William & Mary
CHARTERED 1693

W&M ScholarWorks

VIMS Articles

Virginia Institute of Marine Science

2018

Effects-based spatial assessment of contaminated estuarine sediments from Bear Creek, Baltimore Harbor, MD, USA

SE Hartzell

Michael A. Unger
Virginia Institute of Marine Science

BL McGee

LT Yonkos

Follow this and additional works at: https://scholarworks.wm.edu/vimsarticles

Part of the Aquaculture and Fisheries Commons

Recommended Citation

Hartzell, SE; Unger, Michael A.; McGee, BL; and Yonkos, LT, "Effects-based spatial assessment of contaminated estuarine sediments from Bear Creek, Baltimore Harbor, MD, USA" (2018). *VIMS Articles*. 741.
https://scholarworks.wm.edu/vimsarticles/741

This Article is brought to you for free and open access by the Virginia Institute of Marine Science at W&M ScholarWorks. It has been accepted for inclusion in VIMS Articles by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.

Environ Sci Pollut Res (2017) 24:22158–22172
DOI 10.1007/s11356-017-9667-0

RESEARCH ARTICLE

Effects-based spatial assessment of contaminated estuarine sediments from Bear Creek, Baltimore Harbor, MD, USA

Sharon E. Hartzell[1] · **Michael A. Unger**[2] · **Beth L. McGee**[3] · **Lance T. Yonkos**[1]

Received: 17 February 2017 / Accepted: 27 June 2017 / Published online: 15 July 2017
© Springer-Verlag GmbH Germany 2017

Abstract Estuarine sediments in regions with prolonged histories of industrial activity are often laden to significant depths with complex contaminant mixtures, including trace metals and persistent organic pollutants. Given the complexity of assessing risks from multi-contaminant exposures, the direct measurement of impacts to biological receptors is central to characterizing contaminated sediment sites. Though biological consequences are less commonly assessed at depth, laboratory-based toxicity testing of subsurface sediments can be used to delineate the scope of contamination at impacted sites. The extent and depth of sediment toxicity in Bear Creek, near Baltimore, Maryland, USA, was delineated using 10-day acute toxicity tests with the estuarine amphipod *Leptocheirus plumulosus*, and chemical analysis of trace metals and persistent organic pollutants. A gradient of toxicity was demonstrated in surface sediments with 21 of 22 tested sites differing significantly from controls. Effects were most pronounced (100% lethality) at sites proximate to a historic industrial complex. Sediments from eight of nine core samples to depths of 80 cm were particularly impacted (i.e., caused significant lethality to *L. plumulosus*) even in locations overlain with relatively non-toxic surface sediments, supporting a conclusion that toxicity observed at the surface (top 2 cm) does not adequately predict toxicity at depth. In seven of nine sites, toxicity of surface sediments differed from toxicity at levels beneath by 28 to 69%, in five instances underestimating toxicity (28 to 69%), and in two instances overestimating toxicity (44 to 56%). Multiple contaminants exceeded sediment quality guidelines and correlated positively with toxic responses within surface sediments (e.g., chromium, nickel, polycyclic aromatic hydrocarbon (PAH), total petroleum hydrocarbon). Use of an antibody-based PAH biosensor revealed that porewater PAH concentrations also increased with depth at most sites. This study informs future management decisions concerning the extent of impact to Bear Creek sediments, and demonstrates the benefits of a spatial approach, relying primarily on toxicity testing to assess sediment quality in a system with complex contaminant mixtures.

Keywords Sediment toxicity · Sediment chemistry · *Leptocheirus plumulosus* · Trace metals · Persistent organic pollutants · Pore water

Responsible editor: Ester Heath

Electronic supplementary material The online version of this article (doi:10.1007/s11356-017-9667-0) contains supplementary material, which is available to authorized users.

✉ Lance T. Yonkos
lyonkos@umd.edu

[1] Environmental Science and Technology Department, University of Maryland, 1451 Animal Science Bldg, College Park, MD 20742-2315, USA

[2] Virginia Institute of Marine Science, College of William & Mary, 1208 Greate Road, Gloucester Point, VA 23062, USA

[3] Chesapeake Bay Foundation, Philip Merrill Environmental Center, 6 Herndon Avenue, Annapolis, MD 21403, USA

Introduction

Sediments, particularly those of harbors and other urban estuarine systems surrounded by myriad point and nonpoint sources of chemical pollutants, are the principal repository for persistent environmental contaminants (Burton 2002; Ridgway and Shimmield 2002; Wang et al. 2015). As such, sediments in regions with prolonged histories of industrial activity are often laden to significant depths with complex mixtures of legacy contaminants, including trace metals and

persistent organic pollutants (POPs), which can be disturbed and reintroduced to surface waters, allowing for exposure, uptake, and accumulation in resident biota (Garcia et al. 2011; Palma et al. 2014). To address the complexity in assessing contaminated sediment risks, researchers have relied on integrated "weight-of-evidence" approaches that pair chemical analysis with laboratory and field-based biological effects data (Burton 2002). This is the underlying strategy of the sediment quality triad approach which combines sediment chemistry, laboratory-based sediment toxicity testing, and benthic community structure to determine the relative risk of a particular site (Chapman 1990). Given the complexity of cause-and-effect relationships between multi-contaminant exposures and biological consequences, toxicity evaluations play a central role in sediment quality monitoring, by providing realistic insight into the potential of sediments to elicit deleterious biological effects (Wang et al. 2015).

Progressive improvements in analytical capabilities have led risk assessment efforts to increasingly rely on sediment quality guidelines to predict impacts to biological receptors. However, biological effects cannot easily be extrapolated from chemical measures, especially in instances where multiple contaminants are present. In cases of complex contamination, the inability to identify specific causative agents should not be a barrier for delineation of regions of possible concern. Integrative assessments of chemical characteristics and biological responses within a spatial framework can, at a minimum, identify those regions requiring more intense investigation (Palma et al. 2014). Laboratory-based toxicity tests have been widely used to establish the toxicity of field-collected surface sediments (United States Environmental Protection Agency 1994). However, using toxicity testing of subsurface sediments to delineate the depth of contamination is a relatively uncommon technique. Most assessments of toxicity at depth rely exclusively on comparing chemical data from core segments to sediment quality guidelines (Xu et al. 2014; Chen et al. 2016). Exceptions include Wang et al. (2015) who employed laboratory toxicity tests of sediments from various depths within field-collected cores from the Yangtze River estuary, Shanghai, China; Ingersoll et al. (2009) who used similar methods to investigate toxicity at depth in sediment from the Ashtabula River in Northeastern Ohio, USA; and Borgmann and Norwood (2002) who analyzed metal bioavailability and toxicity to *Hyalella azteca* in sediment cores from Richard Lake in Ontario, Canada.

In the present study, laboratory-based toxicity testing was used as a tool for delineating not only the horizontal but also the vertical extent of sediment contamination at a historically impacted site. Bear Creek in Baltimore, Maryland, was selected as the study site (Fig. 1). The sediments at this site have been extensively studied and found to contain a complex mixture of inorganic (e.g., zinc, chromium, copper, nickel, and lead) and organic (e.g., polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs), and organochlorine pesticides) contaminants (Baker et al. 1997; Ashley and Baker 1999; Hartwell and Hameedi 2007; Graham et al. 2009). Studies of surface sediment toxicity, performed over the past two decades, have also demonstrated persistent acute lethality to benthic organisms (McGee et al. 1999, 2004; Klosterhaus et al. 2006). However, attempts to assign causal links between particular contaminants and observed toxicity, either in laboratory-based toxicity tests or assessments of benthic community structure, have been largely unsuccessful (McGee et al. 1999, 2004).

Bear Creek, like many historically contaminated sites, is in the vicinity of residential and recreational areas; therefore, the potential exists for contaminant-related human health impacts (see supplemental material Fig. S1). In regions such as this, where site remediation is a priority and redevelopment is a likely future scenario, a thorough understanding of the spatial extent of contaminant-induced toxicity is necessary to guide management decisions. Absent an individual causative agent or clear understanding of multi-contaminant interactions, delineation of impacted regions based solely on chemical data is inadequate (Chapman 2002; Fleeger et al. 2007; Kwok et al. 2014; Pourabadehei and Mulligan 2016). The first objective of the present study was to demonstrate the usefulness of toxicity testing as a tool for delineating contaminant effects, particularly the unconventional use of toxicity tests on subsurface sediments from cores. Two hypotheses related to the distribution of chemical contaminants and toxicity were tested: first, that contaminant loads and toxicity of sediments would decrease with distance from legacy sources and second, that contaminant loads and toxicity within sediment cores would correspond with temporal trends of contaminant release from industrial sources. An additional objective of the study was to compare toxicity results at surface and at depth to various contaminant classes, particularly the bioavailable fraction of PAHs found in porewater. In pursuing this objective, conventional analytical techniques were used to measure various whole-sediment metal and POP concentrations, whereas an antibody-based PAH biosensor was used to measure porewater PAH concentrations (Li et al. 2016). The field-portable device required minimal sample volumes (≤10 mL) allowing measurement of total PAHs within the limited porewater available from sediment core segments, demonstrating the value of the technique for low-cost pollution mapping at contaminated sites.

Materials and methods

Site description and sample collection

The area of investigation within Bear Creek was bordered on the east by the Sparrows Point Industrial Complex and on the

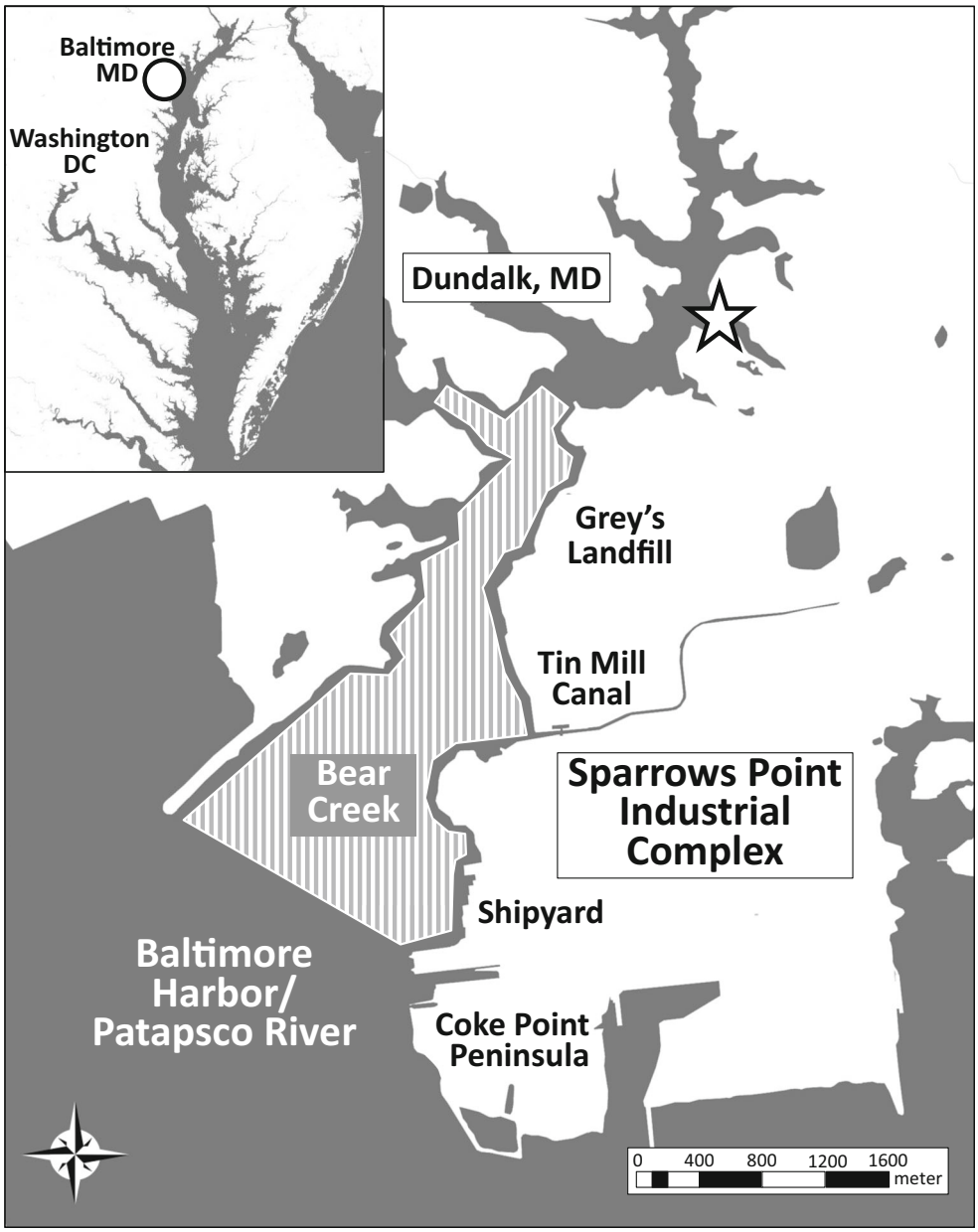

Fig. 1 Bear Creek, a tributary to the Chesapeake Bay near the mouth of the Patapsco River and Baltimore Harbor, MD, USA (*inset*, *circle*); bordered to the east by the Sparrows Point Industrial Complex and the northwest by the community of Dundalk, MD; hashed region indicates area of investigation; *star* indicates location of in-system reference site

northwest by the community of Dundalk, MD (Fig. 1). Twenty two stations for sediment toxicity testing and chemical analysis were established in a grid comprised of roughly east-west oriented transects (Fig. 2). Sites were located in areas of previously established and/or suspected toxicity, and at increasing distances from the shoreline (McGee et al. 1999, 2004). Sites were also constrained to depositional areas dominated by fine-grained sediments (supplemental material Table S2), reflecting expectations that contaminants would sorb to and deposit with sediments, and that fine-grained sediments hold more contaminants than coarse-grained sediments due to their greater surface area (Forstner 1989).

Surface sediments from 20 stations were collected by boat-mounted Ponar® grab sampler on 3–4 June 2015, according to standard protocols (American Society for Testing and Materials 1994). On 11 June 2015, surface sediments from two more sites were sampled (H2 and J1, Fig. 2). Reference surface sediment was collected from Bigwood Cove, a tributary to the Wye River on the eastern shore of Maryland. An in-system reference sample was also collected from a historically non-toxic location within Bear Creek approximately 2 km northeast of the region of investigation. Sediments from the in-system reference site have similar characteristics to those of the impacted sediments in Bear Creek (Table 2). While not pristine, sediments from the reference site have been shown in previous investigations to have lower contaminant loads than other sediments in Bear Creek (Baker et al. 1997), and to not exhibit significant toxicity to benthic organisms during 10-day

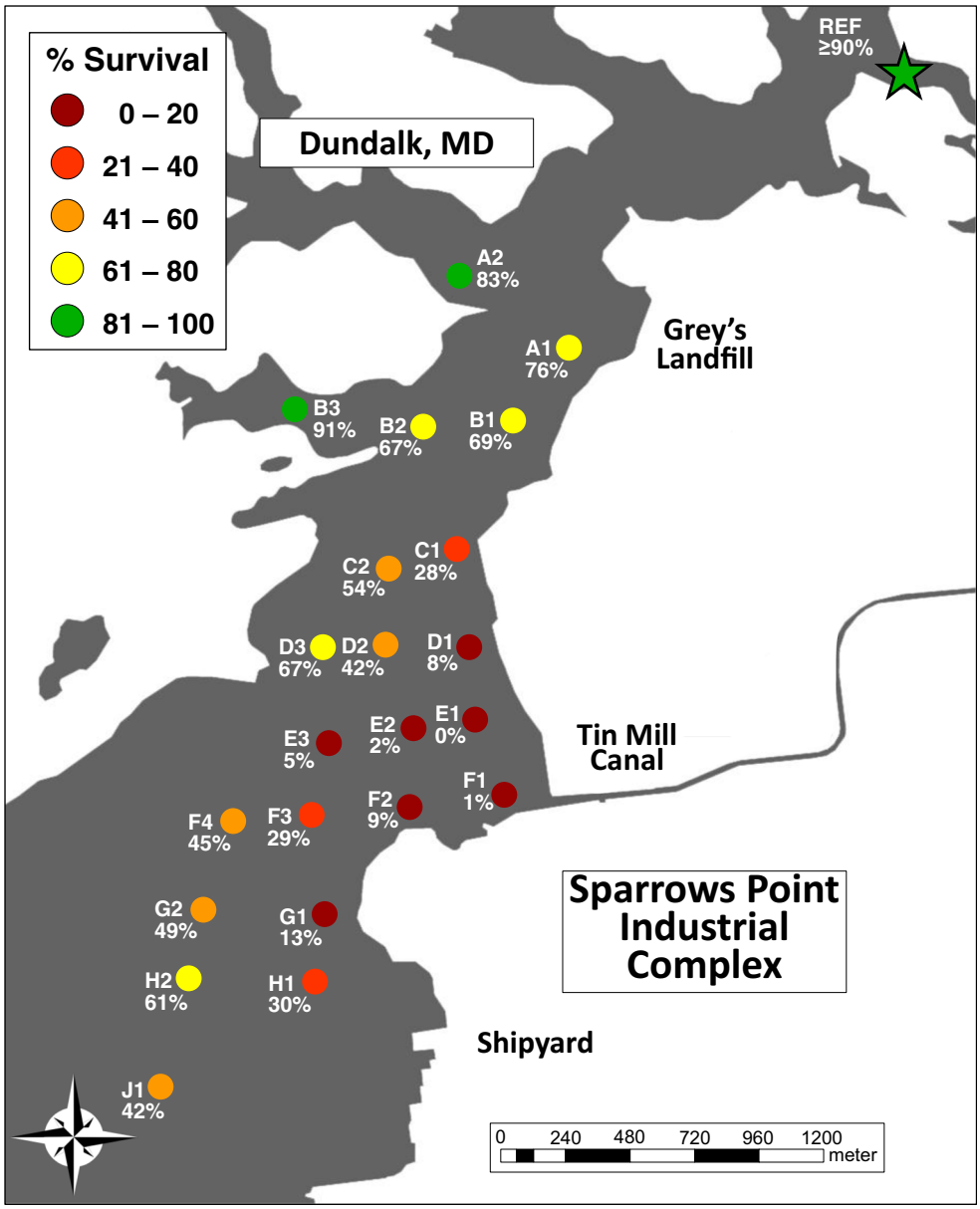

Fig. 2 Spatial display of amphipod *Leptocheirus plumulosus* survival (color-coded in 20% increments) in 10-day toxicity tests of surface sediments collected from Bear Creek, Baltimore, MD, USA, June 2015; *star* indicates in-system reference site

sediment toxicity testing (unpublished data). Each surface sample was generated by combining the oxic layer (approximately the top 2 cm of sediment) from five replicate grab samples and homogenizing in a stainless steel bowl. Surface sediment samples were meant to reflect conditions in the well oxygenated layer where most benthic organisms would reside. Subsamples for chemical analysis were placed into 250-mL certified-clean amber wide-mouth jars (×2) and pre-cleaned 1-L Mason jars. Remaining sediments for toxicity testing were placed in pre-cleaned 2.5-L HDPE containers. All material was held on ice while in the field, then stored at 4 °C (toxicity samples) or −20 °C (chemistry samples) before processing for testing or analysis.

Core samples (10 cm diameter × 80 cm depth) were collected from a subset of six sites (B1, C1, F1, F2, F3, and F4) on 28 July, 2015, and from another three sites (B2, B3, and C2) on 17 September, 2015. Core sites were selected based on results from toxicity tests performed on surface sediments. Sediment cores were collected using an in-house fabricated coring device that employed single-use PVC core housings and a vacuum piston to displace water and retain samples. Sediment cores were capped for transport to the University of Maryland, and then stored vertically at 4 °C until processed. Before the initiation of sediment tests, PVC core housings were split lengthwise with electric sheet-metal shears and cores were segmented by depth, at 10 cm increments from 0 to 40 cm and 20 cm increments from 40 to 80 cm. The resulting volume of sediment in the 10 cm segments was 785 mL. Core segments were thoroughly homogenized, 100 mL subsamples were centrifuged (3500*g* × 15 min), and

supernatant (porewater) was stored in 60-mL pre-cleaned amber glass jars for PAH analysis. Remaining sediments were stored in re-sealable plastic zipper bags at 4 °C before processing for toxicity testing or analysis.

Sediment toxicity tests

Toxicity of sediment samples was investigated using the benthic estuarine amphipod *Leptocheirus plumulosus* following methods described in the *Methods for Assessing the Toxicity of Sediment-associated Contaminants with Estuarine and Marine Amphipods* (United States Environmental Protection Agency 1994), with any deviations noted. Test organisms (2 to 4 mm length) were provided by Chesapeake Cultures, Inc., Hayes, VA. Chambers used for the tests were glass 1-L Mason jars, which were pre-cleaned with soap and water, 50% nitric acid, and acetone, and triple rinsed with deionized water. One day prior to test initiation (i.e., day −1), aged and aerated dechlorinated municipal water was temperature (23 °C) and salinity (15 ppt; Crystal Sea® Bioassay Formula Marinemix) adjusted to approximate estuarine conditions, and was added to test chambers containing 175 mL of sediment to a final volume of 1.0 L. Overlying water was siphoned from chambers and replaced after 24 h before initiating exposures by loading test organisms on day 0. Water quality parameters (pH, DO, temperature, salinity, NH_3) were measured in overlying water from all test chambers prior to and after removal of water on day 0 and daily from one replicate/treatment until the conclusion of the test. Two rounds of toxicity bioassays were performed with surface sediments. These bioassays used five replicate chambers with 20 amphipods loaded in each test chamber. Organisms were not fed during 10-day static tests. The first exposure round began 9 June 2015 and the second began 23 June 2015. An additional test chamber was prepared for each treatment on day −1 for the purpose of determining porewater ammonia. On day 0, water was decanted from these chambers, sediment was centrifuged at 3000*g* for 15 min, and supernatant was collected for porewater ammonia analysis by the Nesslerization method (United States Environmental Protection Agency 1983) using a LaMotte Smart3 Colorimeter (LaMotte, Chestertown, MD).

Toxicity tests for sediment core samples were conducted under the same protocols as surface sample tests (United States Environmental Protection Agency 1994) with several necessary modifications. Given space limitations of the testing table, cores were tested in groups of three, with exposures initiated 4 August 2015 for cores B1, C1, and F1; 28 August 2015 for cores F2, F3, and F4; and 8 October 2015 for cores B2, B3, and C2. Core segments only provided sufficient sediment volume to run four replicate chambers at volumes of 150 mL/chamber. Reference sediments were accordingly prepared with only 150 mL of sediment/chamber but were still tested using five replicate chambers. Limited material also precluded setup of an additional beaker for the sole purpose of porewater ammonia analysis. Instead, ammonia was measured in the porewater from the subsample collected for PAH analysis. Because this material was not allowed to sit in test chambers containing overlying water for 24-h before analysis, porewater values for core segment samples potentially overestimate the ammonia concentrations that were present in beakers at the time that organisms were loaded.

Sediment chemical analysis

Whole sediments from surface samples were transported on ice to the University of Maryland Chesapeake Biological Laboratory, Solomons, MD, for trace metal analysis. Following acid digestion (HNO_3) as described in US Environmental Protection Agency Method 3052, concentrations of chromium (Cr), zinc (Zn), Copper (Cu), nickel (Ni), lead (Pb), vanadium (V), and arsenic (As) were determined by inductively coupled plasma-mass spectrometry (ICP-MS; Hewlett-Packard 4500) using USEPA Method 6020 (United States Environmental Protection Agency 1995, 2014). Whole sediments from all surface samples were analyzed for PAHs and total petroleum hydrocarbon (TPH) and a subset of 10 samples were also analyzed for PCBs. Persistent organic pollutant analyses were performed by the Geochemical and Environmental Research Group (GERG) at Texas A&M University using methods described previously by Kirman et al. (2016). Analysis was by GC/MS with TPH and PAHs measured using a flame ionization detector and PCBs measured in selected ion mode (SIM). Total sediment PAH (PAH_{sed}) was calculated as the sum of 46 targeted PAH congeners and total PCB was calculated as the sum of 175 targeted PCB congeners.

Porewater PAH analysis

Aqueously available PAH (PAH_{pw}) was determined by analyzing porewater fractions from all surface sediment samples and from all segments of sediment cores using an antibody-based PAH biosensor. The antibody used was 2G8, an anti-pyrene-butyric acid monoclonal antibody produced in mice, which is sensitive to all 3- to 5-ring PAHs (Li et al. 2016). Sediment samples were centrifuged (3500*g* × 15 min) to extract porewater, which was then glass-fiber filtered to 0.7 μm. Spier et al. (2011) describe the automated sample-handling program of the KinExA Inline Sensor. Phenanthrene standards were used to create a standard curve to estimate final concentrations in porewater samples. Analysis of PAH_{pw} was conducted according to methods in Li et al. (2016). PAH_{pw} was used as a proxy for other chemicals in sediment cores, given cost, and sediment volume constraints which limited the analysis that could be performed on toxicity-tested sediments.

Statistical analysis

Survival data (the only toxicological endpoint for the *Leptocheirus plumulosus* test) were arcsine square root transformed and tested by one-way analysis of variance (ANOVA) followed by Dunnett's multiple comparison. Data not satisfying normality or homogeneity of variance requirements for parametric statistics were tested using Kruskal-Wallis one-way ANOVA on ranks followed by Dunn's multiple comparison. Treatment means were compared to both in-system and out-of-system reference sediments. Pearson product moment correlation coefficients were computed to measure the strength of association between variables including toxicity test survival and each chemical parameter. Statistical analyses were carried out using Sigma-Stat version 12.0 (Systat Software, Inc., San Jose, CA) with significance set at a minimum probability level of $p = 0.05$.

Results

Toxicity tests—surface sediments

Water quality parameters in overlying water were monitored throughout the surface sediment toxicity tests and all values remained within acceptable parameters (water quality provided in supplemental material Table S3). Survival in the 22 Bear Creek surface sediment toxicity tests ranged from 0 to 91% (Table 1; individual replicate survival data provided in supplemental material Table S4). Of the 11 test sites included in the first toxicity test series, all but 1 (B3) were found to differ significantly from the Bigwood Cove reference sediment, which had 98% survival. Eight of 11 sediments also differed significantly from the in-system reference sediment (90% survival), exceptions were A1, A2, and B3. All 11 test sites included in the second toxicity test series were found to differ significantly from Bigwood Cove and in-system reference sediments which had survivals of 99 and 98%, respectively. Observed toxicity was greatest in surface sediments from sites most proximate to Tin Mill Canal, a drainage swale discharging surface runoff and groundwater leachate from much of the industrial site (Fig. 1), and diminished with increasing distance from this location. Survival of amphipods in 10-day toxicity tests was strongly correlated with distance from the location of the Tin Mill Canal discharge ($r^2 = 0.77$) (Fig. 3). More generally, survival data indicated an east to west gradient along longitudinal transects, with locations nearest the industrial complex causing the most lethality and locations further from the complex causing the least.

Trace metals analyses—surface sediments

Surface sediment metal concentrations are reported in Table 1 with *effects range median* (ERM) values (indicating the concentration at which biological effects were observed 50% of the time in other studies; Long and Morgan 1991) included for comparative purposes. Zinc concentrations ranged from 605 to 2295 μg/g and exceeded the ERM (410 μg/g) throughout the entire system (Table 1; Fig. 4a). Chromium concentrations ranged from 467 to 3196 μg/g and also exceeded the ERM (370 μg/g) throughout the system (Table 1; Fig. 4b). Sediment samples from 14 sites exceeded the 52 μg/g Ni ERM (Table 1, Fig. 4c). Nine sites exceeded the 220 μg/g Pb ERM, four exceeded the 9.6 μg/g Cd ERM, two exceeded the 270 μg/g Cu ERM, and none exceeded the 70 μg/g As ERM (Table 1). In general, within east-west transects, metal concentrations were highest in proximity to the industrial complex and diminished with increasing distance. A notable exception was transect B where most metals concentrations were highest at B3, the westernmost sample site.

Persistent organic pollutants analyses—surface sediments

Total PAH, total PCB, and TPH in surface sediments are reported in Table 1 with sediment reference guidelines provided for comparative purposes (Long and Morgan 1991; Inouye 2014). Individual PAH and PCB congener results are provided in supplemental material (Tables S5 and S6, respectively). For all contaminant classes, several sites within Bear Creek exceeded sediment quality guidelines. Sediment PAH concentrations (PAH_{sed}) ranged from 12.3 to 98.4 μg/g, exceeding the 35.0 μg/g ERM at 7 of the 22 sites. Sites bordering the southern portion of the industrial complex (E1, F1, G1, H1, J1) had the highest concentrations (Fig. 4d). Concentrations of TPH ranged from 1547 to 54,684 μg/g with 17 of the 22 sites exceeding the 3600 μg/g screening level sediment quality standard (SL-SQS) and highest concentrations at F1 and E1 (>15-fold exceedance and >7-fold exceedance, respectively; Fig. 4e). Concentrations of total PCB (only tested in 10 samples) ranged from 0.37 to 1.16 μg/g with 7 of 10 sites exceeding the 0.40 μg/g ERM (highest concentrations again at F1 and E1).

Porewater PAH concentrations from surface sediment samples, measured by antibody-based PAH biosensor, ranged from 0.30 to 6.72 μg/L (Table 1; Fig. 4f). Concentrations were again highest along the southern margin of the industrial facility (E1, F1, G1) and tended to diminish with increasing distance. Currently, no established screening guideline exists for PAH contamination of sediment porewater. Only a small fraction of the total PAH measured in whole sediments was found in porewater, reflecting the dominant partitioning of the hydrophobic compounds to the solid phase. Ratios of PAH_{sed} (measured by GC/MS) to PAH_{pw} (measured by PAH biosensor) for the 22 surface sediment samples were on the order of 10^4 to 10^5, closely reflecting the range of K_{OC} values for 3- to 5-ring PAHs (Chiou et al. 1998).

Table 1 Results of sediment toxicity tests (% survival ± standard deviation) and chemical analysis of surface sediments collected from Bear Creek, Baltimore MD, June 2015

Site	Survival (%)	Whole sediment (dry wt)										Porewater	
		Zn (μg/g)	Cr (μg/g)	Cu (μg/g)	Ni (μg/g)	Pb (μg/g)	As (μg/g)	Cd (μg/g)	PAH (μg/g)	PCB (μg/g)	TPH (μg/g)	PAH (μg/L)	Ammonia (mg/L)
Reference	99 ± 2.2[a] 98 ± 2.7[b]	–	–	–	–	–	–	–	–	–	–	–	4.3
In-system reference	98 ± 2.7[a] 90 ± 17.3[b]	*1018*	*455*	132	45	134	17	5.1	8.9	–	2485	0.49	4.6
A1	**76 ± 16.7**[b]	*1363*	*913*	164	*58*	203	29	7.6	12.3	*0.4*	*5779*	0.47	2.3
A2	83 ± 9.8[b]	*1525*	*966*	196	48	*231*	30	*13.0*	15.3	–	*12,358*	0.41	2.3
B1	**69 ± 14.8**[b]	*1514*	*987*	156	*66*	194	34	*12.7*	13.8	*0.5*	*7504*	0.51	1.5
B2	**67 ± 9.1**[b]	*1586*	*1066*	183	*60*	*239*	30	7.4	16.3	–	*6204*	0.39	1.6
B3	91 ± 8.2[b]	*2230*	*1220*	*316*	*83*	*352*	44	*10.3*	16.1	0.39	*4206*	0.39	1.9
C1	**28 ± 13.0**[b]	*2295*	*2686*	229	*118*	*225*	51	9.3	15.3	0.63	*11,745*	0.51	2.4
C2	**54 ± 17.1**[b]	*1390*	*932*	171	22	*240*	12	2.6	20.6	–	*4436*	0.50	2.2
D1	**8 ± 8.4**[b]	*1569*	*1153*	164	*63*	177	28	6.6	16.2	*0.76*	*19,254*	0.80	3.4
D2	**42 ± 12.0**[b]	*1168*	*853*	148	*53*	190	22	5.0	18.4	0.37	*5680*	0.83	2.7
D3	**67 ± 14.8**[b]	*920*	*547*	79	33	184	36	4.1	26.0	–	1908	0.38	2.1
E1	**0**[a]	*2295*	*2686*	229	*118*	*225*	51	9.3	*45.0*	*1.16*	*25,794*	2.32	9.7
E2	**2 ± 2.7**[a]	*1451*	*1282*	176	*72*	179	28	5.5	15.7	*0.44*	*22,479*	0.51	9.5
E3	**5 ± 5**[a]	*1146*	*882*	138	48	184	31	6.4	34.9	–	*4517*	0.30	6.6
F1	**1 ± 2.2**[a]	*1416*	*3196*	213	*128*	96	25	5.0	*49.9*	*1.09*	*54,684*	4.37	11.8
F2	**9 ± 8.9**[a]	*1655*	*1240*	177	*76*	189	34	6.9	*39.6*	–	*15,258*	0.90	6.8
F3	**29 ± 17.8**[a]	*983*	*720*	140	*59*	175	27	5.3	23.2	0.38	*5759*	0.46	9.9
F4	**50 ± 14.1**[a]	*1379*	*800*	187	44	*485*	62	4.8	*42.2*	–	2504	0.55	8.6
G1	**13 ± 11.1**[a]	*2001*	*1721*	*343*	*69*	*384*	45	*13.2*	*98.4*	–	*12,680*	6.72	4.6
G2	**49 ± 21.6**[a]	*986*	*857*	114	37	*259*	56	2.5	34.9	–	1557	0.43	11.1
H1	**34 ± 12.5**[b]	*706*	*601*	144	*62*	146	30	4.9	*44.4*	–	*4457*	0.62	2.6
H2	**61 ± 13.9**[a]	*605*	*467*	98	39	151	25	2.2	27.3	–	1742	0.42	9.5
J1	**42 ± 29.1**[a]	*775*	*615*	142	44	190	28	2.6	*57.7*	–	1547	0.64	6.2
Sediment reference guideline[c]		410	370	270	52	220	70	9.6	35.0	0.4	*3600*	–	60

PAH polycyclic aromatic hydrocarbons, *PCB* polychlorinated biphenyls, *TPH* total petroleum hydrocarbon

[a] Exposure assay begun June 9, 2015

[b] Exposure assay begun June 23, 2015

[c] Reference guidelines: effects range-median (ERM) for metals, PAH and PCB (Long and Morgan 1991); proposed screening level sediment quality standards (SL-SQS; Inouye 2014) for TPH; water quality standard for ammonia (USEPA 1994); concentrations in italics exceed guidelines; bold survival valued differ significantly from reference treatment (Kruskal-Wallis one-way ANOVA; $p = 0.05$)

Relationship of surface sediment toxicity to contaminant classes

For both inorganic and organic contaminants, sediments in Bear Creek frequently exceeded reference values for sediment quality, indicating the potential of these contaminants, individually or collectively, to cause toxicity within the system (Table 1). Relationships between specific contaminant classes and observed mortality were explored through a Pearson Product Moment Correlation, summarized in Table 2. Mortality had significant positive correlations with Cr, Ni, PAH_{sed}, PAH_{pw}, and TPH, whereas correlations with Zn and Cu were only weakly positive and not statistically significant. Contaminants tended to covary with one another such that sediments that caused the greatest toxicity also carried the greatest loads of virtually all contaminant categories (Table 2). Curiously, PAH_{pw} had significant positive correlations with Cr, Cu, and Ni (but not Zn), while PAH_{sed} did not have a significant correlation with any of the measured metals (Table 2). This suggests, at least within the Bear Creek system, that PAH_{pw} may be a better proxy for estimating trends in concentrations of metals than PAH_{sed}. Porewater ammonia

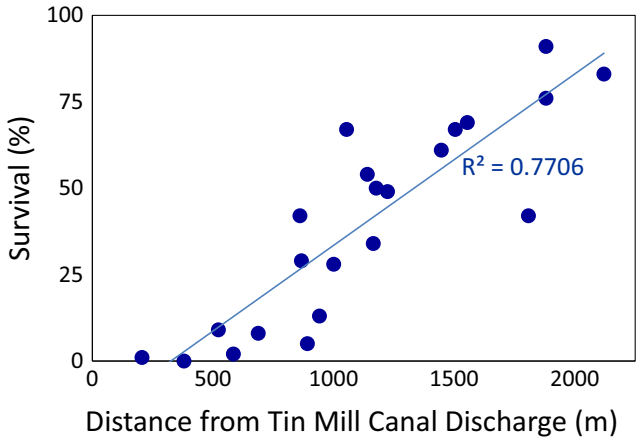

Fig. 3 Linear relationship between *Leptocheirus plumulosus* survival in 10-day sediment toxicity tests and the distance of the Bear Creek surface sediment source from the point of discharge of Tin Mill Canal

(NH_3) also correlated positively with mortality. However, because concentrations were <12 mg/L, well below the 60 mg/L test threshold (United States Environmental Protection Agency 1994), NH_3 likely was not a significant contributor to observed mortality in surface sediments.

Generally, the region of greatest toxicity and greatest contamination was arrayed in proximity to the central portion of the industrial complex (i.e., E1, F1 near the Tin Mill Canal discharge; Fig. 2) with POP concentrations comparatively high to the south and metal concentrations more variable and high to the north (Fig. 4a–f). The spatial relationship between toxicity and contamination was particularly evident for Cr, Ni, PAH_{sed}, PAH_{pw}, and TPH. Concentrations of other trace metals also co-occurred with substantial toxicity in proximity to the Tin Mill Canal discharge; however, several sites exhibiting minimal toxicity (e.g., A2, B3) had disproportionately high metals (e.g., Zn at site B3; Fig. 4a).

Core sediment toxicity, ammonia, and porewater PAHs

Water quality parameters in overlying water were monitored throughout the 10-day core segment toxicity tests (supplemental material Table S7). Dissolved oxygen (DO) remained above the acceptable minimum (≥60 saturation) in all treatments. However, pH decreased over time in the overlying water of several treatments; most notably, segments from the B3 core, and to a lesser extent, segments from the B2 and C2 cores. Ammonia in sediment porewater increased with depth in most cores, exceeding the recommended test level of 60 mg NH_3/L in six of nine cores at lower depths (supplemental material Table S8). Test chambers holding sediments with high porewater ammonia levels saw progressive increases in ammonia in overlying water during 10-day amphipod exposures. Most notable was the F1 core which had porewater NH_3 concentrations substantially in exceedance of 60 mg/L at all depths, causing ammonia in overlying water to reach concentrations as high 65 mg/L.

Bear Creek sediment cores revealed a variety of vertical and horizontal patterns with respect to organism survival (Fig. 5; replicated survival results in supplemental material Table S9). Cores from B1, C1, F1, and F2, all near the industrial complex shoreline, showed substantial mortality throughout entire column depths from surface to 80 cm. Compared to toxicity results from surface grab samples (comprised of the top 2 cm of multiple grabs), cores from B1 and C1 were substantially more toxic below the surface (i.e., homogenized sediment from the top 10 cm). Cores from B2 and C2 had a different pattern with a moderately toxic surface layer underlain by a relatively non-toxic top 10 cm, but with increasing toxicity at depth. Toxicity of sediments from B3 and F3 were relatively non-toxic near the surface, with mortality peaking in middle sections and survival improving with increased depth. The F4 core showed the greatest survival at depth, with only the top 20 cm exhibiting toxicity. Spatially, cores taken proximate to the industrial complex (B1, C1, F1, and F2) were intensely toxic from the surface to 80 cm. At virtually all depths severity of toxicity decreased with increasing horizontal distance from the industrial complex (Fig. 5).

This study also used PAH_{pw} as a proxy for total contaminant loads at depth and found that concentrations followed a similar pattern to mortality. In eight of nine sediment cores, PAH_{pw} concentrations increased with depth, with 60–80 cm segments exceeding 0–10 cm segments by up to 37-fold (Fig. 5). This was most apparent in the F1 core (adjacent to the Tin Mill Canal discharge) which had progressively increasing PAH_{pw} concentrations from 13.6 μg/L in the 0–10 cm segment to 40.1 μg/L in the 60–80 cm segment. In comparison, the PAH_{pw} concentration in the F1 surface sediment sample (top 2 cm) was only 4.4 μg/L. The only exception to this pattern was the F4 core, which had modestly higher PAH_{pw} concentrations at 10–40 cm depths (maximum 6.7 μg/L) than 40–80 cm depths (2.1 μg/L). In several cores, PAH_{pw} concentrations also differed considerably between the surface (top 2 cm) and the homogenized top 10 cm. Cores from B1 and C1, which caused greater mortality in top 10 cm segments than surface samples, also had 1.5-fold and 3.6-fold higher PAH_{pw} concentrations, respectively, in top 10 cm segments than surface sediments. Other cores had similarly large disparities between PAH_{pw} concentrations in the surface layer and the top 10 cm. For example, PAH_{pw} concentra-

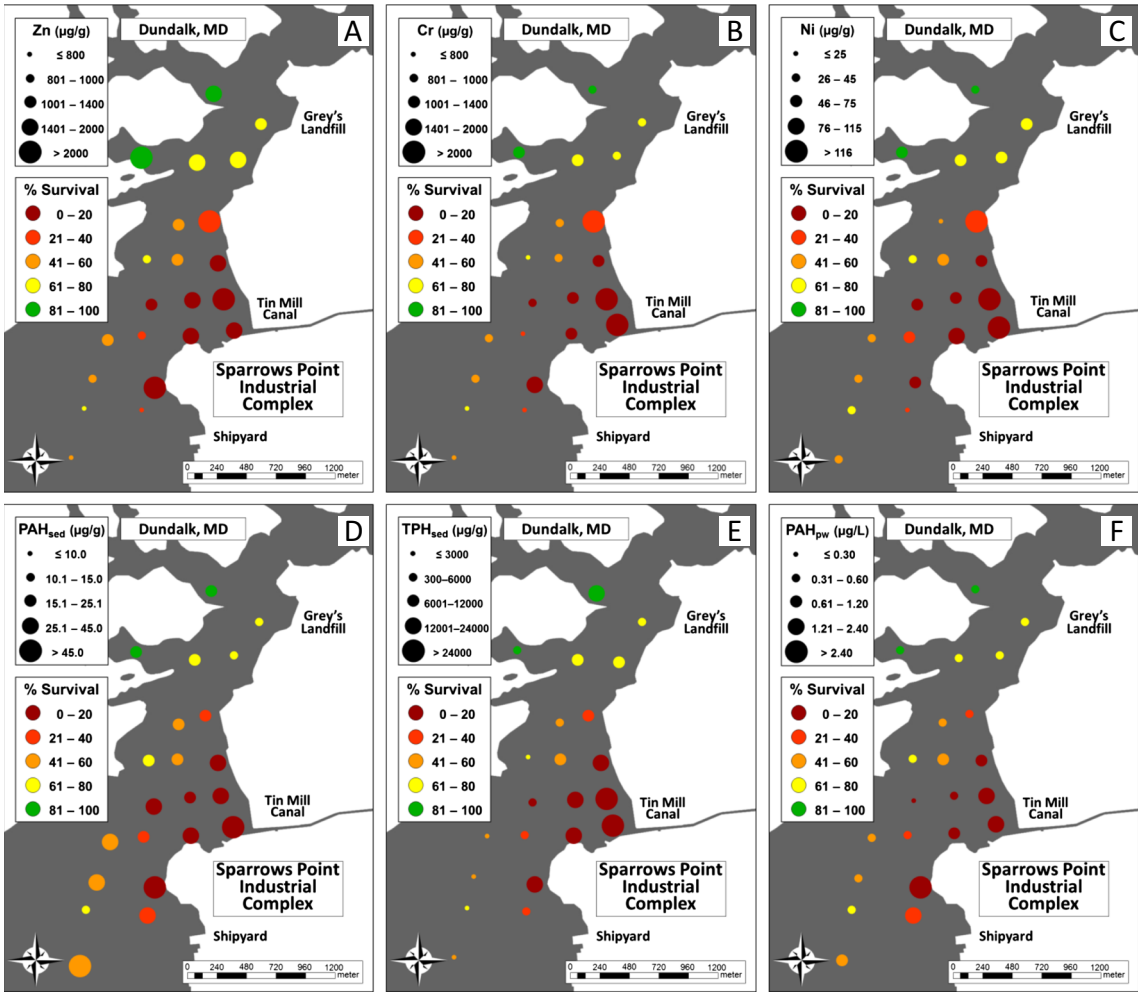

Fig. 4 Spatial relationship of contaminant concentrations in surface sediments (top 2 cm): **a** zinc, **b** chromium, **c** nickel, **d** total sediment PAH (PAH_{sed}), **e** total petroleum hydrocarbon (TPH_{sed}), and **f** porewater PAH (PAH_{pw}); *marker size* indicates relative contaminant concentration (*top legend*); *marker color* indicates observed toxicity (*bottom legend*)

tions in the top 10 cm of cores from F1, F2, and F4 were 3.1-fold, 4.3-fold, and 1.9-fold higher, respectively, than in surface sediments.

Discussion

Results of the current study confirm those of previous investigations in the region, which show that surface sediments are severely toxic and that they contain a spatially variable mixture of trace metals and persistent organic contaminants (McGee et al. 1999, 2004; Baker et al. 1997; Ashley and Baker 1999; Hartwell and Hameedi 2007). Results of surface sediment tests also demonstrate the spatial relationship between observed toxicities and terrestrial features of the Sparrows Point industrial complex, including the Tin Mill Canal drainage feature (Figs. 2 and 3). Results of core sediment tests provide useful information for delineating the area of historical impact within Bear Creek. Specifically, core results support a conclusion that toxicity observed at the surface (top 2 cm) does not adequately predict toxicity at depth. In seven of nine cores, surface toxicity differed from toxicity in the top 10 cm by 20 to 56%, in three instances overestimating toxicity (B2: +20%; F3: +56%; F4: +23%), and in four instances underestimating (C1: −28%; C2: −54%; B1: −54%; B3: −28%). Only in cores F1 and F2 in the immediate vicinity of the Tin Mill Canal discharge, where toxicity was nearly 100% over entire depths of sediment columns, were surface results consistent with subsurface results. This suggests management or mitigation decisions should not be based on results of surface sediment toxicity tests alone, and that an appreciation of toxicity and contaminant loads in sediments at

Table 2 Pearson product moment correlation coefficients for mortality and trace metal and persistent organic pollutant concentrations

	Coefficient (probability)	Mortality[a]	Zn	Cr	Cu	Ni	PAH_{pw}	PAH_{sed}
Zn	*r*	0.18						
	(*p* value)	0.434						
Cr	*r*	*0.56*	*0.68*					
	(*p* value)	*0.007*	*0.000*					
Cu	*r*	0.12	*0.82*	*0.57*				
	(*p* value)	0.600	*0.000*	*0.006*				
Ni	*r*	*0.51*	*0.66*	*0.91*	*0.55*			
	(*p* value)	*0.015*	*0.001*	*0.000*	*0.008*			
PAH_{pw}[b]	*r*	*0.46*	0.34	*0.57*	*0.62*	*0.43*		
	(*p* value)	*0.032*	0.108	*0.005*	*0.002*	*0.044*		
PAH_{sed}[c]	*r*	0.42	0.01	0.26	0.38	0.12	*0.79*	
	(*p* value)	0.051	0.955	0.312	0.083	0.581	*0.000*	
TPH_{sed}[d]	*r*	*0.64*	0.35	*0.82*	0.32	*0.75*	*0.55*	0.18
	(*p* value)	*0.001*	0.108	*0.000*	0.144	*0.000*	*0.008*	0.414

A positive Pearson coefficient indicates a positive correlation; the larger the value the stronger the correlation; values in italics are statistically significant ($p < 0.05$)

[a] Mortality data arcsine square root transformed

[b] PAH_{pw}: porewater polycyclic aromatic hydrocarbon (PAH)

[c] PAH_{sed}: total PAH in whole sediment

[d] TPH_{sed}: total petroleum hydrocarbon in whole sediment

depth is also important. This also raises concerns that redevelopment of the region could lead to resuspension and distribution of contaminated sediments at depth as a consequence of increased marine construction and boating activity.

In the present study, multiple contaminants were shown to correlate positively with toxicity, and most contaminant classes also correlated significantly with one another (Table 2). As previous studies show, finding a causative agent for toxicity in the Bear Creek system has proven difficult (Ashley and Baker 1999; Baker et al. 1997; Hansen et al. 1996; Klosterhaus et al. 2006; McGee et al. 1999, 2004). In actuality, numerous compounds likely contribute to observed effects on test organisms, a common phenomenon at sites with complex contaminant mixtures (Hagopian-Schlekat et al. 2001). Given the high degree of collinearity of measured contaminants, the present study did not attempt to elucidate the proportional contribution of particular chemicals to observed toxicity. Instead, contaminant analysis of surface sediments was directed toward determining the usefulness of PAH_{pw}, as measured by the antibody-based biosensor, to serve as a proxy for other contaminants within the Bear Creek system. PAH_{pw} was selected to measure in core segments because the biosensor required a small sample volume (≤10 mL) for analysis, and sediment material for chemical analysis was limited after removal of material for toxicity testing. The biosensor, a comparatively new technology, has been used effectively to produce near real-time, on-site, quantitative analysis of PAHs in the aqueous environment (Spier et al. 2011). A recent study by Li et al. (2016) compared measurements of PAHs in sediment porewater by gas chromatography-mass spectroscopy (GC-MS) and the antibody-based biosensor and demonstrated a very strong positive correlation between the methods ($r^2 = 0.98$; supplemental material Fig. S2). The biosensor typically yields higher total PAH values than does GC-MS analysis. This is because biosensor results reflect the cumulative contribution of all 3- to 5-ring PAHs, including analytes not targeted or below detection in GC analyses. While the majority of PAHs in sediments are sorbed to particles, the comparatively small fraction dissolved within porewater represents the portion most bioavailable to benthic organisms (ter Laak et al. 2006). Therefore, concentrations of PAHs within sediment porewater may more accurately predict PAH-related toxicity without the need to correct for the influences of grain size and organic carbon content on bioavailability (ter Laak et al. 2006).

Within Bear Creek surface sediments, PAH_{pw} showed significant positive correlations with mortality as well as with most other contaminants (Table 2), making it a good choice as a proxy for screening sediment contamination. At present, there is no sediment criteria for PAH_{pw}, so it is difficult to assess its role in observed toxicity. Even if PAH_{pw} was not a primary contributor to toxicity in Bear

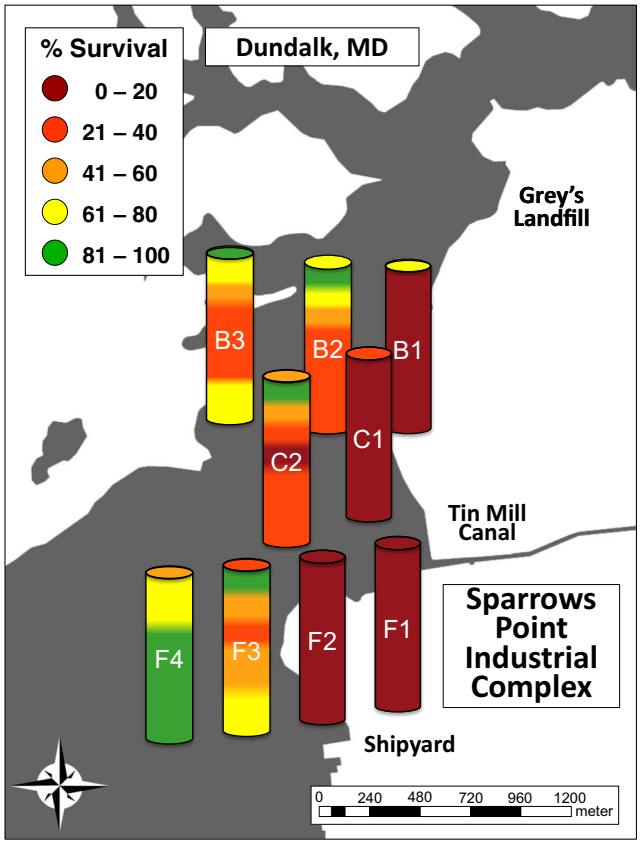

Fig. 5 Percent survival (*color-coded*) of *Leptocheirus plumulosus* following 10-day exposures to sediments from the surface (*top disk*) and at depth intervals (0–10, 10–20, 20–30, 30–40, 40–60, 60–80 cm) within sediment cores, and depth distribution of survival (*left*) and porewater polycyclic aromatic hydrocarbon (PAH_{pw}; *right*) within individual core segments from B, C, and F transects (PAH_{pw} of surface sediments are not shown)

Creek sediments, patterns within sediment cores show that where PAH_{pw} concentrations exceeded approximately 10 μg/L, toxicity was severe (Fig. S3; e.g., <40% survival). At the very least, PAH_{pw} results provide an indication that observed mortality at depth within sediment cores was likely related to increased contamination and not exclusively to other sediment characteristics below the biologically active layer. Although mortality generally increased with depth, high survival was also observed in the lowest levels of several cores (Fig. 5), further supporting the validity of core segment toxicity tests as a screening tool for assessing sediment quality.

It is understood that many other factors can confound sediment bioassay results, especially in subsurface sediments (Kinsman-Costello et al. 2015; Postma et al. 2002). Whole sediment characteristics, such as dissolved oxygen and pH that deviate from acceptable ranges, can have direct deleterious effects on test organisms, or can affect bioavailability and/or toxicity of pollutants targeted for study (Postma et al. 2002). In these instances, distinguishing natural from anthropogenic effects can be difficult. The most common compounds that confound sediment bioassays are dissolved sulfide and ammonia. Both accumulate in the porewater of anoxic sediments where lack of oxygen inhibits their transformation to less toxic oxidized forms (Kinsman-Costello et al. 2015). Moreover, pH influences dissolved sulfide and ammonia in contrasting ways, acidic conditions (pH < 6) favoring formation of more toxic unionized hydrogen sulfide (H_2S; Wang and Chapman 1999), whereas alkali conditions (pH > 8) favoring formation of more toxic unionized ammonia (NH_3; Emerson et al. 1975). These dissolved compounds can also transport to overlying water during bioassays, producing toxicity to test organisms even if they do not burrow into the sediment (Postma et al. 2002).

Porewater NH_3, which was measured in all core segments, increased with depth in most cores, in several far exceeding the recommended test threshold of 60 mg/L (supporting material Table S8). In particular, core F1 exhibited extremely high porewater NH_3 at depth. While segments from this core also had the highest PAH_{pw} concentrations, extraordinarily high NH_3 levels likely contributed substantially to the 100% lethality observed at all depths. Ammonia in other cores ranged from below the level of concern to well above that level. Cores from sites adjacent to the Sparrows Point industrial complex fell in this category (e.g., B1, C1, and F2). The

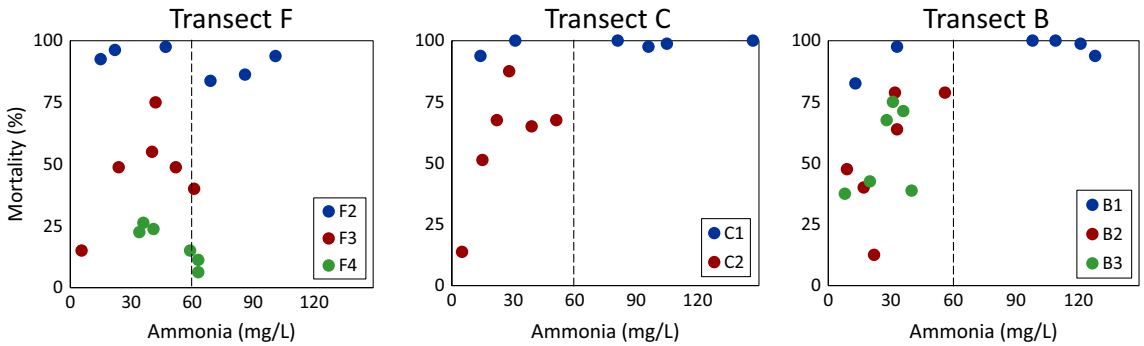

Fig. 6 Relationship between porewater ammonia and *Leptocheirus plumulosus* mortality in sediment core samples from Bear Creek, Baltimore MD, USA

consistent feature in these cores was very high mortality at all depths with no apparent dose-response relationship to NH_3 (Fig. 6). Despite porewater NH_3 levels ranging from <15 to >100 mg/L in sediments from these cores, all segments produced 90–100% lethality. Cores collected from locations more distant from the shore of the industrial complex (e.g., B2, B3, C2, F3, and F4) all demonstrated lower and more variable levels of toxicity. Ammonia in these cores only approached or marginally exceeded the 60 mg/L threshold at the lowest depths, and in only one instance (B2) did the highest porewater NH_3 concentration coincide with the greatest level of mortality. Conversely, many segments with NH_3 concentrations below 30 mg/L produced substantial mortality (Fig. 6). Collectively, these data suggest that NH_3 at shallow depths played, at most, a limited role in observed toxicity, and that other contaminants, including possibly sulfides, were more significant factors.

Neither sulfides nor pH were measured in Bear Creek sediments or porewater. However, the decrease in pH over time in overlying water during bioassay testing of several core segments indicates that sediment pH was very low. Under this condition, H_2S would be in abundance and would likely have contributed to observed toxicity. Whereas sulfides in sediments can be directly toxic to benthic organisms, they also serve to reduce metal toxicity by binding toxic metals to form less bioavailable metal-sulfide complexes (Wang and Chapman 1999). As such, the ratio or relative abundance of simultaneously extractable metals to acid volatile sulfide (SEM/AVS or SEM-AVS, respectively) is used as a predictor of metal-related sediment toxicity (Borgmann and Norwood 2002; Di Toro et al. 1991, 1992, 2005). Numerous studies indicate a more complicated relationship between AVS and metals availability that necessitates correction for organic carbon content (Di Toro et al. 2005). Others cite particle ingestion as an alternative route of metal exposure (Lee et al. 2000), or creation of oxic microenvironments by benthic macroinvertebrates as confounding factors when predicting toxicity based on assumptions of anoxic geochemistry (Campbell et al. 2006). In the present study, surface and core sediments were not analyzed for SEM/AVS. However, information from several previous assessments of Bear Creek sediments (near sites A1, A2, B3, D1, and the in-system reference) indicate AVS is consistently high (≥130 μmol/g) and that SEM/AVS ratios within the system are on the order of 0.10–0.28 (Table S2; Baker et al. 1997; Mason et al. 2004; McGee et al. 2004). Although not all of the sample locations in the present study were analyzed, conditions related to SEM/AVS were found to be reasonably consistent within depositional regions of Bear Creek, suggesting AVS predominates in sediments throughout the system. Sulfides also tend to increase with increasing depth below the oxic surface layer suggesting AVS would predominate in subsurface sediments as well (Wang and Chapman 1999).

Further chemical data, particularly on sulfide concentrations, are necessary to rule out the possibility of confounding factors causing observed toxicity in sediments at depth within cores. However, survival in deeper segments of some Bear Creek cores, and the spatial patterns observed that link toxicity at the surface and at depth to on-land features, suggest anthropogenic causes contribute to observed toxicity in Bear Creek sediments. Although uncommon, other studies have investigated toxicity at depth within sediments. Borgmann and Norwood (2002) sectioned sediment cores from Richard Lake near Sudbury, Ontario, and analyzed them for metal content, metal bioavailability, and toxicity to *Hyalella azteca*. They found that metal bioavailability varied with depth, and noted that toxicity was influenced by the inherent toxic content of the sediments, as well as by geochemical processes that changed contaminant availability (Borgmann and Norwood 2002). This study followed earlier work, including Rosiu et al. (1989) who measured toxicity in core segments from the Detroit River, Stemmer et al. (1990) who observed higher mortality to cladocerans exposed to deeper sediments, and Swartz et al. (1991) who observed greater toxicity of subsurface sediments to amphipods. Most studies of sediments at depth demonstrated toxicity correlating with contaminant concentrations, although Harkey et al. (1995) observed less PAH bioaccumulation in *Lumbriculus variegatus* in sediments with similar total PAH concentrations, indicating reduced PAH bioavailability in surficial sediments.

Despite more than two decades of study, establishing causal relationships between contaminant classes and toxic outcomes in Bear Creek sediments has proved elusive (Hansen et al. 1996; Klosterhaus et al. 2006; McGee et al. 1999, 2004). Metals, although exceeding ERMs in sediments from most sites, are presumed to be unavailable and therefore not contributing to observed toxicity because of the substantial abundance of AVS (United States Environmental Protection Agency 2005a). Likewise, PAHs and PCBs are below toxic levels based on bioavailability models (United States Environmental Protection Agency 2005b) in most surface sediments, despite substantial mortality in benthic test organisms (McGee et al. 2004). TPH exceeded screening guidelines in 17 of 22 samples and showed a strong correlation with amphipod mortality (Tables 1 and 2), so may have contributed to sediment toxicity. The suite of PAH compounds form a toxic subset of the larger TPH category, but other unidentified organic contaminants contained within the TPH category may be contributing to the observed toxicity (Agency for Toxic Substances and Disease Registry 1999). Although not measured in core samples, TPH had a strong correlation with PAH_{pw} in surface sediments, and this correlation may continue within sediments at depth.

Regardless of what contaminant or contaminants drive toxicity in Bear Creek, the most compelling pattern observed in the present study was the contradictory nature of surface

sediment toxicity compared with toxicity at depth. Areas with less toxicity at the surface may reflect decreases in contaminant inputs to the system, a result of recent reductions in manufacturing at the industrial complex. Mason et al. (2004) contend that environmental regulations contributed to the decrease in metals inputs to Bear Creek in recent decades. Studies from other regions demonstrate similar correspondence between periods of intensive industrial activity and contaminant patterns within sediment cores (Ingersoll et al. 2009; Wang et al. 2015). Vertical and horizontal patterns in sediment toxicity and contamination may also relate to bulk sediment transport dynamics within the system, driven by resuspension and distribution of surface sediments to more or less toxic neighboring areas. Alternatively, differences may reflect biologically mediated transformation processes associated with bioturbation of surface sediments by benthic invertebrates (Aller 1994; Amato et al. 2016), or microbial degradation in the case of petroleum hydrocarbons (Cappello et al. 2006). Such biological processes would be most active in surface sediments capable of supporting a diverse benthic and microbial community (Roling et al. 2002; Millward et al. 2004). Whatever mechanisms are at play in modifying surface sediment toxicity, they are unlikely to have meaningful effects below the biotic zone where bioturbation, biodegradation, and bulk sediment transport are slow or non-existent. Therefore, current subsurface contaminant conditions are likely to persist, and must be taken into account when planning for site redevelopment. The application of toxicity tests within sediment cores has proved useful in delineating this vertical reservoir of legacy contamination, and defining areas of impact more clearly than could have been accomplished through chemical analysis alone. Subsequent investigations should include measurements of pH, sulfides, and AVS/SEM to determine more clearly the role of anthropogenic contaminants in observed toxicity.

The present study, and others, highlights the importance of biological effects-based monitoring to inform environmental management decisions in such situations of uncertainty (Chapman 2002, 2016; Wang et al. 2015; Pourabadehei and Mulligan 2016). While analytical capabilities are ever improving, ecological and health-based screening criteria will always require site-specific adjustments to address unique confounding factors, particularly in cases of complex contaminant mixtures (Caeiro et al. 2009; Cesar et al. 2009; Mehler et al. 2011). In such scenarios, biological effects-based assessments of sediment toxicity are necessary tools to guide remediation and management decisions, especially at locations where complex and spatially variable contaminant mixtures make large-scale chemical assessment impractical. At Bear Creek and similar regions subject to redevelopment, progress may best be accomplished by applying effects-based assessments as screening tools to identify the extent of regions of concern.

Acknowledgements The authors thank Elizabeth Friedel, Allison Satterfield, Amy Wherry, Wenqi Hou, Alex MacLeod, and Barrett Wessel at the University of Maryland Department of Environmental Science and Technology (ENST) for laboratory and field assistance; Terry Wade of the Texas A&M University Geochemical and Environmental Research Group for analysis of persistent organic contaminants; and Andrew Heyes of the UMD Chesapeake Biological Laboratory for metals analysis. The authors recognize Mary Ann Vogelbein and George Vadas of the Virginia Institute of Marine Sciences (VIMS) for assistance with porewater PAH analyses, Daniel Fisher (ENST) for internal manuscript review, and Sacoby Wilson (University of Maryland School of Public Health) for input on public health aspects of Bear Creek research. This research was supported by an ENST graduate assistantship and by a grant from the Chesapeake Bay Foundation. Biosensor analysis of porewater samples at VIMS was supported by the NIEHS-SRP grant RO1ES024245.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

References

Agency for Toxic Substances and Disease Registry (1999) ToxFAQs™ for total petroleum hydrocarbons (TPH). US Department of Health and Human Services, Public Health Service https://www.atsdr.cdc.gov/toxfaqs/tf.asp?id=423&tid=75. Accessed 12 Feb 2017

Aller RC (1994) Bioturbation and remineralization of sedimentary organic matter: effects of redox oscillation. Chem Geol 114:331–345

Amato ED, Simpson SL, Remaili TM, Spadaro DA, Jarolimek CV, Jolley DF (2016) Assessing the effects of bioturbation on metal bioavailability in contaminated sediments by diffusive gradients in thin films (DGT). Environ Sci Technol 50:3055–3064

American Society for Testing and Materials (1994) Standard guide for collection, storage, characterization and manipulation of sediments for toxicological testing. ASTM annual book of standards, vol 11.04. Amer Soc Testing Materials, Philadelphia

Ashley JTF, Baker JE (1999) Hydrophobic organic contaminants in surficial sediments of Baltimore Harbor: inventories and sources. Environ Toxicol Chem 18:838–849

Baker JE, Mason RP, Cornwell JC, Ashley JTF, Halka J, Hill J (1997) Spatial mapping of sedimentary contaminants in the Baltimore Harbor/Patapsco River/Back River system. Final Report. Maryland Department of the Environment, Baltimore http://www.aquaticcommons.org/2883/1/97_142ocr.pdf. Accessed 12 Feb 2017

Borgmann U, Norwood WP (2002) Metal bioavailability and toxicity through a sediment core. Environ Pollut 116:159–168

Burton GA (2002) Sediment quality criteria in use around the world. Limnology 3:65–75

Caeiro S, Costa MH, DelValls A, Repolho R, Gonçalves M, Mosca A, Coimbra AP, Ramos TB, Painho M (2009) Ecological risk assessment of sediment management areas: application to Sado Estuary, Portugal. Ecotoxicology 18:1165–1175

Campbell PGC, Chapman PM, Hale BA (2006) Risk assessment of metals in the environment. Issues Environ Sci Technol 21:102–131

Cappello S, Caruso G, Zampino D, Monticelli LS, Maimone G, Denaro R, Tripodo B, Troussellier M, Yakimov M, Giuliano L (2006) Microbial community dynamics during assays of harbor oil spill bioremediation: a microscale simulation study. J Appl Microbiol 102:184–194

Cesar A, Abessa DMS, Pereira CDS, Santos AR, Fernández N, Choueri RB, Del Valls TA (2009) A simple approach to integrate the

ecotoxicological and chemical data for the establishment of environmental risk levels. Braz Arch Biol Technol 52:233–240

Chapman PM (1990) The sediment quality triad approach to determining pollution-induced degradation. Sci Total Environ 97/98:815–825

Chapman PM (2002) Integrating toxicology and benthic ecology: putting the eco back into ecotoxicology. Mar Pollut Bull 44:7–15

Chapman PM (2016) Environmental quality benchmarks—the good, the bad, and the ugly. Environ Sci Pollut Res. doi:10.1007/s11356-016-7924-2

Chen C, Chen C, Ju Y, Dong C (2016) Vertical profile, source apportionment, and toxicity of PAHs in sediment cores of a wharf near the coal-based steel refining industrial zone in Kaohsiung, Taiwan. Environ Sci Pollut Res 23:4786–4796

Chiou CT, McGroddy SE, Kile DE (1998) Partition characteristics of polycyclic aromatic hydrocarbons on soils and sediments. Environ Sci Technol 32:264–269

Di Toro DM, Mahony JD, Hansen DJ, Scott KJ, Carlson AR, Ankley GT (1992) Acid volatile sulfide predicts the acute toxicity of cadmium and nickel in sediments. Environ Sci Technol 26:96–101

Di Toro DM, McGrath JA, Hansen DJ, Berry WJ, Paquin PR, Mathew R, Wu KB, Santore RC (2005) Predicting sediment metal toxicity using a sediment biotic ligand model: methodology and initial application. Environ Toxicol Chem 24:2410–2427

Di Toro DM, Zarba CS, Hansen DJ, Berry WJ, Swartz RC, Cowan CE, Pavlou SP, Allen HE, Thomas NA, Paquin PR (1991) Technical basis for establishing sediment quality criteria for nonionic organic chemicals by using equilibrium partitioning. Environ Toxicol Chem 10:1541–1583

Emerson K, Russo RC, Lund RE, Thurston RV (1975) Aqueous ammonia equilibrium calculations: effect of pH and temperature. J Fish Res Board Can 32:2379–2383

Fleeger JW, Gust KA, Marlborough SJ, Tita G (2007) Mixtures of metals and polynuclear aromatic hydrocarbons elicit complex, nonadditive toxicological interactions in meiobenthic copepods. Environ Toxicol Chem 26:1677–1685

Forstner U (1989) Contaminated sediments: lectures on environmental aspects of particle-associated chemicals in aquatic systems. In: Bhattacharji S, Fridman GM, Neugebauer HJ, Seilacher A (eds) Lecture notes in earth sciences, vol 21. Springer-Verlag, Berlin, pp 1–157

Garcia CA, Passos EA, Alves JP (2011) Assessment of trace metals pollution in estuarine sediments using SEM-AVS and ERM-ERL predictions. Environ Monit Assess 181:385–397

Graham AM, Wadhawan AR, Bouwer EJ (2009) Chromium occurrence and speciation in Baltimore Harbor sediments and porewater, Baltimore, Maryland, USA. Environ Toxicol Chem 28:471–480

Hagopian-Schlekat T, Chandler GT, Shaw TJ (2001) Acute toxicity of five sediment-associated metals, individually and in a mixture, to the estuarine meiobenthic harpacticoid copepod *Amphiascus tenuiremis*. Mar Environ Res 51:247–264

Hansen DJ, Berry WJ, Mahony JD, Boothman WS, Di Toro DM, Robson DL, Ankley GT, Ma D, Yan Q, Pesch CE (1996) Predicting the toxicity of metal-contaminated field sediment using interstitial concentrations of metals and acid-volatile sulfide normalizations. Environ Toxicol Chem 15:2067–2079

Harkey GA, Van Hoof PL, Landrum PF (1995) Bioavailability of polycyclic aromatic hydrocarbons from a historically contaminated sediment core. Environ Toxicol Chem 14:1551–1560

Hartwell IS, Hameedi J (2007) Magnitude and extent of contaminated sediment and toxicity in Chesapeake Bay. NOAA technical memorandum NOS NCCOS, 47. NOAA/NOS/National Centers for Coastal Ocean Science/Center for Coastal Monitoring and Assessment, Silver Spring

Ingersoll CG, Kemble NE, Kunz JL, Brumbaugh WG, MacDonald DD, Smorong D (2009) Toxicity of sediment cores collected from the Ashtabula River in northeastern Ohio, USA, to the amphipod *Hyalella azteca*. Arch Environ Contam Toxicol 57:315–329

Inouye L (2014) DMMP issue paper: implementation of revised freshwater sediment screening values. Public review draft. Washington Department of Ecology, for the DMMP and RSET agencies, Lacey http://www.nws.usace.army.mil/Portals/27/docs/civilworks/dredging/Updates/2003-BCOC_with_alphaBHC_comments.pdf. Accessed 11 May 2017

Kinsman-Costello LE, O'Brien JM, Hamilton SK (2015) Natural stressors in uncontaminated sediments of shallow freshwaters: the prevalence of sulfide, ammonia, and reduced iron. Environ Toxicol Chem 34:467–479

Kirman DZ, Sericano JL, Wade TL, Bianchi TS, Marcantonio F, Kolker AS (2016) Composition and depth distribution of hydrocarbons in Barataria Bay marsh sediments after the Deepwater Horizon oil spill. Environ Pollut 214:101–113

Klosterhaus S, Baker J, Ziegler G, Fisher D (2006) Toxicity identification and evaluation and long-term contaminant trends in the Baltimore Harbor. Final Report. Technical and regulatory services administration. Maryland Department of the Environment, Baltimore

Kwok KW, Batley GE, Wenning RJ, Zhu L, Vangheluwe M, Lee S (2014) Sediment quality guidelines: challenges and opportunities for improving sediment management. Environ Sci Pollut Res Int 21:17–27

Lee B, Griscom SB, Lee J, Choi HJ, Koh C, Luoma SN, Fisher NS (2000) Influences of dietary uptake and reactive sulfides on metal bioavailability from aquatic sediments. Science 287:282–284

Li X, Kaattari SL, Vogelbein MA, Vadas GG, Unger MA (2016) A highly sensitive monoclonal antibody based biosensor for quantifying 3–5 ring polycyclic aromatic hydrocarbons (PAHs) in aqueous environmental samples. Sens Biosensing Res 7:115–120

Long ER, Morgan LG (1991) The potential for biological effects of sediment-sorbed contaminants tested in the National Status and Trends Program. NOAA technical memorandum NOS OMA 52. Second printing. Seattle, WA, USA. https://docs.lib.noaa.gov/noaa_documents/NOS/OMA/TM_NOS_OMA/nos_oma_52.pdf. Accessed 11 May 2017

Mason RP, Kim E, Cornwell J (2004) Metal accumulation in Baltimore Harbor: current and past inputs. Appl Geochem 19:1801–1825

McGee BL, Fisher DJ, Wright DA, Yonkos LT, Ziegler GP, Turley SD, Farrar JD, Moore DW, Bridges TS (2004) A field test and comparison of acute and chronic sediment toxicity tests with the estuarine amphipod *Leptocheirus plumulosus* in Chesapeake Bay, USA. Environ Toxicol Chem 23:1751–1761

McGee BL, Fisher DJ, Yonkos LT, Ziegler GP, Turley S (1999) Assessment of sediment contamination, acute toxicity, and population viability of the estuarine amphipod *Leptocheirus plumulosus* in Baltimore Harbor, Maryland, USA. Environ Toxicol Chem 18:2151–2160

Mehler WT, Li H, Lydy MJ, You J (2011) Identifying the causes of sediment associated toxicity in urban waterways of the Pearl River Delta, China. Environ Sci Technol 45:1812–1819

Millward RN, Carman KR, Fleeger JW, Gambrell RP (2004) Mixtures of metals and hydrocarbons elicit complex responses by a benthic invertebrate community. J Exp Mar Biol Ecol 310:115–130

Palma P, Ledo L, Soares S, Barbosa IR, Alvarenga P (2014) Integrated environmental assessment of freshwater sediments: a chemical and ecotoxicological approach at the Alqueva reservoir. Environ Geochem Health 36:209–223

Postma JF, De Valk S, Dubbeldam M, Maas JL, Tonkes M, Schipper CA, Kater BJ (2002) Confounding factors in bioassays with freshwater and marine organisms. Ecotoxicol Environ Saf 53:226–237

Pourabadehei M, Mulligan CN (2016) Selection of an appropriate management strategy for contaminated sediment: a case study at a shallow contaminated harbor in Quebec, Canada. Environ Pollut 219:846–857

Ridgway J, Shimmield G (2002) Estuaries as repositories of historical contamination and their impact on shelf seas. Estuar Coast Shelf Sci 55:903–928

Roling WFM, Milner MG, Jones DM, Lee K, Daniel F, Swannell RJP, Head IM (2002) Robust hydrocarbon degradation and dynamics of bacterial communities during nutrient-enhanced oil spill bioremediation. Appl Environ Microbiol 68:5537–5548

Rosiu CJ, Giesy JP, Kreis RG Jr (1989) Toxicity of vertical sediments in the Trenton Channel, Detroit River, Michigan, to *Chironomus tentans* (Insecta: Chironomidae). J Great Lakes Res 15:570–580

Spier CR, Vadas GG, Kaattari SL, Unger MA (2011) Near real-time, on-site, quantitative analysis of PAHs in the aqueous environment using an antibody-based biosensor. Environ Toxicol Chem 30:1557–1563

Stemmer BL, Burton GA, Sasson-Brickson G (1990) Effect of sediment spatial variance and collection method on cladoceran toxicity and indigenous microbial activity determinations. Environ Toxicol Chem 9:1035–1044

Swartz RC, Schults DW, Lamberson JO, Ozretich RJ, Stull JK (1991) Vertical profiles of toxicity, organic carbon, and chemical contaminants in sediment cores from the Palos Verdes Shelf and Santa Monica Bay, California. Mar Environ Res 31:215–225

ter Laak TL, Barendregt A, Hermens JLM (2006) Freely dissolved pore water concentrations and sorption coefficients of PAHs in spiked, aged, and field-contaminated coils. Environ Sci Technol 40:2184–2190

United States Environmental Protection Agency (1983) Methods of Chemical Analysis of Water and Waste (MCAWW). EPA/600/4-79/020. Office of Research and Development, Washington

United States Environmental Protection Agency (1994) Methods for assessing the toxicity of sediment-associated contaminants with estuarine and marine amphipods. EPA/600/R-94/025. Report. US Environmental Protection Agency, Office of Water, Washington

United States Environmental Protection Agency (1995) Method 3052: microwave assisted acid digestion of siliceous and organically based matrices. Test methods for evaluating solid waste. US Environmental Protection Agency, Washington

United States Environmental Protection Agency (2005a) Procedures for the derivation of equilibrium partitioning sediment benchmarks (ESBs) for the protection of benthic organisms: metal mixtures (cadmium, copper, lead, nickle, silver, zinc). EPA/600/R-02/011. US Environmental Protection Agency, Office of Research and Development, Washington

United States Environmental Protection Agency (2005b) Predicting toxicity to amphipods from sediment chemistry. EPA/600/R-04/030. US Environmental Protection Agency, Office of Research and Development, Washington

United States Environmental Protection Agency (2014) Method 6020B: inductively coupled plasma-mass spectrometry. Test methods for evaluating solid waste. US Environmental Protection Agency, Washington

Wang F, Chapman PM (1999) Biological implications of sulfide in sediment—a review focusing on sediment toxicity. Environ Toxicol Chem 18:2526–2532

Wang P, Zhang L, Liu L, Chen L, Gao H, Wu L (2015) Toxicity of sediment cores from Yangtze River estuary to zebrafish (*Danio rerio*) embryos. Environ Sci Pollut Res 22:16423–16433

Xu J, Guo J, Liu G, Shi G, Guo C, Zhang Y, Feng Y (2014) Historical trends of concentrations, source contributions and toxicities for PAHs in dated sediment cores from five lakes in western China. Sci Total Environ 470(471):519–526