SCIENTIFIC REPORTS
nature research

OPEN

DNA epigenetic marks are linked to embryo aberrations in amphipods

Elena Gorokhova[1], Giulia Martella[1], Nisha H. Motwani[1], Natalia Y. Tretyakova[2], Brita Sundelin[1] & Hitesh V. Motwani[1*]

Linking exposure to environmental stress factors with diseases is crucial for proposing preventive and regulatory actions. Upon exposure to anthropogenic chemicals, covalent modifications on the genome can drive developmental and reproductive disorders in wild populations, with subsequent effects on the population persistence. Hence, screening of chemical modifications on DNA can be used to provide information on the probability of such disorders in populations of concern. Using a high-resolution mass spectrometry methodology, we identified DNA nucleoside adducts in gravid females of the Baltic amphipods *Monoporeia affinis*, and linked the adduct profiles to the frequency of embryo malformations in the broods. Twenty-three putative nucleoside adducts were detected in the females and their embryos, and eight modifications were structurally identified using high-resolution accurate mass data. To identify which adducts were significantly associated with embryo malformations, partial least squares regression (PLSR) modelling was applied. The PLSR model yielded three adducts as the key predictors: methylation at two different positions of the DNA (5-methyl-2′-deoxycytidine and N^6-methyl-2′-deoxyadenosine) representing epigenetic marks, and a structurally unidentified nucleoside adduct. These adducts predicted the elevated frequency of the malformations with a high classification accuracy (84%). To the best of our knowledge, this is the first application of DNA adductomics for identification of contaminant-induced malformations in field-collected animals. The method can be adapted for a broad range of species and evolve as a new omics tool in environmental health assessment.

Deoxyribonucleic acid (DNA) in an organism can be structurally modified on its nucleobase moieties (G, C, A, and T) by both exogenous and endogenous factors. Covalent DNA adducts can be formed by chemicals that are electrophilic or form reactive metabolites[1–3]. Another pathway for modification of the DNA is through oxidation by reactive oxygen species (ROS) formed under oxidative stress[4,5]. If not repaired, such DNA lesions can interfere with the accuracy of DNA polymerases during replication which could lead to mutations. Alternatively, environmental exposure can cause epigenetic deregulation by altering the patterns of DNA methylation, affecting histone marks, and disrupting chromatin structure[6–9]. In this work we focused on the DNA methylated adducts, such as 5-methylcytosine and N^6-methyladenine. Unlike the genotoxic DNA adducts, epigenetic DNA modifications are introduced enzymatically and are involved in regulating the levels of gene expression. Broadly, exposure to environmental stressors may cause both genetic and epigenetic modifications, leading to growth and developmental pathologies in wild populations, which may, in turn, result in reproductive failure and population instability. However, studies linking genotoxic exposures to reproductive disorders are rare[10,11], and little is known about the relationships between specific types of DNA modifications and organism-level responses.

A comprehensive analysis of the modifications on the nucleobase moieties of DNA is addressed by DNA adductomics, an emerging field in systems toxicology. High-resolution mass spectrometry (HRMS) methods employing Orbitrap MS instrumentation have gained popularity in the adductomics analysis due to their ability to detect adducts with high mass accuracy in complex mixtures[12,13], and thus enabling more accurate elemental composition determination of the adducts. However, DNA adductomics being a relatively young research area, presents many challenges, including optimization of chromatographic separations, structural assignments for novel adducts, sampling design, and data interpretation.

Human activities over the past decades have aggravated the environmental pollution in many estuarine systems, including the Baltic Sea[14]. Responding to the need in sensitive and robust warning signals of exposure, there

[1]Department of Environmental Science and Analytical Chemistry, Stockholm University, Stockholm, SE, 10691, Sweden. [2]Department of Medicinal Chemistry and Masonic Cancer Center, University of Minnesota, Minneapolis, MN, 55455, United States. *email: hitesh.motwani@aces.su.se

has been an interest in the development of omics techniques and their application in ecotoxicology. Active search for stress biomarkers that are informative about the modes of action has contributed to the application of metabolomics in the environmental sciences[15]. In a similar manner, environmental adductomics can be used to study the effects of environmental stress – such as pollution and climate change – on the health of wild populations.

In the aquatic environments, physical adsorption causes hydrophobic contaminants to become highly enriched in sediments and porewater as compared to the water column[16]. This makes the benthic animals, such as amphipods, polychaetes, and clams, particularly vulnerable to continuous exposure to the sediment-bound compounds. Such organisms are thus useful as sentinels in environmental monitoring programs.

The amphipod *Monoporeia affinis* is a keystone species in soft bottoms of the Baltic Sea, where it often dominates benthic communities. Their great abundance, high quality as prey, and active feeding on organic material in the sediments make these animals an important trophic link in the transfer of sediment-associated contaminants to the benthic and pelagic food webs. In this semelparous amphipod, contaminant exposure can induce various reproductive disorders[17], including embryo aberrations[18]. Moreover, several aberration types observed in the amphipods were linked to specific contaminants such as polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs) and heavy metals in the ambient sediments, providing validation for application of the embryo aberration analysis in the environmental assessment[19,20]. Currently, embryo aberrations in the amphipods are used as an indicator for biological effects of contaminants within the Swedish National Marine Monitoring Program (SNMMP), and relative abundance of the aberrant embryos is a supporting indicator for Descriptor 8 (Hazardous substances) in the Marine Strategy Framework Directive (MSFD).

Towards the development of the screening methods that would facilitate detecting biological effects of environmental contaminants, we applied HRMS-based DNA adductomics in *M. affinis* collected in the northern Baltic Sea. We hypothesized that embryo aberrations in these amphipods are a manifestation of toxicity related to specific DNA modifications. To test this hypothesis, we evaluated whether certain modifications of the DNA occurring in both females and their broods were associated with high frequency of embryo aberrations.

Methods

Chemicals and other materials. Deoxyribonucleic acid from calf thymus (ctDNA) sodium salt, 2′-deoxyguanosine (dG), 2′-deoxycytidine (dC), 2′-deoxyadenosine (dA), thymidine (T), 5-methyl-2′-deoxycytidine (5-me-dC), 8-oxo-7, 8-dihydro-2′-deoxyguanosine (8-oxod-G), N^6-methyl-2′-deoxyadenosine (N^6-me-dA), nuclease P_1 from *Penicillium citrinum* (NP1), phosphodiesterase I from *Crotalus adamanteus* (snake) venom (SVPDE), alkaline phosphatase from *Escherichia coli* (AKP), ammonium acetate, ammonium bicarbonate, 2, 2, 6, 6-tetramethyilpiperidine-1-oxyl (Tris-buffer), zinc chloride and formic acid were obtained from Sigma-Aldrich (St. Louis, MO). Chelex-100 resin was purchased from Bio-Rad (Solna, Sweden). All solvents used were of HPLC grade. Experiments concerning DNA were carried out in DNA LoBind tubes, 1.5 mL (Eppendorf).

Survey area and amphipods. In the middle of January 2017, gravid *M. affinis* were collected at 30 stations along the Swedish coast, from the Quark in the north of the Bothnian Sea to Western Gotland Basin as a part of SNMMP yearly sampling programme (Supplementary Information, Fig. S1). These sampling sites are considered as representative habitats of *M. affinis* with regard to depth, sediment type and salinity[19,20]. Contamination levels with regard to PAHs, PCBs and trace and heavy metals in surface sediments at these sites are reported elsewhere[20], wherein total PAH and PCB concentrations in the sediments varied 13–1427 ng/g dwt (mean 496 ng/g dwt) and 1–54 pg/g dwt (18 pg/g dwt), respectively. The PAHs had medium high to high concentrations at most of the sites, whereas the PCB concentrations were relatively low at all sites. High to very high deviations of Cd, Cr, Cu, Ni and Zn were found compared to the pristine levels according to the analysis provided by Löf and co-workers[20].

M. affinis reproduces in late-November/early-December and carry their embryos in a brood pouch (marsupium) until offspring release in February-March. In the end of January, the embryos reach advanced developmental stage (stage 4–8; see full description of the embryo development by Sundelin and Eriksson, 1998[18]), making it possible to identify developmental aberrations. Gravid females were collected by a benthic sled[21], sorted, and transported in ambient water (temperature 4–7 °C) to the laboratory for embryo analysis.

Embryo aberration analysis. Eggs and embryos were removed from the marsupium of the female, and individual embryos were examined using a stereo-microscope (Leica M10 80× magnification) with polarized cold light as previously described[18] (Supplementary Information, Note S1). For each female, we recorded fecundity (eggs per female), embryo developmental stage, and number of aberrant embryos, if present (Table S1). The aberration types considered were malformed embryos, embryos with damaged membrane, embryos with arrested development and dead eggs, which are all incompatible with survival[18–20]. For both de-brooded females and their embryos, two groups were created: (*i*) healthy (aberration frequency $\leq$5%; 21 females and their broods), and (*ii*) unhealthy (aberration frequency 8–41%; 19 females and their broods). When assigning the test animals to these groups, the threshold for aberration frequency of 6%, i.e., percentage of malformed embryos of the total number of embryos in the brood, was applied based on the evaluation of the background frequency of embryo aberrations in this species[22].

For DNA adducts analysis, the de-brooded females and their embryos were analyzed separately; a single female sample and one brood sample originated from a single gravid female. A total of 40 gravid females were used generating 40 samples for de-brooded females and 40 samples for embryo broods. To have approximately equal number of females with relatively high and low levels of the embryo aberrations from the entire study area, the number of amphipods used per station varied 1 to 3. A care was taken to select a cohort of the amphipods with relatively synchronized brood development to minimize the variability of the adducts due to the embryo development. Hence, in the sample set, the mean embryo developmental stage was 5 ± 1 (mean and standard deviation), and there was no significant difference in the stage distribution between the healthy and unhealthy

groups (Student's t-test, p-value > 0.05). Each female was snap frozen in liquid nitrogen directly after the brood removal, and the dissected broods were treated in the same way directly after the embryo analysis. The samples were stored at −80 °C until the extraction of DNA.

DNA extraction. Individual samples were manually homogenized using a Kontec pestle, and genomic DNA was extracted with 350 µL of 6% Chelex-100 solution[23]. Briefly, the homogenate and chelex solution were incubated at 65 °C for 3 h and intermittently vortexed. After centrifugation at 14000 rpm for 10 min, the supernatant was transferred to an Eppendorf tube and stored at 4 °C overnight. DNA concentrations and purity were determined with a Nanophotometer™ (Implen); in all samples, A_{260}/A_{280} varied 1.7–1.8. The extracted DNA was then stored at −20 °C until sample preparation for liquid chromatography mass spectrometry (LC-MS) analysis.

Enzymatic digestion. Amphipod DNA (15 µg) extracted from the individual females and their broods was mixed with Tris-buffer (1 mM, pH 7.4) to arrive to a total volume of 300 µL. Ammonium acetate (0.1 M, 30 µL), zinc chloride (10 mM, 12 µL) and ammonium bicarbonate (1 M, 35 µL) were added along with the enzymes for digestion, which were NP1 (0.1 U/µL, 8 µL), SVPDE (0.000126 U/µL, 7 µL) and AKP (0.029 U/µL, 7 µL). The mixture was incubated at 37 °C for 2.5 h, followed by centrifugation at 14000 rpm (4 °C, 10 min). The supernatant containing the nucleosides was transferred to a septum sealed vial and analyzed by LC-MS. Further, blank samples were obtained with the procedure being repeated in triplicates but without any DNA.

Liquid chromatography conditions. The LC-MS system used consisted of a Dionex UltiMate 3000 LC device interfaced to an Orbitrap Q Exactive HF HRMS (Thermo Fisher Scientific, MA). The mobile phase for high-pressure liquid chromatography (HPLC) consisted of a mixture of water-methanol; system A with 5% methanol and system B with 95% methanol, each containing 0.1% formic acid. The HPLC column used was a Superlco Ascentis® Express F5 2.7 Micron HPLC column (15 cm × 2.1 mm) from Sigma-Aldrich. Sample injection volume was 20 µL. A flow rate of 120 µL/min was employed with the column temperature maintained at 25 °C. The LC gradient consisted of an initial 2 min equilibration at 5% of B, which was increased to 30% in 6 min and then to 100% in 2 min. After holding at this composition for 2 min, a ramping to the initial condition of 5% B was done in 1 min, and the system was re-equilibrated for 3 min before the next injection. An automated switch-valve was connected between the LC column and the Orbitrap MS, which was set to allow the eluent from the column to enter the waste during the first minute after injection to remove polar impurities in the samples. After this one minute, the valve was used to divert the eluent into the ion source of the mass spectrometer.

Orbitrap HRMS/MS analysis. HRMS analysis was carried out using an Orbitrap Q Exactive HF mass spectrometer equipped with a heated electrospray ionization (HESI) source. The HESI MS parameters were initially tuned for maximal signal intensity by infusion using a standard mixture prepared that contained dG, dA, dC and T (10 fmol/µL of each). The optimized MS parameters that were used for all further experiments were as follows. Spray voltage, 3.5 kV; spray current, 22 µA; capillary temperature, 275 °C; sheath gas, 20 arbitrary units (au); auxiliary gas, 10 au; S-Lens RF level, 60%; and probe heater temperature, 240 °C. The mass spectrometer was operated in the positive ion mode using normalized collision energy of 30 eV. Screening for the DNA modifications as nucleoside adducts was performed using full scan (FS)/data-independent acquisition (DIA) mode, wherein MS/MS fragmentation is performed on a sequential scan range. The full MS scanning was conducted with a mass resolution of 120000, automatic gain control (AGC) target $3e^6$, maximum ion injection time (IT) 200 ms, and scan range from 110 to 1200 m/z. The DIA was set to have resolution 60000, AGC target $5e^5$, maximum ion IT 120 ms, loop count 31, and scan range from 195 to 355 m/z, which was divided into 16 discrete m/z ranges with an isolation window of 10 m/z, in the form 200 ± 5, 210 ± 5, and up to 350 ± 5 m/z.

Data processing. The generated MS data was processed using XCalibur 3.1 (Thermo Fisher Scientific). Each mass range from the DIA mode was manually evaluated in the Qual browser by monitoring for the fragment ion with m/z 117.0552 ± 5 ppm, corresponding to the protonated deoxyribose ion, $[dR]^+$, to select the adduct candidates. The candidates were filtered as possible nucleoside adducts when two others ions, in addition to the $[dR]^+$, were detected in the form of parent ion $[M]^+$ and a specific fragment ion $[M\text{-}dR]^+$ (Fig. S2) within the same chromatographic peak, for which the following equation sustained within a 5 ppm mass tolerance for each ion; $[M]^+ = [M\text{-}dR]^+ + 116.0473$. The corresponding FS spectra were used for confirmation of the parent ions. This processing generated a list of 23 putative adducts, referred here as A1 to A23. Quantification was performed in XCalibur's Quan browser for the putative adducts (A1–A23) in all samples. Peak areas were measured using extracted ion chromatograms (EIC) of the adducts corresponding to m/z (±5 ppm) of the specific fragments which had higher abundance than the parent ions under the employed MS conditions. A normalized maximum intensity level (NL) of 1.0E3 was set as an apparent limit of quantification for individual peak areas. All adducts below this limit (A5 and A10 in females; A1, A5, A10, and A22 in embryo broods) were excluded from further statistical analysis. The measured peak areas were normalized to that of dG (adduct area × 10^2/dG area) from the same sample, followed by statistical analysis. Normalization to dG, which has also been successfully applied in other LC-MS based studies including those investigating changes in 5-me-dC levels[24,25], circumvents for loses during sample preparation and avoids the need for expensive isotope-labeled internal standards.

Identification of DNA modifications. Preliminary identification of each nucleoside adduct was based on the elemental composition obtained from the high mass accuracy data of the parent ions, and their comparison with the corresponding theoretical values calculated from ChemDraw Professional 16.0. The identification included calculating for presence of protonated and sodiated adducts. In addition to the four unmodified nucleosides, the identities of 5-me-dC, N^6-me-dA, and 8-oxo-dG were confirmed using authentic reference standards. Standard solutions of each reference compound (100 fmol/µL) were prepared in deionized water and analyzed by

LC-Orbitrap MS employing the same settings as described above. The resulting HRMS data, as well as retention time on the chromatogram, were compared with the corresponding data obtained from the amphipod samples. Spiking of amphipod samples with respective standard solution was performed to further confirm the identification of 5-me-dC, N^6-me-dA, and 8-oxo-dG.

Evaluation of DIA method for quantification. Adduct measurements based on the DIA method were evaluated for repeatability and variability using ctDNA, a commercially available DNA standard. Five replicates of ctDNA (15 µg) were dissolved in Tris-buffer (1 mM, pH 7.4, 300 µL) and enzymatically digested in the same way as the amphipod DNA. The resulting mixture was analyzed by LC-Orbitrap MS using the same DIA settings as described above. Peak areas were measured using EIC corresponding to the specific fragments, and normalized to dG, for five nucleoside adducts covering the selected mass range and including the identified adducts. These included parent ions having m/z 242.1146 (5-me-dC), 284.0983 (8-oxo-dG), 289.1758 (A7), 300.1299 (A12), 306.0603 (A17); N^6-me-dA was below quantification limit in the ctDNA samples and hence not selected in the evaluation. The standard deviation (SD) and coefficient of variation (CV) were calculated for each measured analyte from the ctDNA (Fig. S3).

Statistical analysis. The full dataset consisted of the putative nucleoside adduct profile in the female and their embryos and the health status as a group variable. Adducts A4, A13, A16, and A18 were excluded from the statistical modelling, since they were attributed as sodiated adducts and had a high correlation with their corresponding protonated forms. To evaluate differences in frequencies of specific adducts between the females and embryos and between healthy and unhealthy individuals, we applied the following statistical methods.

First, to compare the differences in the profiles of adducts between the females and their embryos, the data variability across the samples was visualized using non-dimensional multiple scaling (NMDS) based on Bray–Curtis distances. Permutational Multivariate Analysis of Variance using distance matrices, PERMANOVA, with 1000 Monte Carlo permutations was performed based on Bray–Curtis distances[26] and using software PRIMER+ v.6 (Primer-e, Ivybridge, UK).

Second, to produce a predictive model for the amphipod health status using adducts, we combined projection to latent structures using partial least squares discriminant analysis (PLS-DA) and a follow-up logistic regression. Whereas the standard logistic regression model has difficulty handling a large number of variables, PLS-DA can be constructed using all adducts, without any prior variable selection step. The standard logistic regression model is then constructed using the selected adducts identified by PLS-DA[27] as the most influential. Therefore, PLS-DA was applied on both females and embryos to search for discriminative adducts that contributed to the separation between the adduct profiles (DNA modifications across an individual) belonging to different groups (healthy vs. unhealthy) as implemented in MetaboAnalyst[28]. The influential predictors were selected based on their variable importance in projection (VIP) values with a cutoff of 0.8[29]. The predictors were then used in the logistic model with the health status as a binomial response variable, binomial error structure, log-link function, and the area under the receiver operating characteristic (ROC) curve (AUC) score as the selection criterion. Odds ratios of the logistic regression were used as a measure of associations between the treatment and the predictors. The sensitivity (i.e., true positives, or the proportion of cases correctly identified by the model as being unaffected by stress) and the specificity (e.g., true negatives, the proportion of cases correctly identified by the model as experiencing stress) were extracted from the ROC curves.

Results

Detection of DNA modifications using high-resolution mass spectrometry. To screen for DNA modifications, released as nucleoside adducts in the enzymatic digests, we took advantage of the characteristic dR moiety in the LC-HRMS analysis on the Orbitrap MS system. Twenty-three putative nucleoside adducts (abbreviated as A1 to A23; Table 1) were detected in the amphipods. The detection was based on fragmentation of the nucleoside adducts, using electrospray ionization in positive ion mode, to m/z 117.0552 (corresponding to protonated dR) and a specific fragment for each nucleoside adduct corresponding to nucleobases with structural modifications at the purine or pyrimidine ring (Fig. S2). High mass accuracy within a 5 ppm range was employed, facilitating structural identification of the unknown adducts based on their elemental composition.

The amount of the digested DNA that was injected on-column was kept constant (0.75 µg), which was judged to be sufficient for adduct detection. No nucleosides were detected in the blank samples prepared by the same procedure but in the absence of DNA. Representative chromatograms corresponding to adducts detected in the amphipod samples using LC-HRMS/MS analysis are shown in Fig. 1. The figure also highlights the advantage of using HRMS for resolving nucleoside adducts from the complex set of ions present in the samples (Fig. 1C), as high background from sample matrix was observed when using low mass resolution settings (Fig. 1B).

Structural identification of DNA modifications. Structural characterization of each nucleoside adduct was based on the molecular weight information, characteristic fragment ions containing modified free base (Fig. S2), and HPLC retention time, which were compared to those of authentic standards (Table 1). High resolution accurate mass (HRAM) measurements on the Orbitrap MS allowed for determination of possible elemental composition of the unknown adducts. Out of the 23 detected nucleoside adducts, 8 were structurally identified. Figure 2 shows the structures of nucleoside adducts that were identified in the amphipod DNA, along with their elemental composition and calculated masses. The mass difference between data recorded for the standards (5-me-dC, N^6-me-dA, and 8-oxo-dG) and the corresponding analytes from amphipod samples was <1 ppm. When the digested DNA samples were spiked with standard solutions of 5-me-dC, N^6-me-dA, and 8-oxo-dG, complete co-elution of the analyte peaks was observed before and after spiking, confirming their identification in the amphipod DNA (Fig. S4).

Detected nucleosides	Molecular ion, m/z $[M]^+$	Specific fragment, m/z $[M\text{-}dR]^+$	Retention time, min	Identified nucleosides
A1	295.1203	179.0739	3.51	
A2	278.1605	162.1126	3.59	
dC	228.0978	112.0507	3.81	2′-Deoxycytidine, dC[a]
A3	242.1146	126.0664	3.97	5-Methyl-2′-deoxycytidine, 5-me-dC[a]
A4	264.0961	148.0484	3.97	Na adduct of 5-me-dC[b]
A5	244.0932	128.0457	4.48	5-Hydroxy-2′-deoxycytidine, 5-OH-dC
A6	229.0818	113.0348	5.49	2′-Deoxyuridine, dU
A7	289.1758	173.1288	5.60	
A8	268.1043	152.0569	6.41	8-Hydroxy-2′-deoxyadenosine, 8-OH-dA
dA	252.1094	136.0620	6.41	2′-Deoxyadenosine, dA[a]
A9	239.1139	123.0667	6.41	
A10	274.0915	158.0440	6.42	Na adduct of dA[b]
A11	284.1358	168.0883	6.43	
A12	300.1299	184.0832	6.43	
A13	275.0753	159.0280	7.29	Na adduct of dI[b]
A14	253.0931	137.0460	7.30	2′-Deoxyinosine, dI
dG	268.1033	152.0568	7.51	2′-Deoxyguanosine, dG[a]
A15	274.1148	158.0676	7.51	Guanidinohydantoin, Gh
A16	290.0852	174.0389	7.52	Na adduct of dG[b]
A17	306.0603	190.0129	7.61	
T	243.0970	127.0504	7.99	Thymidine, T[a]
A18	265.0798	149.0324	7.99	Na adduct of T[b]
A19	327.0504	211.0029	8.03	
A20	236.1281	120.0809	8.48	
A21	284.0989	168.0518	8.75	8-Oxo-7, 8-dihydro-2′-deoxyguanosine, 8-oxo-dG[a]
A22	266.1257	150.0777	9.42	N^6-Methyl-2′-deoxyadenosine, N^6-me-dA[a]
A23	252.1228	136.0759	11.34	

Table 1. Characterization of the detected nucleoside adducts (A1–A23), and the 4 unmodified nucleosides, using a HRAM adductomics approach. Observed mass-to-charge ratio of molecular ion and that of its specific fragment with loss of dR, and retention time under the employed chromatographic conditions are given for each detected nucleoside, along with names of those identified. For the identified compounds, the mass difference between that observed (reported in this table) and calculated m/z was less than 3 ppm. Empty space indicates adduct not identified. [a]Identification confirmed by comparison with respective standards. [b]These were assigned as sodiated adducts of the respective compounds (based on retention time and molecular masses), with molecular ion and specific fragment represented as $[M+Na]^+$ and $[M\text{-}dR+Na]^+$, respectively. All others are assigned as protonated adducts with molecular ion and specific fragment represented as $[M+H]^+$ and $[M\text{-}dR+H]^+$, respectively.

Quantitative evaluation of DNA modifications in females and embryos. When all samples were considered, the composition and abundance of nucleoside adducts were largely similar between the females and the embryos for a majority of the modifications (Figs. S5 and S6). The frequent adducts were distributed similarly across the samples, with A11 and A20 being the most commonly detected nucleoside adducts in both females and embryos (Fig. 3). However, differences in the detection frequencies between the females and embryos were also found for some specific adducts. For A15, A20, and A23, the detection frequencies correlated significantly positively between the females and their embryos (Table S2). Moreover, significant correlations were observed between the same adduct types in both females and embryos (Table S3), such as between A7 and A20 [$r = 0.86$ (females), 0.74 (embryos)], and that between A9 and A11 [$r = 0.88$ (females), 0.75 (embryos)]. The NMDS plot of the Bray–Curtis distances confirmed that samples clustered primarily in a source-dependent manner along the NMDS2 axis, with variability between the broods (embryo samples) being significantly higher compared to that between the individual females (Figs. 4 and S7). Both NMDS stress value and ANOSIM output indicated significant differences in the nucleoside adduct profile between the females and their embryos (Fig. 4).

DNA modifications as predictors of amphipod health status. PLS-DA using normalized peak areas for nucleoside adducts as predictors and amphipod health status (healthy vs. unhealthy) successfully discriminated between the females carrying high number of malformed embryos (8–41%; unhealthy) and those with the background levels of malformations ($\leq$5%; healthy) (Fig. 5A); the between-group difference was validated by the permutation test ($p < 0.013$; Fig. 5B). Three component matrices accounted for 32.1, 15.2, and 5.9% of the total variance (Fig. 5A). Top contributors to the PLS-DA included A22 (N^6-me-dA), A9, A3 (5-me-dC), A12, and A1 in the females and A15 (Gh) and A7 in the embryos. Individually, among the female-associated adducts, A22, A9, and A3 were the most elevated, while A12 and A1 were the most depleted adducts in the unhealthy individuals. Among the embryo-associated adducts, A7 was elevated while A15 was depleted in the unhealthy individuals (Fig. 5C). These adducts were the greatest contributors to the PLS-DA components 1–3 based on their VIP scores (Fig. 5C).

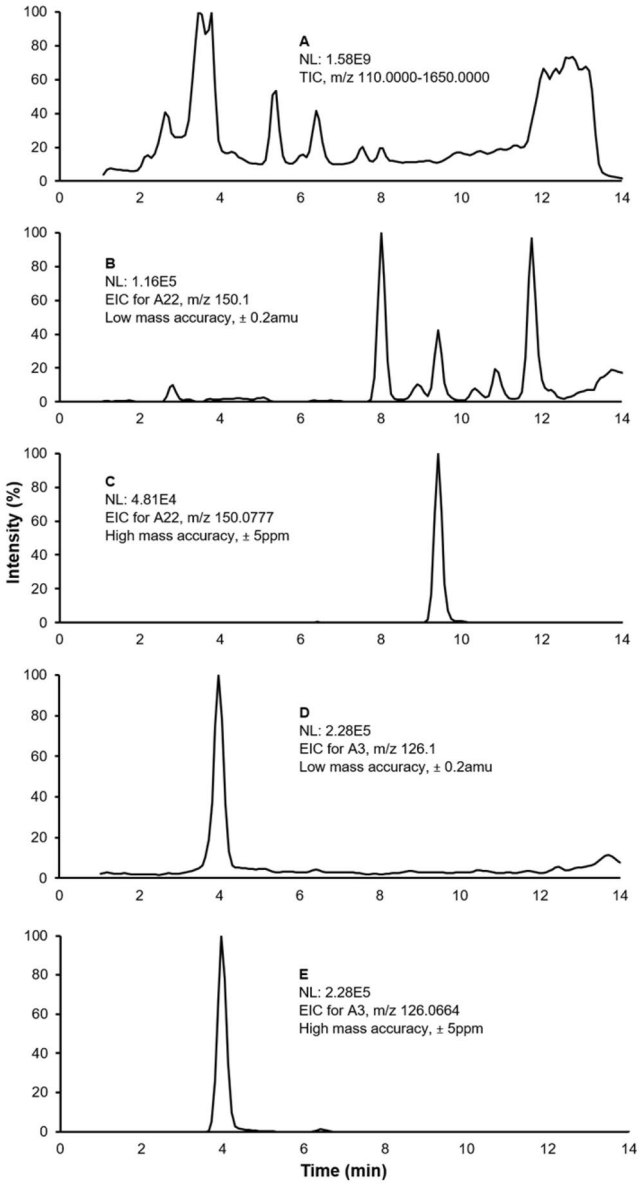

Figure 1. Representative chromatograms from an amphipod showing differences in high and low mass accuracy for resolving of nucleoside adducts. (**A**) Total ion chromatogram (TIC) of the full scan data showing high background signal from the sample matrix, which does not allow for detection of the adducts; NL, normalized maximum ion intensity level. (**B**) Extracted ion chromatogram (EIC) corresponding to $[M\text{-}dR + H]^+$ for A22 (N^6-me-dA) at a mass tolerance typical of quadrupole instrument (± 0.2 amu). The peak corresponding to A22 is not clearly distinguishable. (**C**) EIC corresponding to $[M\text{-}dR + H]^+$ for A22 at a mass tolerance of 5 ppm. The peak corresponding to the adduct is clearly resolved. (**D**,**E**) Similar to (**B**,**C**), respectively, but for A3 (5-me-dC). The peak is resolved in both panes.

A logistic model also discriminated between healthy and unhealthy females with high classification accuracy (Table S4). Using this approach, the area under the receiver operating characteristic (ROC) curve (AUC) identified A22 (N^6-me-dA), A9, and A3 (5-me-dC) as having the greatest specificity and sensitivity for distinguishing between the reproductively unhealthy and healthy females (Fig. S8). For all the three adducts, the increased levels were associated with high frequency of embryo aberrations. None of the adducts in the embryos that were identified as significant predictors for PLS-DA were included in the winning logistic model.

Discussion

Despite the biological plausibility of causal relationships between exposure to environmental contaminants, reproductive abnormalities, and DNA modifications, our knowledge of these linkages is sparse, particularly for non-model species and in wild populations. The evidence is accumulating that environmental chemicals that exhibit reproductive toxicity do so, at least in part, through genotoxic mechanisms and epigenetic alterations

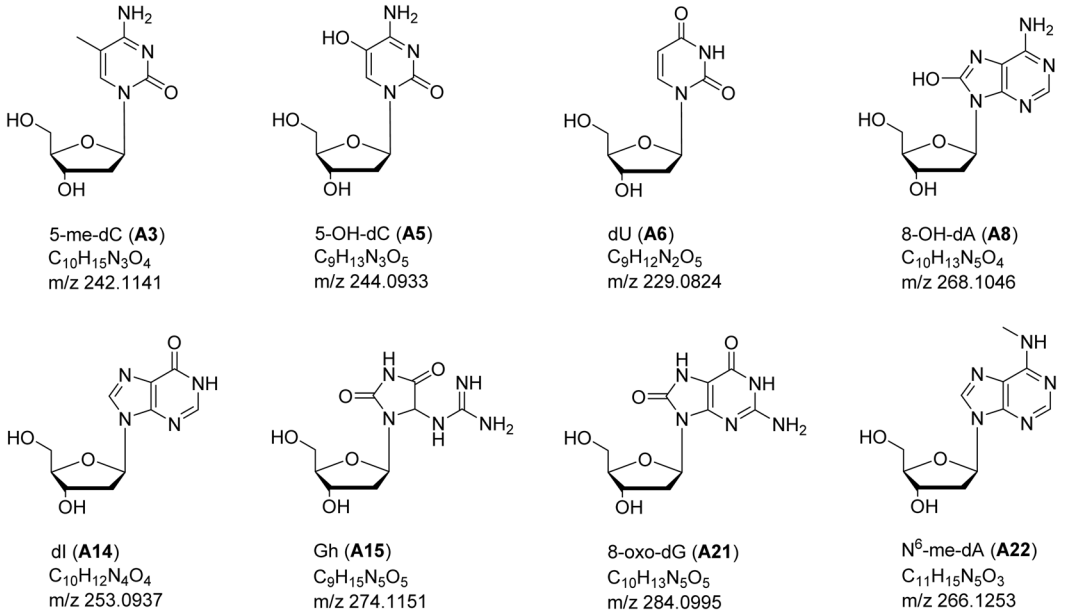

Figure 2. Chemical structure, elemental composition and calculated m/z $[M+H]^+$ of nucleoside adducts identified in the amphipods. The identification of A3, A21 and A22 were confirmed by comparison with respective standards; the proposed structures of the others are based on the HRAM data.

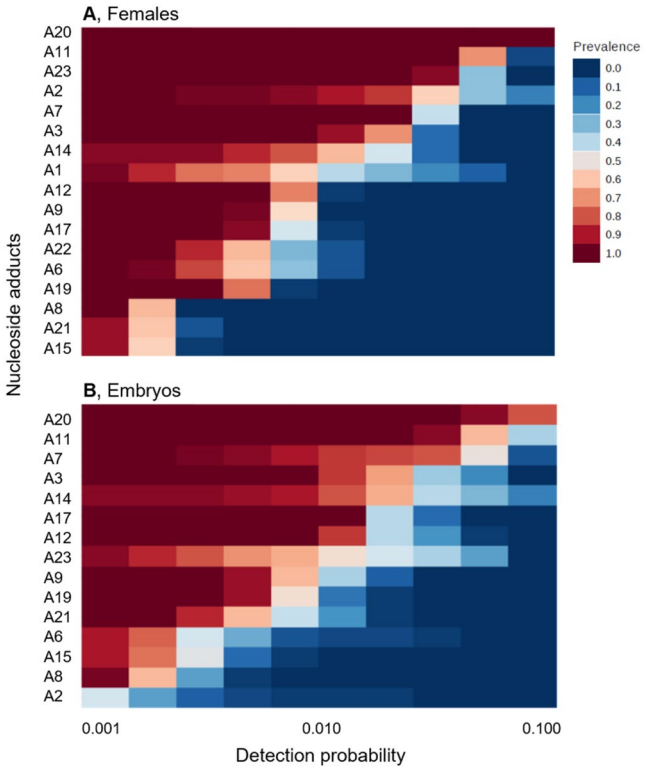

Figure 3. Profiles of nucleoside adducts in the females (**A**) and their embryos (**B**) expressed as their detection probability in 40 individual samples. See Table 1 for additional information on specific adducts.

in genomic DNA[10,30]. Using an untargeted adductomics approach with HRMS, we found that the frequency of embryo aberrations in Baltic Sea amphipods is related to the levels of certain DNA modifications in females and their embryos. To the best of our knowledge, this is the first study demonstrating the utility of DNA adductomics in wild populations.

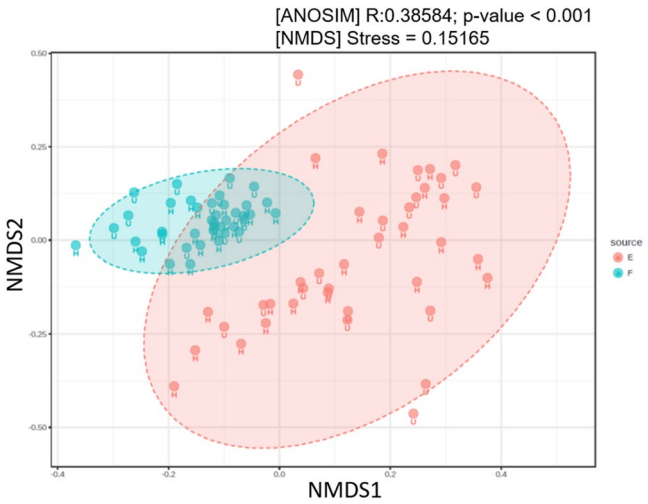

Figure 4. NMDS ordination diagram based on the nucleoside adducts in the females and their embryos with healthy (H) and unhealthy (U) data combined. Analysis of group Similarities (ANOSIM) results on the amount of variation in the adduct data attributable to the source (females vs. embryos) are based on Bray-Curtis dissimilarity index. See Fig. S7 for the evaluation of the samples grouped according to their health status.

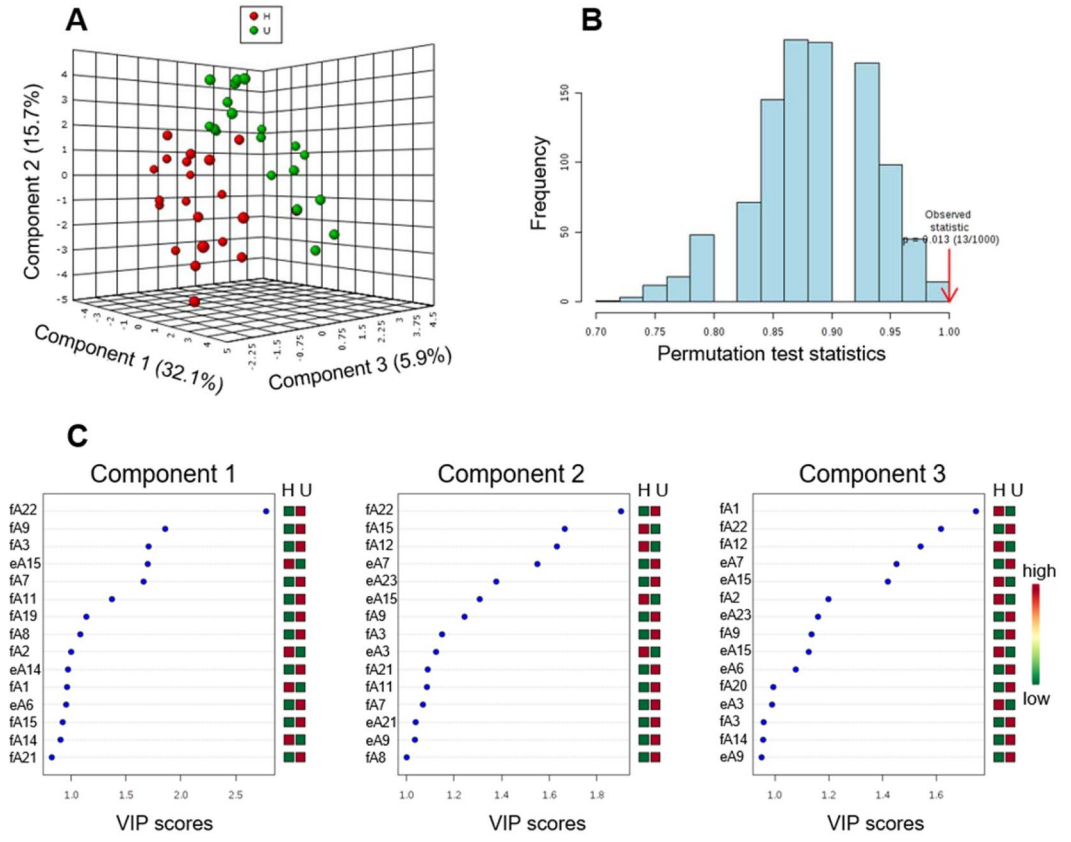

Figure 5. PLS-DA of nucleoside adducts in healthy (≤5% embryo aberrations) and unhealthy (8–41%) females and embryos. (**A**) 3-Dimensional score plot of PLS-DA using components 1, 2, and 3, accounting for 32.1, 15.7, and 5.9% of the total variance; (**B**) Validation of PLS-DA by permutation test ($p < 0.013$); and (**C**) Variable importance in projection (VIP) scores of 15 top contributors to PLS-DA, components 1–3; note the differences in the scale of the x-axis between the panels. The adducts (Table 1) measured in females and embryos are denoted as fA# and eA#, respectively.

Methodological considerations. LC-MS/MS has become the preferred method for measuring DNA adducts due to the high selectivity obtained in comparison to, e.g., ^{32}P-postlabeling technique[31–33]. Moreover, recent advances in MS instrumentation providing HRAM data have resulted in gaining more information on the chemical structure of the adduct based on its fragmentation pattern and more accurate elemental composition determination[13,34]. We employed HRMS in DIA mode that permitted stepwise acquisition for the detection of adducts and is useful for its archival nature as samples that have been analyzed earlier can be re-examined for new adducts. Even with the fragmentation data being limited to MS/MS, DIA approach has made significant progress in proteomics[35,36] and metabolomics[37,38]. Development of bioinformatics tools related to adduct analysis would further enhance the use of this method in adductomics.

In the detection strategy, we utilized the common dR fragment present on nucleosides. This analytical approach did not allow us to capture rapidly depurinating adducts[39] and phosphate-adducts formed on the phosphodiester backbone of DNA[40]. Further, bulky adducts such as those formed from covalent binding of nucleophilic sites on DNA to reactive metabolites of benzo[a]pyrene and other PAHs, having high molecular mass (>350 g/mol), were not targeted in this work. These adducts require metabolic biotransformation of the precursor PAH[41], the ability of which remains to be explored for amphipods and other relevant species.

A relatively large amounts of DNA from individual amphipods and the whole broods were obtained for this work; 30 to 105 µg/female and 15 to 95 µg/brood. We did not attempt linking the nucleoside adducts to a specific type of embryo malformation, because the DNA was extracted from the whole broods comprised of 13 to 58 embryos, often with mixed malformation types within a brood. For evaluating whether specific DNA modifications occur in an embryo with a certain malformation type, the samples have to be prepared using fewer embryos carrying only the malformation type in question. This would require a higher sensitivity for detection of adducts and, for instance, could be done by using nano-LC flow condition utilizing the inverse relationship between flow rate and electrospray sensitivity[42].

Relationships between reproductive pathology and DNA modifications. There were significant differences in the nucleoside adduct profile between the amphipods with high and low frequencies of the abnormal embryos in the brood; we identified the adducts contributing most to these differences. Thus, the hypothesized relationship between the developmental pathologies in the amphipods and the occurrence of DNA modifications was supported. The occurrence of the aberrant embryos was positively associated with increased levels of 5-me-dC (A3) and N^6-me-dA (A22), which are the epigenetic DNA modifications that are introduced enzymatically and have a regulatory role in cell functioning. We suggest that environmental exposure lead to epigenetic deregulation, which in turn caused developmental toxicity in the animals.

Amongst the 23 putative adducts detected, the chemical structure of 8 nucleoside modifications was proposed (Fig. 2) based on the elemental composition determined from the high mass accuracy, and the identity of 3 adducts was confirmed by comparison with commercially obtained standards. Significantly, the two epigenetic marks in the female DNA, 5-me-dC (A3) and N^6-me-dA (A22), were identified in this work as predictors for high malformation rate. DNA methylation is a characteristic feature for control of gene expression and maintaining genomic stability. This type of epigenetic alteration has been implicated in numerous biological processes, such as cell differentiation, X-chromosome inactivation, embryogenesis and tumorigenesis[9,43–45]. Many environmental contaminants in the sediments, including metals (e.g., Cd, Zn, and Hg), PCBs and PAHs[19,20], might interfere with DNA methylation status. Specifically, 5-me-dC has commonly been used as an epigenetic marker in evaluating the association of environmental chemicals to adverse health outcomes[7,46,47]. The N^6-me-dA adduct is comparatively less studied, and its biological role is still unclear. However, the recent identification of N^6-me-dA in frog *Xenopus laevis*, mouse *Mus musculus* and human tissues[48], and in mouse embryonic stem cells[49], suggest that this type of DNA methylation is more prevalent in the eukaryotes than previously thought. An unidentified DNA adduct (A9), having *m/z* 239.1139 of the nucleoside adduct, was the most significant predictor of the health status. This mass does not seem to correspond to any methylated derivate of the nucleosides, suggesting another pathway contributing to the pathological development.

Also, oxidative adducts 8-oxo-dG (A21), 5-OH-dC (A5), 8-OH-dA (A8) and Gh (A15), and nitrosative deamination adducts dU (A6) and dI (A14) were detected, which might be formed due to presence of reactive oxygen and nitrogen species, respectively. Such type of DNA damage arising from oxidative and nitrosative stresses have been mechanistically linked to inflammation[50,51]. In particular, chromosomal aberrations and mutations, particularly GC to TA transversions, have been linked to the formation of 8-oxo-dG[5,52]. The oxidized guanine might also be associated with detrimental effects on cell function, such as microsatellite instability and acceleration of telomere shortening[4]. However, no significant correlation was observed between the oxidative- or nitrosative-adducts and embryo damage.

Ontogenetic variability in DNA adduct profile. We observed significant ontogenetic differences in the nucleoside adduct profiles. Comparison of the adduct profiles between the females and their embryos showed that some adducts (A15, A20, and A23) correlated significantly positively between the mothers and their offspring. However, no such correlations were observed for the adducts that were identified as significant predictors for the high frequency of the aberrant embryos (Table S2). The overall variability of the adduct profiles was significantly higher among the broods than among the females (Fig. 4), which was most likely related to both the frequency of the aberrations and the development-stage variability of the embryos. In this study, a care was taken to minimize the variability in the DNA adduct profile due to activation/deactivation of numerous pathways involved in normal embryogenesis by targeting a relatively narrow span of the embryo development (stage 4–8) when selecting the gravid females for the adduct analysis. Nevertheless, it is possible that organogenesis that occurs during the ontogenetic development during these stages was sufficient to induce the high variability in the relative distribution of the adducts in the embryo samples. It is also possible that embryonic cells are repair-efficient, and only a small number of DNA lesions remain unrepaired. As a result, none of those adducts that were detected in

the embryos were retained in the predictive model that discriminated between the females carrying low and high proportions of the aberrant embryos.

Future directions. The validation of DNA modifications for use as exposure biomarkers in environmental status assessment requires establishing causal relationships with pathological conditions, such as tumorigenesis or developmental aberrations. The latter was addressed in our study using field-collected animals, but more controlled experimental data relating profiles of DNA modifications, oxidative balance, and reproductive pathologies are needed. In the field, biological effects of contaminants typically result from exposure to complex mixtures of environmental contaminants with different bioavailability and a myriad of confounding factors, such as embryo developmental stage, temperature, oxygenation, and feeding conditions. All these factors as well as variables, such as adequate nutrition and possible parasite infestation, may affect rates of adduct formation and DNA repair. Assessing the role of embryo stage and other biotic and abiotic confounding factors in the laboratory under controlled conditions is thus necessary in future studies as a part of the validation. Further, standardizing exposure conditions would facilitate identification of unique, measurable nucleoside adducts indicative of exposure, elucidate which chemicals have an impact on these adducts and provide information on the dose-response relationships. However, despite current knowledge gaps, DNA adductomics represents a promising approach for application in risk assessment and monitoring surveys. When more extensive data become available, they might be used to establish effective biological dose and predict the potential for irreversible toxicity, confirm suspected exposures and improve assessment of human impacts on the quality status of the marine environment.

Data availability

All data is available in the main article or in the Supplementary Information.

Received: 14 April 2019; Accepted: 26 December 2019;
Published online: 20 January 2020

References

1. Hemminki, K., Koskinen, M., Rajaniemi, H. & Zhao, C. Y. DNA adducts, mutations, and cancer 2000. *Regul Toxicol Pharm* **32**, 264–275 (2000).
2. La, D. K. & Swenberg, J. A. DNA adducts: Biological markers of exposure and potential applications to risk assessment. *Mutat Res* **365**, 129–146 (1996).
3. Miller, E. C. & Miller, J. A. Mechanisms of chemical carcinogenesis. *Cancer* **47**, 1055–1064 (1981).
4. Evans, M. D. & Cooke, M. S. Factors contributing to the outcome of oxidative damage to nucleic acids. *Bioessays* **26**, 533–542 (2004).
5. Grollman, A. P. & Moriya, M. Mutagenesis by 8-Oxoguanine - an Enemy Within. *Trends Genet* **9**, 246–249 (1993).
6. Choi, J. D. & Lee, J. S. Interplay between Epigenetics and Genetics in Cancer. *Genomics Inform* **11**, 164–173 (2013).
7. Stein, R. A. Epigenetics and environmental exposures. *J Epidemiol Commun H* **66**, 8–13 (2012).
8. Vandegehuchte, M. B. & Janssen, C. R. Epigenetics in an ecotoxicological context. *Mutat Res* **764**, 36–45 (2014).
9. Jackson-Grusby, L. *et al*. Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. *Nat Genet* **27**, 31–39 (2001).
10. Anderson, S. L. & Wild, G. C. Linking genotoxic responses and reproductive success in ecotoxicology. *Environ Health Perspect* **102 S**, 9–12 (1994).
11. Pampanin, D. M. *et al*. DNA adducts in marine fish as biological marker of genotoxicity in environmental monitoring: The way forward. *Mar Environ Res* **125**, 49–62 (2017).
12. Hemeryck, L. Y., Decloedt, A. I., Vanden Bussche, J., Geboes, K. P. & Vanhaecke, L. High resolution mass spectrometry based profiling of diet-related deoxyribonucleic acid adducts. *Anal Chim Acta* **892**, 123–131 (2015).
13. Balbo, S., Hecht, S. S., Upadhyaya, P. & Villalta, P. W. Application of a high-resolution mass-spectrometry-based DNA adductomics approach for identification of DNA adducts in complex mixtures. *Anal Chem* **86**, 1744–1752 (2014).
14. HELCOM Baltic Sea Environ. Proc. No. 120B. Hazardous substances in the Baltic Sea – An integrated thematic assessment of hazardous substances in the Baltic Sea. (Helsinki, Finland, 2010).
15. Bundy, J. G., Davey, M. P. & Viant, M. R. Environmental metabolomics: a critical review and future perspectives. *Metabolomics* **5**, 3–21 (2009).
16. Forstner, U. Sediment-Associated Contaminants - an Overview of Scientific Bases for Developing Remedial Options. *Hydrobiologia* **149**, 221–246 (1987).
17. Jacobson, T., Sundelin, B., Yang, G. & Ford, A. T. Low dose TBT exposure decreases amphipod immunocompetence and reproductive fitness. *Aquat Toxicol* **101**, 72–77 (2011).
18. Sundelin, B. & Eriksson, A. K. Malformations in embryos of the deposit-feeding amphipod Monoporeia affinis in the Baltic Sea. *Mar Ecol Prog Ser* **171**, 165–180 (1998).
19. Löf, M. *et al*. Biomarker-enhanced assessment of reproductive disorders in Monoporeia affinis exposed to contaminated sediment in the Baltic Sea. *Ecol Indic* **63**, 187–195 (2016).
20. Löf, M., Sundelin, B., Bandh, C. & Gorokhova, E. Embryo aberrations in the amphipod Monoporeia affinis as indicators of toxic pollutants in sediments: A field evaluation. *Ecol Indic* **60**, 18–30 (2016).
21. Blomqvist, S. & Lundgren, L. A benthic sled for sampling soft bottoms. *Helgolander Meeresun* **50**, 453–456 (1996).
22. HELCOM supplementary indicator report. Reproductive disorders: malformed embryos of amphipods (Helsinki, Finland, 2018).
23. Giraffa, G., Rossetti, L. & Neviani, E. An evaluation of chelex-based DNA purification protocols for the typing of lactic acid bacteria. *J Microbiol Meth* **42**, 175–184 (2000).
24. Song, L. G., James, S. R., Kazim, L. & Karpf, A. R. Specific method for the determination of genomic DNA methylation by liquid chromatography-electrospray ionization tandem mass spectrometry. *Anal Chem* **77**, 504–510 (2005).
25. Vandegehuchte, M. B., Lemiere, F. & Janssen, C. R. Quantitative DNA-methylation in Daphnia magna and effects of multigeneration Zn exposure. *Comp Biochem Phys C* **150**, 343–348 (2009).
26. Anderson, M. J. Permutational Multivariate Analysis of Variance (PERMANOVA). *Wiley StatsRef: Statistics Reference Online*, 1–15 (2017).
27. Xi, B., Gu, H., Baniasadi, H. & Raftery, D. Statistical analysis and modeling of mass spectrometry-based metabolomics data. *Methods Mol Biol* **1198**, 333–353 (2014).
28. Xia, J. & Wishart, D. S. Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. *Nat Protoc* **6**, 743–760 (2011).
29. Wold, S., Sjostrom, M. & Eriksson, L. PLS-regression: a basic tool of chemometrics. *Chemometr Intell Lab* **58**, 109–130 (2001).

30. Woodruff, T. J., Schwartz, J. & Giudice, L. C. Research agenda for environmental reproductive health in the 21st century. *J Epidemiol Commun H* **64**, 307–310 (2010).
31. Tretyakova, N., Villalta, P. W. & Kotapati, S. Mass Spectrometry of Structurally Modified DNA. *Chem Rev* **113**, 2395–2436 (2013).
32. Singh, R. & Farmer, P. B. Liquid chromatography-electrospray ionization-mass spectrometry: the future of DNA adduct detection. *Carcinogenesis* **27**, 178–196 (2006).
33. Andrews, C. L., Vouros, P. & Harsch, A. Analysis of DNA adducts using high-performance separation techniques coupled to electrospray ionization mass spectrometry. *J Chromatogr A* **856**, 515–526 (1999).
34. Villalta, P. W. & Balbo, S. The Future of DNA Adductomic Analysis. *Int J Mol Sci* **18** (2017).
35. Tsou, C. C. *et al*. DIA-Umpire: comprehensive computational framework for data-independent acquisition proteomics. *Nat Methods* **12**, 258–264 (2015).
36. Song, Y. M. *et al*. Data-Independent Acquisition-Based Quantitative Proteomic Analysis Reveals Potential Biomarkers of Kidney Cancer. *Proteom Clin Appl* **11** (2017).
37. Zhou, J. T., Li, Y. H., Chen, X., Zhong, L. J. & Yin, Y. X. Development of data-independent acquisition workflows for metabolomic analysis on a quadrupole-orbitrap platform. *Talanta* **164**, 128–136 (2017).
38. Fenaille, F., Saint-Hilaire, P. B., Rousseau, K. & Junot, C. Data acquisition workflows in liquid chromatography coupled to high resolution mass spectrometry-based metabolomics: Where do we stand? *J Chromatogr A* **1526**, 1–12 (2017).
39. Cavalieri, E. *et al*. Mechanism of DNA depurination by carcinogens in relation to cancer initiation. *Iubmb Life* **64**, 169–179 (2012).
40. Haglund, J., Van Dongen, W., Lemiere, F. & Esmans, E. L. Analysis of DNA-phosphate adducts *in vitro* using miniaturized LC-ESI-MS/MS and column switching: Phosphotriesters and alkyl cobalamins. *J Am Soc Mass Spectr* **15**, 593–606 (2004).
41. Xue, W. & Warshawsky, D. Metabolic activation of polycyclic and heterocyclic aromatic hydrocarbons and DNA damage: a review. *Toxicol Appl Pharmacol* **206**, 73–93 (2005).
42. Smith, R. D., Shen, Y. F. & Tang, K. Q. Ultrasensitive and quantitative analyses from combined separations-mass spectrometry for the characterization of proteomes. *Accounts Chem Res* **37**, 269–278 (2004).
43. Wolffe, A. P. & Matzke, M. A. Epigenetics: Regulation through repression. *Science* **286**, 481–486 (1999).
44. Santos, K. F., Mazzola, T. N. & Carvalho, H. F. The prima donna of epigenetics: the regulation of gene expression by DNA methylation. *Braz J Med Biol Res* **38**, 1531–1541 (2005).
45. Schubeler, D. Function and information content of DNA methylation. *Nature* **517**, 321–326 (2015).
46. Ruiz-Hernandez, A. *et al*. Environmental chemicals and DNA methylation in adults: a systematic review of the epidemiologic evidence. *Clin Epigenetics* **7**, 55 (2015).
47. Chappell, G., Pogribny, I. P., Guyton, K. Z. & Rusyn, I. Epigenetic alterations induced by genotoxic occupational and environmental human chemical carcinogens: A systematic literature review. *Mutat Res* **768**, 27–45 (2016).
48. Koziol, M. J. *et al*. Identification of methylated deoxyadenosines in vertebrates reveals diversity in DNA modifications. *Nat Struct Mol Biol* **23**, 24–30 (2016).
49. Wu, T. P. *et al*. DNA methylation on N(6)-adenine in mammalian embryonic stem cells. *Nature* **532**, 329–333 (2016).
50. Lonkar, P. & Dedon, P. C. Reactive species and DNA damage in chronic inflammation: reconciling chemical mechanisms and biological fates. *Int J Cancer* **128**, 1999–2009 (2011).
51. Pang, B. *et al*. Lipid peroxidation dominates the chemistry of DNA adduct formation in a mouse model of inflammation. *Carcinogenesis* **28**, 1807–1813 (2007).
52. Joenje, H. Genetic Toxicology of Oxygen. *Mutation Research* **219**, 193–208 (1989).

Acknowledgements

This work was financially supported by the Swedish Research Council Formas (2017-00864), the Swedish Institute (23716/2018), the Erasmus+ program, the Department of Environmental Science and Analytical Chemistry at Stockholm University, and the United States National Institutes of Health (R01CA095039).

Author contributions

H.V.M. and E.G. conceptualized the project; B.S. performed the embryo analysis; N.H.M. performed the DNA extraction; G.M. and H.V.M. performed the adductomics; E.G. performed the statistical analysis; G.M., H.V.M., and N.Y.T. conducted the adduct identification; H.V.M. and E.G. wrote the manuscript with input from N.Y.T. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-57465-1.

Correspondence and requests for materials should be addressed to H.V.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.