of magnitude more concentrated than in other measured estuarine sediments in the south west of England and other contaminated areas world-wide [7, 8]. Accordingly, the sediment is demonstrably toxic to macro-fauna [4, 9], with the exception

58 that a population of *N. diversicolor* has flourished [4]. The individuals are tolerant to the acute
59 toxic effects of copper, zinc and cadmium[4, 10-12], but generally have a smaller size and a reduced
60 fecundity as a trade off towards the metabolic cost of tolerance [13]. The trait is not readily lost or
61 gained in the laboratory [4, 12, 14] and is heritable in the absence of metal [11] indicating an underlying
62 genetic adaptation.

63 Hypotheses for the molecular basis of the Cu tolerance include (1) the recruitment and
64 enhancement of existing Cu handling pathways, leading to an increased rate of detoxification; (2)
65 the enhancement of cellular repair pathways to repair copper-induced cell damage; and (3)
66 physiological traits including an increase in mucus production by tolerant animals to reduce
67 exposure [12]. It has been shown that the tolerant individuals take up ambient Cu [15, 16] and produce
68 insoluble detoxificatory deposits within their tegument [12, 17, 18], suggesting enhanced
69 detoxification. Specimens collected from Restronguet Creek have up to 91% of the total body Cu
70 concentration present in the insoluble fraction, compared with up to 30% in organisms collected
71 from sites with unremarkable Cu levels [19]. Pook [20] found a significant increase in the reduced
72 glutathione (GSH) pool supporting a model in which copper is associated with GSH after
73 entering the cell and Cu-thiol complexes are then metabolised in lysosomes, leading to the
74 formation of the detoxificatory deposits. It is unclear whether this process might be sufficient to
75 prevent copper toxicity or if additional adaptation to copper-induced cell damage might also be
76 important; for example, the role of antioxidant systems and cellular repair pathways is unclear.
77 Furthermore, the mucus secreted by *N. diversicolor* only adsorbs small amounts of dissolved
78 metal ions and may not represent a major defence mechanism [15].

79 Here we describe the first functional genomics-based study of the mechanism enabling *N.*
80 *diversicolor* to survive under perpetually Cu stressed conditions. Contextual data, including

metal concentration and copper tolerance, was collected by analysing samples from Restronguet creek in tandem with those collected at nearby reference sites. We report the first transcriptome sequence data for *N. diversicolor*, and the identification of a suite of gene transcripts orthologous to those encoding Cu homeostasis and general stress response proteins including antioxidants and DNA repair enzymes. Transcriptional regulation was measured using real-time PCR, enabling for the first time the regulation of Cu homeostasis to be described in Restronguet Creek.

Materials and Methods

Animal Collection and Husbandry

N. diversicolor were collected by hand in April / May 2012 at 3 locations within the Fal estuary system; Restronguet Creek (+50° 12' 32.97", -5° 5' 21.13"), Mylor Bridge, which lies approximately 3.2 Km to the south (+50° 11' 1.05", -5° 4' 31.15") and Cowlands Creek, which lies approximately 3.5 Km to the east (+50° 13' 39.53", -5° 2' 36.47"). Specimens with a wet weight of 400-600 mg were used for experiments. Animals were either (1) snap-frozen in liquid nitrogen on site or (2) maintained in glass aquaria containing clean sediment (Specialist Aggregates; this material contained only trace amounts of metal) and artificial seawater (ASW; Tropic Marin; 16 PSU) at 11 ^{o}C with a 12 hour light:dark photoperiod.

Transmission Electron Microscopy

The tegument was removed from anaesthetised animals (in 5 % ethanol / ASW) and fixed in 3 % glutaraldehyde (in 0.2 M phosphate buffer, pH 7.2) for 2 hours; stained with 1 % osmium tetroxide (aqueous) for 1 hour; dehydrated in alcohol; embedded in medium viscosity resin (TAAB Laboratories Equipment Ltd); and thin sections (80 nm) were collected on an ultra-microtome. Sections were pre-stained with uranyl acetate and Reynold's lead citrate (for conventional viewing) or collected on gold grids (unstained) for energy dispersive X-ray analysis. Stained sections were viewed with the JEOL 1400 TEM. Unstained sections for X-ray analysis were viewed on a Philips/FEI CM12 TEM operating at 80kV fitted with an X-ray microanalysis detector (EM-400 Detecting Unit, EDAX UK). Images were recorded using a SIS MegaView III digital camera. Analysis of metal containing granular deposits within the tissue was performed using ImageJ [21], using the sizing and densitometry functions.

Metal Measurements

Sediment cores were collected at each study site (2-5 cm depth), dehydrated in a vacuum oven and sieved with a 500 µm mesh. Then, 0.5 g was acid-digested in near-boiling Aqua Regia for 1 hour. A certified reference material (LGC 6137 Estuarine Sediment) was processed in parallel.

N. diversicolor (frozen) were ground into a fine powder and 0.5 g was acidified with 10 % nitric acid and microwave digested at 200 ^{o}C for 10 minutes (ETHOS EZ, Milestone). A certified reference material (NRC TORT-2 Lobster hepatopancreas) was processed in parallel.

The concentration of As, Cu, Sn and Zn in the digests was measured using a Varian 725 ES ICP-OES (Agilent). A five point calibration curve was constructed using an appropriate series of multi-element calibration standards and quality assurance was provided for by the analysis of a second multi-element standard prepared from a different stock solution and by the analysis of the certified reference materials.

Copper Toxicity Assay

Animals were placed individually into acid washed glass troughs (15 cm x 5 cm x 5 cm) containing 200 mL of ASW and a 10 cm length of 6 mm internal diameter polyvinylchloride tubing (refuge); 11^{o}C; and 12 hour photo-period. After 24 hours the water was replaced with a solution of $CuCl_2$, renewed every 24 hours. Animals were considered dead if their anterior segments didn't respond to a mechanical stimulus.

RNA Extraction

Animals were rinsed in 0.2 µm filtered ASW, snap-frozen and ground into a fine powder. Total RNA was extracted using Trizol (Ambion) and the RNeasy Mini Kit (Qiagen). Quality assurance was provided for by resolving the RNA on a non-denaturing agarose gel to visualise a

sharp ribosomal RNA band, and by measuring the absorbance of the sample using a Nanodrop spectrophotometer (Thermo Scientific). Residual genomic DNA was digested with RNase free DNase (RQ1, Promega) prior to clean-up using the RNeasy Mini Kit. Complete DNA digestion was confirmed by a null result from a Taq polymerase-based PCR amplification reaction using the digest sample as a template.

454 Sequencing

RNA was collected from 12 animals; 6 from Restronguet Creek (copper tolerant) and 6 from Cowlands Creek (copper sensitive). The samples were combined in equal proportions and a cDNA library was prepared from 2 µg using the MINT First Strand cDNA Synthesis Kit with oligodT primers, followed by cDNA amplification and Duplex Specific Nuclease-based library normalisation using the TRIMMER and EncycloPCR kits (Evrogen). Normalised cDNA was fragmented by nebulisation followed by adapter sequence ligation and emulsion PCR, and the cDNA library was sequenced on full GS FLX and GS FLX+ Titanium series sequencing plates according to the standard protocol of Roche (Eurofins Genetic Services Ltd).

The raw 454 data reads were assembled *de novo*. Adapter sequence trimming, quality filtering and the initial assembly of sequence contigs were carried out using GS De Novo Assembler (version 2.3, Roche). Additional assemblies were generated using Mira [22] and Cap3 [23]. For each unique contig open reading frames were filtered using DeconSeq [24] and annotated using the blastx programme conducted against the non-redundant protein sequence database (NCBI) with an E-value cut-off of 10^{-5}.

Gene Expression Analysis

Gene expression analysis was performed using real-time PCR (see below). For the *in situ* analysis we collected at least 20 undamaged individuals from each site and immediately snap-

froze them in liquid nitrogen to stabilize the mRNA pool. For each site the frozen specimens were combined and homogenized for RNA isolation. This was repeated on 3 different days, sampling at low tide. For analysis in the absence of metal we placed individuals from each site into clean sediment (as described above) and removed at least 20 individuals every 7 days. Replicate measurements were taken from 3 different holding tanks.

Real Time PCR

Complementary DNA was prepared by oligodT-primed reverse transcription using the Thermoscript RT-PCR System (Invitrogen). The relative abundance of gene transcripts between different individuals was determined by real-time PCR using the CFX96 Real-time PCR Detection System (Biorad) and SYBR green DNA detection chemistry. Transcript sequence-specific primers (designed using Primer 3 [25]) amplify 150-250 bp within the 5' region of the target sequence; see supporting information. All RT-PCR reactions were performed in triplicate, each reaction containing 12.5 μL of IQ SYBR supermix (Biorad), 7.5 μL of nuclease free water, 2 μL of forward and reverse primers (10 pmol/μL) and 1 μL of template cDNA; the total reaction volume was 25 μL.

The thermal cycling program was 95^{o}C for 5 minutes followed by 40 cycles of 95^{o}C for 15 seconds, 55-60^{o}C for 30 seconds and 72^{o}C for 1 minute. At the end of the programme product specificity was analysed from a dissociation curve (55^{o}C to 95^{o}C at 1^{o}C increments for 5 seconds each). Threshold cycle (C_t) values were determined within the exponential potion of the reaction curves. Relative transcript abundance was calculated by normalising the C_t values for the target to the C_t values for a panel of endogenous reference transcript sequences. A panel of 7 endogenous reference sequences were tested, and those that gave consistent results were selected for the analysis of the real-time PCR data; the data presented represents the mean value obtained

using each valid reference. The relative transcript abundance was calculated from E^n, where n was the difference between the C_t values obtained and E is the efficiency of the primer set used for the amplification reaction. Primer efficiency was derived as $E = 10^{\frac{-1}{-m}}$ where m is the slope of a linear fit for C_t values obtained from a 10-fold dilution series of cDNA template in Real Time PCR reactions.

Results and Discussion

Heavy metal contamination in the Fal Estuary System and Copper Tolerance in *Nereis diversicolor*

We sampled 3 estuarine locations in the south west of England, which contained a rich population of *Nereis diversicolor* (more than 1,000 burrows / m^2). These were Restronguet Creek (RC), where metal contamination is famously high [4], Mylor Bridge (MB) and Cowlands Creek (CC). The concentration of Arsenic (As), Copper (Cu), Tin (Sn) and Zinc (Zn) in sediment samples was determined by ICP-OES to be in the order Cu>Zn>As>Sn and RC>MB>CC, as shown Figure 1A. At the time of sampling RC sediment contained more Cu than in any previously measured estuarine location within south west England [6], the Rio Tinto in Peru [7] and the Aznalcollar Tailings in the USA [8]. We use the Kelly Indices (former GLC guidelines) for contaminated soil (IRCL 59/83) to describe the extent of Cu contamination at each site; RC being unusually heavily contaminated; MB being heavily contaminated; and CC being uncontaminated and therefore used as a reference site for the gene expression experiments. To ascertain the effect of the metal pollution on the ecological diversity of the sediment the benthic macro-fauna was measured according to the method of Simpson [26]. Biodiversity at RC was significantly reduced in line with previous determinations [4], specifically in the upper reaches where we only found *N. diversicolor* (data not shown).

In *N. diversicolor* tissue As, Cu and Sn were accumulated to the same order as the ambient concentration, however tissue Zn concentration showed no significant difference between the populations (t-test, $p = > 0.05$) despite clear differences in the sediment Zn concentration at each site, Figure 1B. Historical comparison with measurements made by Bryan and Gibbs in 1983 [4] (for sediment) and Berthet *et al* in 2003 [12] (for tissue), wherein the same

sampling locations were accessed, indicate that despite a modest increase in the sediment Cu concentration the tissue Cu concentration has decreased by more than an order of magnitude. Recent work by Rainbow *et al* [15, 16] to quantify the kinetics of Cu uptake by *N. diversicolor* indicates that Cu accumulation rate is proportional to the biologically available copper in their immediate environment, and therefore we conclude that the bioavailable portion of the Cu in RC may have decreased dramatically in the last few decades. Historical data for As, Sn and Zn was unavailable for comparison. Possible mechanisms for the decline in sediment Cu include tidal "flushing" together with reduced inputs after cessation of local mining operations, or changes to the sediment chemistry (e.g. the concentration of dissolved organic matter).

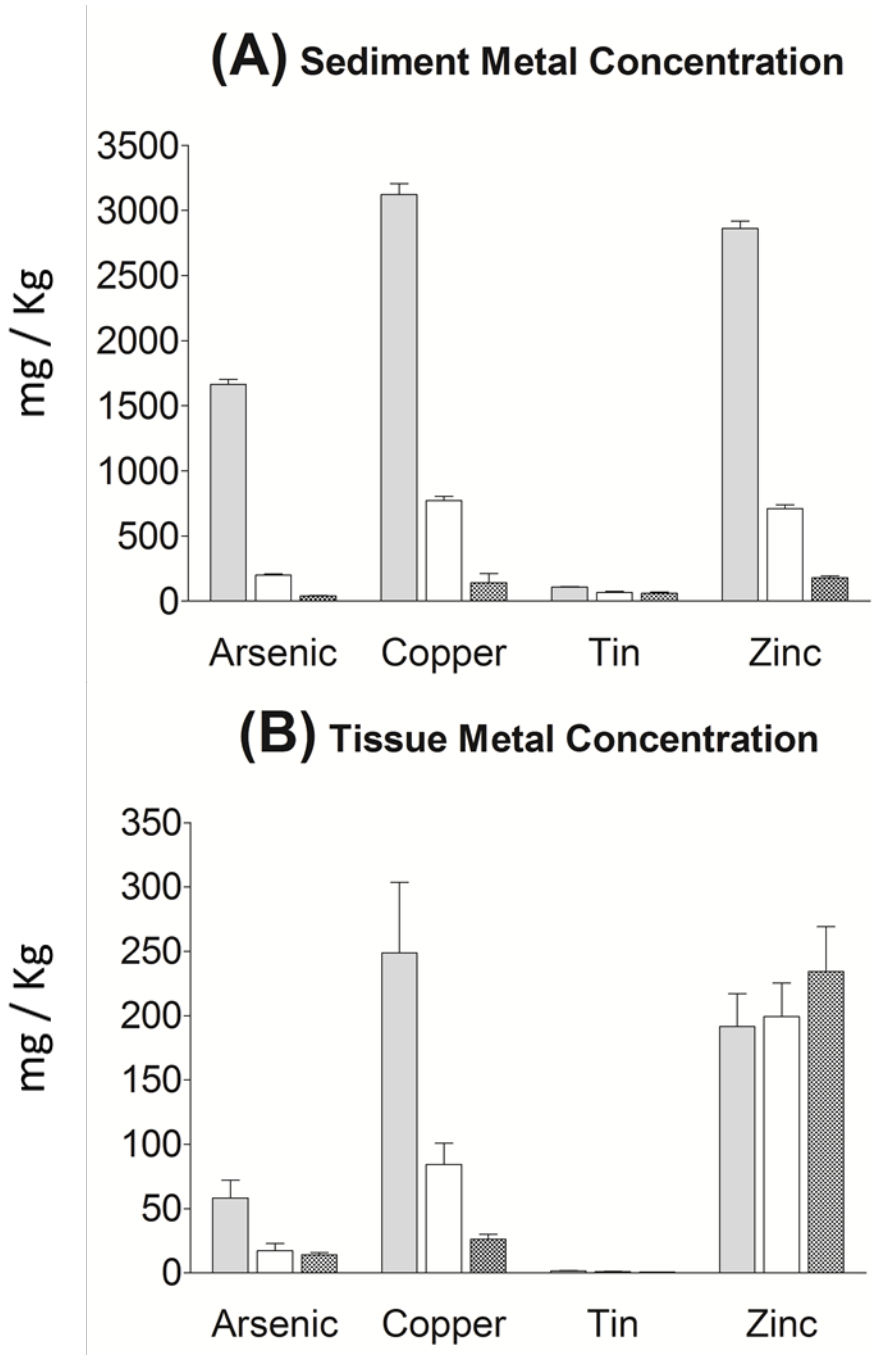

Figure 1. Results of the ICP-OES determination of metal concentration in (A) the sediment and (B) the tissue of *N. diversicolor* in Restronguet Creek (grey bars), Mylor Bridge (white bars) and Cowlands Creek (shaded bars). The error bars represent the standard error on the mean from ≥ 25 sediment samples or ≥ 10 individuals.

The fate of the bio-accumulated copper can be investigated using transmission electron microscopy and energy dispersive X-ray spectroscopy, which shows that copper co-localises with sulphur in granular deposits close to the tegument epiculticle [12, 18], as shown in Figure 2. In similar work carried out in 2003 Mourneyrac *et al* [12] found that Cu tolerant animals from RC presented with an increase in the number and density of the Cu-containing granules indicating a role as detoxificatory Cu stores and a potential role in Cu tolerance. In this study, however, image analysis showed no difference in the number of granules, their size or their density between the populations tested. This corresponds to our comparison of the tissue concentration measured in this study compared to another study carried out using the sane sampling location in 2003 [19].

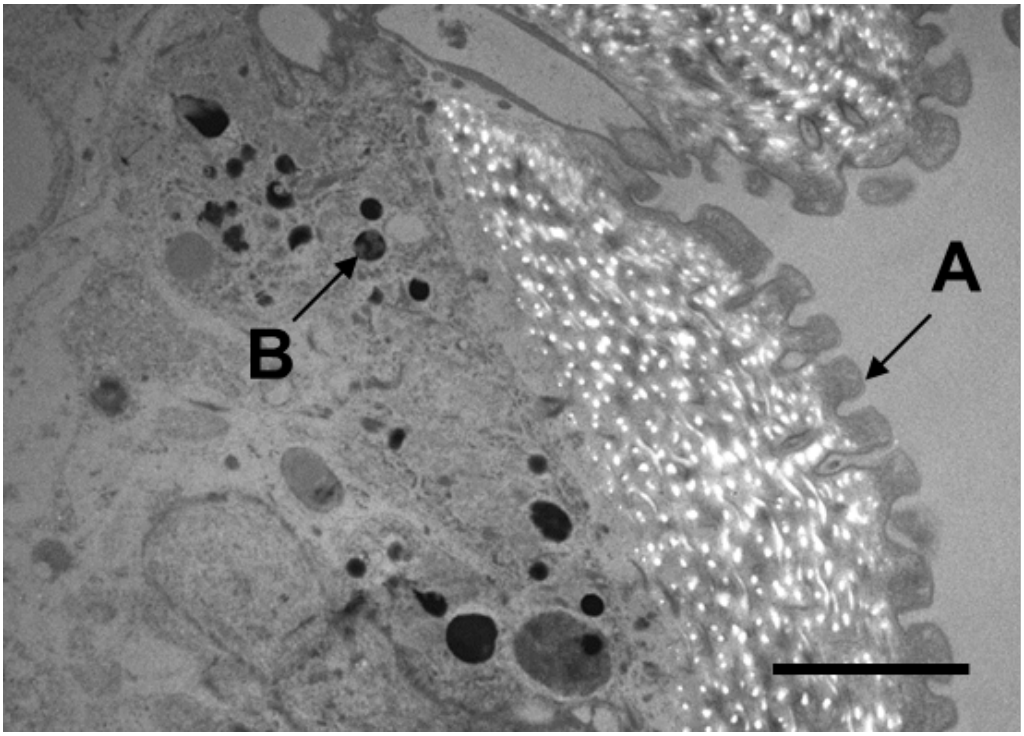

Figure 2. A Transmission Electron Micrograph of a thin sectioned *N. diversicolor* Tegument, excised from a specimen collected at Restronguet Creek, showing the Epicuticle Layer (A) and the underlying tissue with Cu / S containing granules (B). Sections were analysed unstained by energy dispersive X-ray spectroscopy to measure the elemental composition of the granular structures, which were predominantly copper and sulphur (data provided in the supplementary information). A typical image is shown. Bar = 2 µm.

The Cu contamination in RC has driven the selection of a Cu tolerance phenotype that was first reported over 30 years ago [4]. To confirm that this adaptation was still present we compared the Cu tolerance of animals from the differentially contaminated sediments by placing individuals into a solution of $CuCl_2$ and measuring their survival over several days. At a Cu

concentration of 2 µg/mL the individuals found at RC were significantly (two-way ANOVA $P \leq 0.05$) more tolerant to the $CuCl_2$ than those found at either MB or CC, as shown in Figure 3, despite the decrease in the tissue burden of Cu indicating that the selective pressure may have diminished. Crucially, the RC model for the study of rapid adaptation and tolerance was still available for gene expression studies.

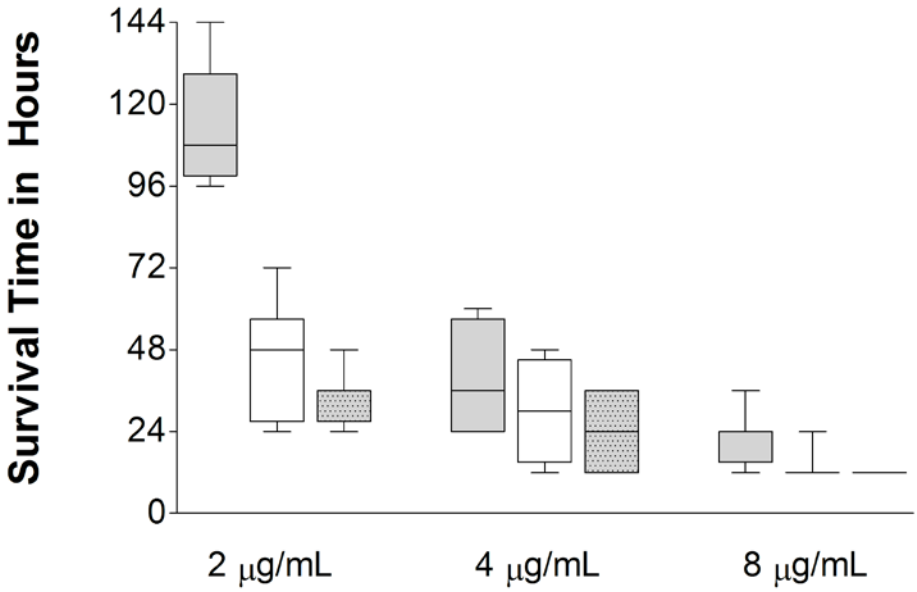

Figure 3. *N. diversicolor* from the different sites ($n \geq 12$) were placed individually into solutions of $CuCl_2$. The survival of individuals from Restronguet Creek (unusually heavy Cu contamination; grey bars), Mylor Bridge (heavy Cu contamination; white bars) and Cowlands Creek (uncontaminated; shaded bars) was measured every 12 hours. The data is presented as a Box-Whiskers plot.

Sequencing and Transcriptome Assembly

To enable a functional genomics-based investigation of the mechanisms underlying the Cu tolerance trait we generated transcriptome sequence datasets using the 454 pyrosequencing technology. A normalised cDNA library was prepared from a pool of the Cu tolerant and Cu sensitive *N. diversicolor* found in RC and CC respectively. The library was sequenced using two sequencing plates (GS FLX Titanium Series, Roche), generating 2,151,516 sequence reads with an average read length of 440 bp and 307 bp of which 1,923,224 (89.4 %) were used for *de novo* assembly of the *N. diversicolor* transcriptome. Read assembly was carried out using multiple assembly algorithms (Newbler, Mira and Cap3; [22, 23, 27]) to maximise the potential coverage [28]. All sequence data underwent post-assembly filtering using DeconSeq [24] and unique transcript sequences of greater than 300 bp in length were queried against the non-redundant protein database using blastx with an E-value cut-off of 10^{-5}. Manual annotation, where necessary, was carried out using local tblastn searches of the transcriptome assembly data with query sequences corresponding to target genes or proteins. Transcriptome coverage was assessed using Transrate [29] against a set of "eukaryotic conserved" proteins (CEGMA [30]). The assembly statistics, blast results for each assembly and transcriptome coverage are summarised in Table 1 (additional information is available in the supplementary material, Table S2). The annotated sequence assemblies were used to identify orthologues to copper homeostasis systems and associated detoxificatory and repair pathways including glutathione metabolism, antioxidants and DNA repair proteins.

Table 2. Transcriptome Assembly and Annotation Statistics.

Assembly Statistics			
Assembler	Mira	Newbler	Cap3
Unique Contigs (< 300 bp)	60,335	31,770	59,769
Largest (bp)	6801	6996	12,565
Number > 1 Kbp	11,873	14,720	18,343
Mean Contig Length	751	1,065	887
N_{50}	837	1,195	1,018
Number with ORF	16,151	15,442	20,690
Annotation Results			
Blast Database	uniprot_sprot	uniprot_sprot	uniprot_sprot
Blast Program	BlaxtX	BlaxtX	BlaxtX
E<10^{-5}	23964	15318	25426
E<10^{-10}	20496	13555	21954
E<10^{-30}	10339	8301	11897
Transcriptome Coverage			
RBH* (as % of Reference)	0.21 %	0.33 %	0.47 %
Reference Coverage	0.26 %	0.34 %	0.39 %
Number of Chimeras	0	0	0

*Reciprocal Blast Hits

Copper Genes

We identified transcript sequences with orthology to copper transporters CTR1 (solute carrier family 31 protein) and ATP7A (P-Type ATPase family protein), which function to move copper across the plasma membrane, into or out of the cell respectively [31]. The ATP7A protein has an additional role in directing intracellular copper into the trans-Golgi network for

incorporation into newly synthesised cuproproteins and may relocate to eliminate copper from the cell under conditions of copper excess [32]. The ATP7 protein in humans has 2 isoforms, ATP7A and ATP7B, but is represented by a single homolog in lower organisms including insects [33]; and in our polychaete sequences we found evidence for a single ATP7 gene, which had more similarity to the ATP7A isoform. Any free Cu^{+} in the cell is highly toxic, so Cu trafficking is mediated by specialised chaperone proteins [34]. We identified Nereid orthologues to the protein CCS (*C*opper *C*haperone to *S*uperoxide dismutase [35]), which delivers Cu^{+} cofactor ions to SOD1 (Superoxide dismutase), and ATOX1 (Antioxidant protein 1), which interacts with ATP7 [36] to support copper export or compartmentalisation.

Additional systems augment copper homeostasis. Cellular thiols including those presented on glutathione and metallothionein / metallothionein-like proteins (MTLPs) sequester excess Cu^{+} away from the metabolism and contribute to the formation of detoxificatory stores [37, 38]. In addition to the identification of the principle components of glutathione metabolism we found evidence for 2 MTLP genes, encoding a typical (< 200 amino acids) and atypically large predicted protein. To our knowledge MTLPs in this species have, until this data, been unconfirmed. Interestingly, the discovery of 2 putative MTLP genes encoding different sized proteins sheds light on the work by Poirier *et al* who found a bi-modal distribution of Cu within heat stable, size fractionated *N. diversicolor* cytosol [18]. Additional analysis of the large MTLP sequence reveals a 404 amino acid protein (confirmed proteomically; data not shown) containing 5 repeats each comprising 10 cysteine residues with the consensus GCGC-X_5-G-XX-CC-X-G-XX-C-X_{11}-G-XX-G-XX-C-XX-$CX_7CX_5CX_7CX_5GX_4K$; the total cysteine composition was 13 %. We also looked for the major antioxidant proteins, which neutralise the effect of Cu^{+}-catalysed formation of reactive oxygen species. In addition to Catalase, *N. diversicolor* has at

316 least 2 isoforms of superoxide dismutase which are identical, or nearly identical to those found in
317 other polychaete species including *Perinereis nuntia*. Finally, we looked for DNA repair
318 pathway proteins for which we found a complement of candidate sequences. The sequences
319 described in this study are summarised in Table 2.
320
321
322

Table 2. Transcript sequences used in this study.

Name[1]	Protein Description	%ID[2]	Organism[3]	Genbank	E[4]
Copper Transporters					
ATP7A	Wilsons disease protein; ATPase family	68	*Daphnia pulex*	EFX90405.1	10^{-39}
CTR1	High affinity copper uptake transporter; Solute carrier family	36	*Harpegnathos saltator*	EFN85594.1	$3x10^{-24}$
Copper Chaperones					
ATOX1	Copper transport protein	51	*Crassostrea ariakensis*	AEJ08756.1	
CCS	Copper chaperone for superoxide dismutase	51	*Salmo salar*	NP_001133786.1	
Metallothionein-like proteins					
-	Cd/Se Metallothionein	55	*Schistosoma mansoni*	XP_002575981.1	10^{-13}
-	Atypical Metallothionein-like protein	26	*Trichomonas vaginalis*	XP_001321197.1	$9x10^{-09}$
Antioxidants					
SOD1	Cu/Zn Superoxide Dismutase	100	*Perinereis nuntia*	ADM64420.1	$3x10^{-10}$
SOD2	Mn Superoxide Dismutase	80	*Perinereis nuntia*	ADM64421.1	10^{-99}
CAT	Catalase	78	*S. kowalevskii*	XP_002738841.1	10^{-99}
Glutathione metabolism					
GCL	Glutamate cysteine ligase	92	*Laeonereis acuta*	AAV48595.2	10^{-122}
GSS	Glutathione synthetase	40	*Tribolium castaneum*	XP_968070.1	$2x10^{-72}$
GPX1	Glutathione peroxidase	68	*Hyriopsis cumingii*	ACY72387.1	$4x10^{-41}$
GSTM1	Glutathione S-transferase (Mu)	56	*Reishia clavigera*	ACD13785.1	$3x10^{-35}$
GSTT1	Glutathione S-transferase (Theta)	84	*Neanthes succinea*	ABQ82132.1	10^{-111}
GSTO1	Glutathione S-transferase (Omega)	46	*Neanthes succinea*	ABR24228.1	$2x10^{-48}$
DNA Repair					
DDB1	Damage DNA binding protein 1	78	*Aedes aegypti*	XP_001655231.1	$5x10^{-46}$
XRCC3	DNA excision repair protein	39	*Bos taurus*	NP_001071585.1	10^{-11}
ABH6	DNA alkylation repair protein	58	*Danio rerio*	NP_001005390.1	10^{-67}
ERCC3	DNA repair helicase	74	*Ixodes scapularis*	XP_002399857.1	$7x10^{-99}$

323 [1]Name in Homo sapiens; [2]% Identity with top blast result; [3]Top blast result; [4]E-value.

Expression *in situ*

The expression of these principle copper homeostasis genes was measured *in situ* (snap-freezing the animals at site to stabilise the mRNA pool), by comparing the abundance of mRNA transcripts using real-time PCR. We compared representatives ($n \geq 20$) of the 2 Cu-exposed populations, at RC and MB, with the unexposed reference population at CC on 3 different days during 2011. The results are compiled into a single figure for clarity, as shown in Figure 4. Unsurprisingly, *N. diversicolor* found in RC had significantly increased (one way ANOVA, $p < 0.05$) levels of gene transcripts corresponding to the copper transporters ATP7 and CTR1 and copper chaperones ATOX1 and CCS. These genes have an established role in protecting the cell from excess copper and their up-regulation in RC probably reflects the elevated sediment copper as shown in Figure 1A. Up-regulation of CTR1, which encodes a copper uptake transporter, is consistent with previous data indicating that the animals living in RC have increased Cu uptake rates despite the Cu-contamination [15, 16]. Our current theory is that the enhanced Cu uptake can protect the organism from Cu damage by increasing the rate of elimination of Cu in their immediate surroundings as it moves through detoxificatory pathways including the formation of the Cu / S containing granules [12, 18]. This process makes *N. diversicolor* a candidate for the bioremediation of contaminated sediments, as noted by others [39].

Putative MTLP transcripts were also significantly more abundant in RC; specifically the atypically large MTLP sequence showed up-regulation by up to 16-fold compared to individuals from the reference site. Conversely, we found no significant difference between RC and the reference site in the amount of mRNA transcripts corresponding to the antioxidant proteins and the DNA repair enzymes.

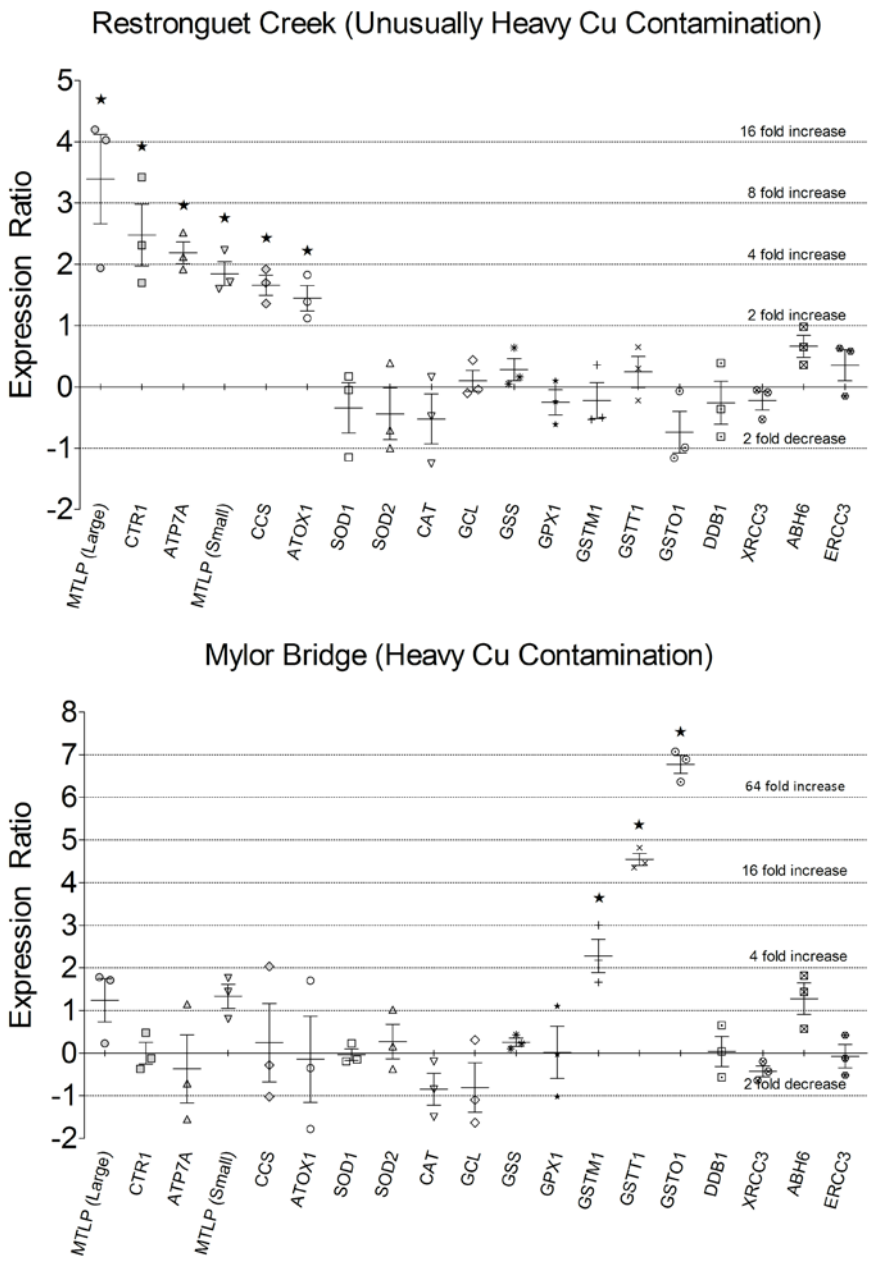

Figure 4. The *in situ* abundance of transcripts in representatives of the Cu contaminated populations in comparison to the uncontaminated reference population. The analysis was performed on a pool of at least 20 individuals on 3 different days. The horizontal line shows the mean expression ratio versus the reference, the error bars are the SEM from 3 replicate measurements (days) and the symbols show the values obtained for each replicate. Statistically significant differences (one way ANOVA, $p < 0.05$) are marked with a star.

It is curious that the gene expression data gave no evidence for the activation of these systems which have an established and significant role in protecting the cell from the effects of heavy metals [40]. Our previous biochemical assay data [20], comparing the activity of Superoxide dismutase and Catalase between animals in RC and the reference population, agreed with our genetic data that the exposed / tolerant individuals do not experience additional oxidative stress. In contrast, exposure of *Laeonereis acuta* (Nereididae), to Cu led to increased activities of Superoxide dismutase, Catalase, and Glutathione S-transferase [41], and enhanced antioxidant functions have been associated with copper exposure in a number of other annelid species (for example [42]). In this instance it is possible that the hyper activation of the Cu detoxificatory pathways, conferring the tolerance trait, is sufficient to mitigate any Cu-catalysed radical formation. It then follows that DNA repair pathways are similarly unaffected by the high ambient Cu levels.

The population in MB had also been exposed to a high concentration of Cu in the sediment but this was far lower than the extreme concentration measured in RC, and these individuals did not display a significant increase in their ability to tolerate Cu indicating that the selective pressure posed by this contamination is lower (compared to RC). We found no significant differences in the expression of the aforementioned copper homeostasis genes between individuals collected from MB and the reference site. It is, however, noted that the abundance of some gene transcripts was highly variable on the different days with, for example, the putative ATOX1 gene transcript fluctuating between approximately 4-fold up- and down-regulated compared to the reference population. In contrast to our data for RC, we found that the animals collected from MB displayed significant up-regulation of Glutathione S-transferase (GST) enzymes, by up to approximately 128-fold, however other antioxidants including

Superoxide dismutase and Catalase were unchanged compared to the reference. The GST enzymes have important roles in cellular detoxification and anti-lipid peroxidation processes and are frequented as biomarkers of environmental pollution. However, the response of GST expression to metals remains unclear; some studies reporting that they are induced by metal-contamination [43] and others that they are repressed [44].

Genetic Response to Cu depletion

We measured the stability over time of the observed differences in transcriptional regulation between the tolerant (RC) and non-tolerant (CC / reference) animals when in the absence of Cu. To achieve this, a second series of real-time PCR experiments were carried out wherein animals from each population were placed into tanks containing clean sediment and ASW for up to 3 weeks prior to RNA extraction. The relative abundance of mRNA transcripts matching the copper homeostasis genes (i.e. those showing significant regulation between the populations *in situ*) was measured every 7 days by sacrificing a random sample of the individuals within each tank. The observed *in situ* differences in transcript abundance were reduced after 7 days in the clean sediment. There was no significant difference in the abundance of the transcripts after 14 days, with the exception that the transcript encoding CTR1 remained significantly more abundant in the Cu-tolerant animals from RC over the duration of the experiment. The CTR1 gene encodes a Cu uptake membrane transporter and may supply Cu to sustain normal cellular processes in addition to detoxificatory pathways.

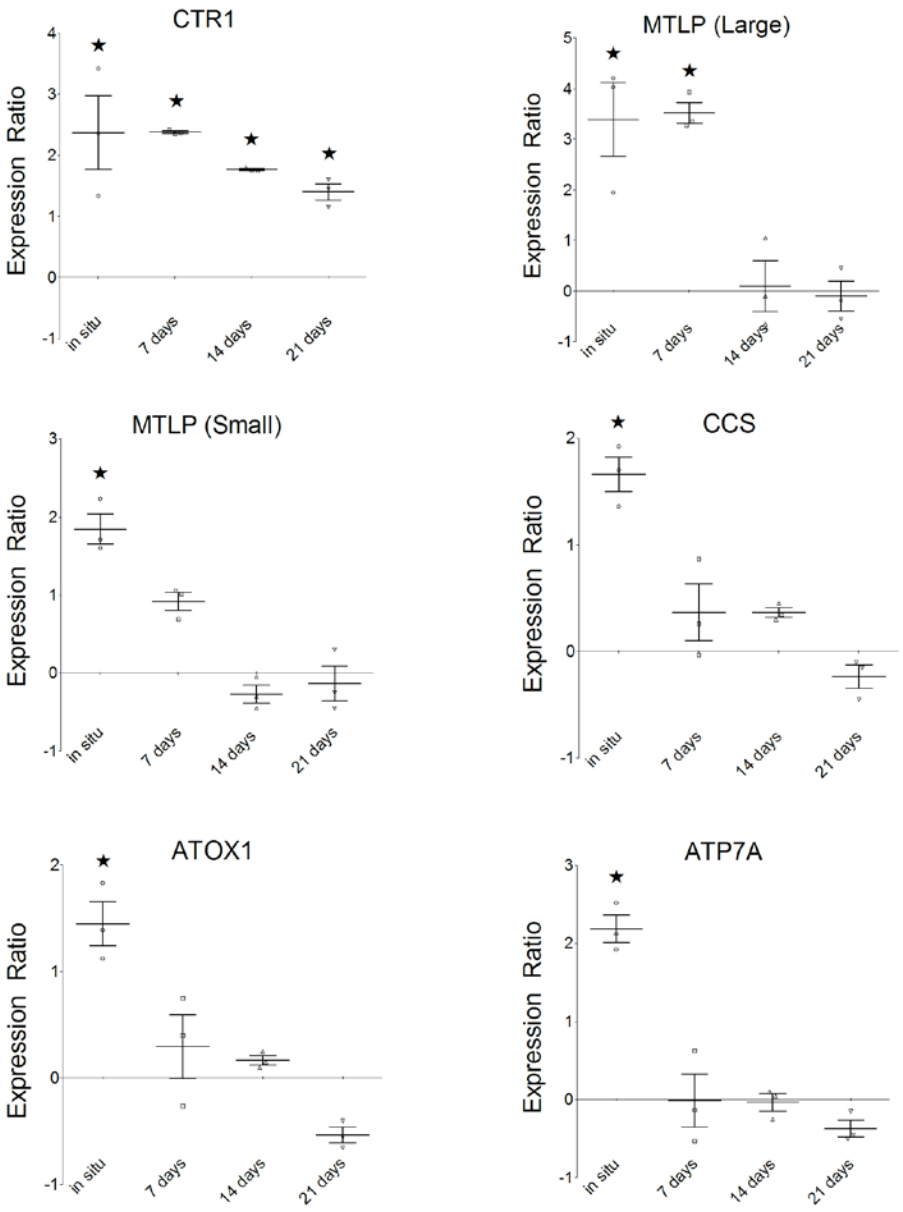

Figure 4. A comparison of the abundance of select transcripts in representatives of the Cu contaminated population from RC in comparison to the uncontaminated reference population. The analysis was performed on a pool of at least 20 individuals from 3 different holding tanks. The horizontal line shows the mean expression ratio versus the reference, the error bars are the SEM from 3 replicate measurements (tanks) and the symbols show the values obtained for each replicate. Statistically significant differences (one way ANOVA, $p < 0.05$) are marked with a star.

Interestingly, the copper toxicity assay, as described in Figure 3, was repeated using animals subjected to the 21 day experiment (see supplementary information Figure S2) wherein the animals from RC remained significantly more tolerant to $CuCl_2$. This supports the hypothesis that the mechanism for Cu tolerance in the RC population can persist in the absence of Cu, and is encoded by an underlying genetic adaptation.

The use of next generation sequencing in environmental science has expanded rapidly, facilitating functional genomics-based investigations using non-model (i.e. non-sequenced) organisms. This study was intended as a primer to facilitate molecular understanding of the processes that enable a species to rapidly adapt to pressure from anthropogenic contaminants such as waste metals. In summary, we compared 3 populations of *N. diversicolor*, either living with or without prolonged Cu stress, and either with or without ecological tolerance. The results indicate that whilst the tolerant individuals display up-regulated expression of Cu homeostasis genes, they do not make a measurable response to the product of Cu toxicity; namely oxygen radicals and cellular (DNA) damage. This supports the hypothesis that the Cu tolerance trait is facilitated by the activity of a detoxificatory pathway as opposed to some enhanced cellular repair. There is mounting evidence for the role of the detoxificatory granules in metal tolerance in invertebrates, and further work should be carried out to identify the genes and proteins that function in their formation. The non-tolerant but heavily exposed individuals in MB did not display a measurable change in the expression of Cu homeostasis genes compared to individuals from the reference population, but did produce significantly greater quantities of gene transcripts corresponding to antioxidants; specifically Glutathione S-transferase enzymes. Whilst this response cannot be attributed to the metal contamination alone, the absence of the response to

Cu, as observed in the tolerant population, is good evidence that the Cu homeostasis systems in this species are central to the molecular basis for the tolerance.

Acknowledgements: JM would like to thank Darren Rowe for assistance with maintaining *N. diversicolor* holding tanks and Clare Redshaw for providing technical and scientific input. We thank Rob Clough for ICP-OES analysis.

Supporting Information: The supporting information submitted with this manuscript provides details of the primer sequences used for real-time PCR experiments including those specific for internal references (Table S1), transcriptome assembly and validation statistics (Table S2), Energy Dispersive X-ray Spectroscopy analysis data for the identification of Cu- and S-containing granules in *N. diversicolor* thin sections (Figure S1), and the results of a Cu toxicity assay for animals that had been placed in clean sediment for 21 days (Figure S2). This material is available free of charge via the Internet at http://pubs.acs.org.

References

1. Bryan, G. W.; Langston, W. J., Bioavailability, accumulation and effects of heavy metals in sediments with special reference to United Kingdom estuaries: a review. *Environmental Pollution* **1992,** *76*, (2), 89-131.
2. Langston, W. J.; Bebianno, M. J., *Metal Metabolism in Aquatic Environments*. Springer: 1998.
3. Scaps, P., A review of the biology, ecology and potential use of the common ragworm <i>Hediste diversicolor</i> (O.F. Müller) (Annelida: Polychaeta). *Hydrobiologia* **2002,** *470*, (1), 203-218.
4. Bryan, G., W.; Gibbs, P., E., Heavy Metals in the Fal Estuary, Cornwall: A Study of Long Term Contamination by Mining Waste and its Effects on Estuarine Organisms. *Occasional Publications. Marine Biological Association of the United Kingdom* **1983,** *2*.
5. Lai, F. M.; Cobuzzi, A.; Spinelli, W.; Williams, P. R.; Attrill, M. J.; Nimmo, M., Heavy Metal Concentrations and Bioaccumulation Within the Fal Estuary, UK: A Reappraisal. *Marine Pollution Bulletin* **1998,** *36*, (8), 643-645.
6. Pirrie, D.; Power, M. R.; Rollinson, G.; Camm, G. S.; Hughes, S. H.; Butcher, A. R.; Hughes, P., The spatial distribution and source of arsenic, copper, tin and zinc within the surface sediments of the Fal Estuary, Cornwall, UK. *Sedimentology* **2003,** *50*, (3), 579-595.
7. Hudson-Edwards, K. A.; Schell, C.; Macklin, M. G., Mineralogy and geochemistry of alluvium contaminated by metal mining in the Rio Tinto area, southwest Spain. *Applied Geochemistry* **1999,** *14*, (8), 1015-1030.
8. Leblanc, M.; Morales, J. A.; Borrego, J.; Elbaz-Poulichet, F., 4,500-Year Old Mining Pollution in Southwestern Spain: Long Term Implications for Modern Mining Pollution. *Economic Geology* **2000,** *95*, (3), 655-662.
9. Austen, M. C.; Somerfield, P. J., A community level sediment bioassay applied to an estuarine heavy metal gradient. *Marine Environmental Research* **1997,** *43*, (4), 315-328.
10. Bryan, G. W.; Hummerstone, L. G., Adaptation of the Polychaete Nereis Diversicolor to Estuarine Sediments Containing High Concentrations of Zinc and Cadmium. *Journal of the Marine Biological Association of the United Kingdom* **1973,** *53*, (04), 839-857.
11. Grant, A.; Hateley, J. G.; Jones, N. V., Mapping the ecological impact of heavy metals on the estuarine polychaete Nereis diversicolor using inherited metal tolerance. *Marine Pollution Bulletin* **1989,** *20*, (5), 235-238.
12. Mouneyrac, C.; Mastain, O.; Amiard, J. C.; Amiard-Triquet, C.; Beaunier, P.; Jeantet, A. Y.; Smith, B. D.; Rainbow, P. S., Trace-metal detoxification and tolerance of the estuarine worm Hediste diversicolor; chronically exposed in their environment. *Marine Biology* **2003,** *143*, (4), 731-744.
13. Pook, C.; Lewis, C.; Galloway, T., The metabolic and fitness costs associated with metal resistance in Nereis diversicolor. *Mar Pollut Bull* **2009,** *58*, (7), 1063-71.
14. Burlinson, F. C.; Lawrence, A. J., A comparison of acute and chronic toxicity tests used to examine the temporal stability of a gradient in copper tolerance of Hediste diversicolor from the Fal estuary, Cornwall, UK. *Mar Pollut Bull* **2007,** *54*, (1), 66-71.

15. Geffard, A.; Smith, B. D.; Amiard-Triquet, C.; Jeantet, A. Y.; Rainbow, P. S., Kinetics of trace metal accumulation and excretion in the polychaete Nereis diversicolor. *Marine Biology* **2005,** *147*, (6), 1291-1304.
16. Rainbow, P. S.; Smith, B. D.; Luoma, S. N., Differences in trace metal bioaccumulation kinetics among populations of the polychaete Nereis diversicolor from metal-contaminated estuaries. *Marine Ecology Progress Series* **2009,** *376*, 173-184.
17. Brown, B. E., THE FORM AND FUNCTION OF METAL-CONTAINING 'GRANULES' IN INVERTEBRATE TISSUES. *Biological Reviews* **1982,** *57*, (4), 621-667.
18. Poirier, L.; Berthet, B.; Amiard, J.-C.; Jeantet, A.-Y.; Amiard-Triquet, C., A suitable model for the biomonitoring of trace metal bioavailabilities in estuarine sediments: the annelid polychaete Nereis diversicolor. *Journal of the Marine Biological Association of the United Kingdom* **2006,** *86*, (01), 71-82.
19. Berthet, B.; Mouneyrac, C.; Amiard, J. C.; Amiard-Triquet, C.; Berthelot, Y.; Le Hen, A.; Mastain, O.; Rainbow, P. S.; Smith, B. D., Accumulation and Soluble Binding of Cadmium, Copper, and Zinc in the Polychaete <i>Hediste diversicolor</i> from Coastal Sites with Different Trace Metal Bioavailabilities. *Archives of Environmental Contamination and Toxicology* **2003,** *45*, (4), 468-478.
20. Pook, C. The bioenergetic cost of metal resistance and its consequences for reproduction in the harbour ragworm, Nereis diversicolor. University of Exeter, 2009.
21. Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W., NIH Image to ImageJ: 25 years of image analysis. *Nat Meth* **2012,** *9*, (7), 671-675.
22. Chevreux, B.; Pfisterer, T.; Drescher, B.; Driesel, A. J.; Muller, W. E.; Wetter, T.; Suhai, S., Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. *Genome Res* **2004,** *14*, (6), 1147-59.
23. Huang, X.; Madan, A., CAP3: A DNA sequence assembly program. *Genome Res* **1999,** *9*, (9), 868-77.
24. Schmieder, R.; Edwards, R., Fast identification and removal of sequence contamination from genomic and metagenomic datasets. *PLoS One* **2011,** *6*, (3), e17288.
25. Rozen, S.; Skaletsky, H., Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* **2000,** *132*, 365-86.
26. Simpson, E. H., Measurement of diversity. *Nature* **1949,** *163*, 688.
27. Margulies, M.; Egholm, M.; Altman, W. E.; Attiya, S.; Bader, J. S.; Bemben, L. A.; Berka, J.; Braverman, M. S.; Chen, Y. J.; Chen, Z.; Dewell, S. B.; Du, L.; Fierro, J. M.; Gomes, X. V.; Godwin, B. C.; He, W.; Helgesen, S.; Ho, C. H.; Irzyk, G. P.; Jando, S. C.; Alenquer, M. L.; Jarvie, T. P.; Jirage, K. B.; Kim, J. B.; Knight, J. R.; Lanza, J. R.; Leamon, J. H.; Lefkowitz, S. M.; Lei, M.; Li, J.; Lohman, K. L.; Lu, H.; Makhijani, V. B.; McDade, K. E.; McKenna, M. P.; Myers, E. W.; Nickerson, E.; Nobile, J. R.; Plant, R.; Puc, B. P.; Ronan, M. T.; Roth, G. T.; Sarkis, G. J.; Simons, J. F.; Simpson, J. W.; Srinivasan, M.; Tartaro, K. R.; Tomasz, A.; Vogt, K. A.; Volkmer, G. A.; Wang, S. H.; Wang, Y.; Weiner, M. P.; Yu, P.; Begley, R. F.; Rothberg, J. M., Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **2005,** *437*, (7057), 376-80.
28. Kumar, S.; Blaxter, M. L., Comparing de novo assemblers for 454 transcriptome data. *BMC genomics* **2010,** *11*, 571.
29. Smith-Unna, R., D.; Boursnell, C.; Kelly, S.; Hibberd, J., M. *Transrate v0.2.0*, 2014.
30. Parra, G.; Bradnam, K.; Korf, I., CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* **2007,** *23*, (9), 1061-1067.

31. Prohaska, J. R., Role of copper transporters in copper homeostasis. *The American Journal of Clinical Nutrition* **2008,** *88*, (3), 826S-829S.
32. La Fontaine, S.; Mercer, J. F. B., Trafficking of the copper-ATPases, ATP7A and ATP7B: Role in copper homeostasis. *Archives of Biochemistry and Biophysics* **2007,** *463*, (2), 149-167.
33. Zhou, H.; Cadigan, K. M.; Thiele, D. J., A Copper-regulated Transporter Required for Copper Acquisition, Pigmentation, and Specific Stages of Development in Drosophila melanogaster. *Journal of Biological Chemistry* **2003,** *278*, (48), 48210-48218.
34. Markossian, K. A.; Kurganov, B. I., Copper chaperones, intracellular copper trafficking proteins. Function, structure, and mechanism of action. *Biochemistry. Biokhimiia* **2003,** *68*, (8), 827-37.
35. Culotta, V. C.; Klomp, L. W. J.; Strain, J.; Casareno, R. L. B.; Krems, B.; Gitlin, J. D., The Copper Chaperone for Superoxide Dismutase. *Journal of Biological Chemistry* **1997,** *272*, (38), 23469-23472.
36. Hamza, I.; Schaefer, M.; Klomp, L. W. J.; Gitlin, J. D., Interaction of the copper chaperone HAH1 with the Wilson disease protein is essential for copper homeostasis. *Proceedings of the National Academy of Sciences* **1999,** *96*, (23), 13363-13368.
37. Suzuki, K. T.; Someya, A.; Komada, Y.; Ogra, Y., Roles of metallothionein in copper homeostasis: responses to Cu-deficient diets in mice. *Journal of Inorganic Biochemistry* **2002,** *88*, (2), 173-182.
38. Freedman, J. H.; Ciriolo, M. R.; Peisach, J., The role of glutathione in copper metabolism and toxicity. *J Biol Chem* **1989,** *264*, (10), 5598-605.
39. Emmerson, M., Remedial habitat creation: does Nereis diversicolor play a confounding role in the colonisation and establishment of the pioneering saltmarsh plant, Spartina anglica? *Helgoland Marine Research* **2000,** *54*, (2-3), 110-116.
40. Lushchak, V. I., Environmentally induced oxidative stress in aquatic animals. *Aquatic Toxicology* **2011,** *101*, (1), 13-30.
41. Geracitano, L. A.; Bocchetti, R.; Monserrat, J. M.; Regoli, F.; Bianchini, A., Oxidative stress responses in two populations of Laeonereis acuta (Polychaeta, Nereididae) after acute and chronic exposure to copper. *Marine Environmental Research* **2004,** *58*, (1), 1-17.
42. Lukkari, T.; Taavitsainen, M.; Soimasuo, M.; Oikari, A.; Haimi, J., Biomarker responses of the earthworm Aporrectodea tuberculata to copper and zinc exposure: differences between populations with and without earlier metal exposure. *Environmental Pollution* **2004,** *129*, (3), 377-386.
43. Regoli, F.; Principato, G., Glutathione, glutathione-dependent and antioxidant enzymes in mussel, Mytilus galloprovincialis exposed to metals under field and laboratory conditions: Implications for the use of biochemical biomarkers. *Aquatic Toxicology* **1995,** *31*, (2), 143-164.
44. Cunha, I.; Mangas-Ramirez, E.; Guilhermino, L., Effects of copper and cadmium on cholinesterase and glutathione S-transferase activities of two marine gastropods (Monodonta lineata and Nucella lapillus). *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* **2007,** *145*, (4), 648-657.