Multiple reaction monitoring mass spectrometry for the discovery of environmentally modulated proteins in an aquatic invertebrate sentinel species, Gammarus fossarum

Maxime Leprêtre, Olivier Geffard, Anabelle Espeyte, Julien Faugere, Sophie Ayciriex, Arnaud Chaumot, Nicolas Delorme, Arnaud Salvador, Davide Degli-Esposti

Environmental Pollution 315 (2022) 120393

Contents lists available at ScienceDirect

Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol

Multiple reaction monitoring mass spectrometry for the discovery of environmentally modulated proteins in an aquatic invertebrate sentinel species, *Gammarus fossarum*☆

Maxime Leprêtre [a], Olivier Geffard [a], Anabelle Espeyte [a], Julien Faugere [b], Sophie Ayciriex [b], Arnaud Salvador [b], Nicolas Delorme [a], Arnaud Chaumot [a], Davide Degli-Esposti [a,*]

[a] INRAE, UR RiverLy, Laboratoire d'écotoxicologie, F-69625, Villeurbanne, France
[b] Université de Lyon, Université Claude Bernard Lyon 1, Institut des Sciences Analytiques, CNRS UMR 5280, 5 rue de la Doua, F-69100, Villeurbanne, France

ARTICLE INFO

Keywords:
Active biomonitoring
Dynamic MRM
Housekeeping peptides
Environmental molecular markers
Gammarus fossarum.

ABSTRACT

Multiple reaction monitoring (MRM) mass spectrometry is emerging as a relevant tool for measuring customized molecular markers in freshwater sentinel species. While this technique is typically used for the validation of protein molecular markers preselected from shotgun experiments, recent gains of MRM multiplexing capacity offer new possibilities to conduct large-scale screening of animal proteomes. By combining the strength of active biomonitoring strategies and MRM technologies, this study aims to propose a new strategy for the discovery of candidate proteins that respond to environmental variability. For this purpose, 249 peptides derived from 147 proteins were monitored by MRM in 273 male gammarids caged in 56 environmental sites, representative of the diversity of French water bodies. A methodology is here proposed to identify a set of customized housekeeping peptides (HKPs) used to correct analytical batch effects and allow proper comparison of peptide levels in gammarids. A comparative analysis performed on HKPs-normalized data resulted in the identification of peptides highly modulated in the environment and derived from proteins likely involved in the environmental stress response. Overall, this study proposes a breakthrough approach to screen and identify potential proteins responding to relevant environmental conditions in sentinel species.

1. Introduction

Thousands of pollutants are routinely disposed in cocktails into the aquatic environment, threatening the health of both wildlife and humans (Schwarzenbach et al., 2006). The importance of ecosystem services (biodiversity, drinking water, recreation) provided by natural water bodies leads to strong societal expectations regarding water quality and protection of this resource. Over the world, these expectations led to the development of legislative framework that requires states to achieve a good chemical and ecological status of their freshwater ecosystems (Hödl et al., 2018), in particular their compliance to environmental quality standard (EQS). EQSs are available for few chemical compounds and current chemical approaches focus on a limited number of selected priority compounds, oversighting a majority of contaminants and the relative cocktail effects (Milinkovitch et al., 2019; Beyer et al., 2014). In this context, the use of biological responses in sentinel organisms is a promising alternative that allows to integrate the effects of chemical cocktails in aquatic environments (Milinkovitch et al., 2019; Wernersson et al., 2015). Over the last decades, considerable efforts have been made to identify biomarkers that provide valuable information on the level of pollution (Adams, 2001; Palos Ladeiro et al., 2017; Catteau et al., 2021; Butcherine et al., 2022). Most ecotoxicological biomarkers developed on invertebrate sentinel species were directly transposed from toxicological studies performed on vertebrate model animals (e.g. zebrafish, mouse), limiting the extrapolation to the biological diversity of ecosystems, as reviewed by Trapp et al. (2014). For example, the measure of several enzyme activities (i.e. catalase, phenoloxydase, acetylcholinesterase) may have different sensitivities to substrates and inhibitors among species. In the same way, the development phase of biomarkers is usually achieved through laboratory investigations, focusing on a single chemical compound often tested at high concentrations compared to those found in the environment

☆ This paper has been recommended for acceptance by Dr. Da Chen.

* Corresponding author.
E-mail address: davide.degli-esposti@inrae.fr (D. Degli-Esposti).

https://doi.org/10.1016/j.envpol.2022.120393
Received 4 May 2022; Received in revised form 3 October 2022; Accepted 5 October 2022
Available online 9 October 2022

(Piedade et al., 2020; Aguirre-Martínez et al., 2016; Giarratano et al., 2014). While laboratory experiments are required to validate the sensitivity and specificity of identified biomarkers, this approach does not take into account the environmental complexity, which includes complex cocktails of pollutants of unknown composition and low contamination levels, for the discovery of markers of interest. For this purpose, field-based studies represent a promising alternative to develop biomarkers that integrate the environmental complexity. However, such challenge requires (i) the use of controlled and calibrated organisms to limit the effect of biotic factors (i.e. active biomonitoring approach) (Catteau et al., 2021; Ji et al., 2010) and (ii) the development of analytical tools to explore and compare properly biological modulations in samples collected at different sites and dates.

The crustacean *Gammarus fossarum* has emerged as a relevant sentinel species for tracking the pollution of freshwater ecosystems. Found in most of European waterbodies, many studies have shown that these organisms bioaccumulate different pollutants in a dose- and time-dependent manner (Alric et al., 2019; Ciliberti et al., 2017; Wattier et al., 2020). Gammarids are also sensitive to pollutants, so several ecotoxicological biomarkers have been developed for this organism. For example, feeding and reproductive assays have been used to assess the biological impact of environmental pollution on *G. fossarum* (Coulaud et al., 2011; Coulaud et al., 2015; Lopes et al., 2020). Easy to handle, field caging of *G. fossarum* is well mastered, standardized (AFNOR XP T90-721) and is now part of the operational tools of French Water Agencies to evaluate environmental quality standards (Babut et al., 2020). This active biomonitoring strategy involve field caging of calibrated organisms (same age, size, sex) collected from one unstressed and uncontaminated population (Besse et al., 2012). In contrast to passive sampling, this methodology has provided the standardization of several biotic parameters which guarantees a reliable comparison of biological responses between organisms exposed *in situ* to large-scale studies (Besse et al., 2012; Oikari, 2006). Therefore, active biomonitoring methodologies are increasingly employed to evaluate the impact of pollution on aquatic organisms, through the measure of ecotoxicological biomarkers (Barjhoux et al., 2018; Palais et al., 2012).

By combining RNA sequencing (RNA-Seq) and MS-based shotgun proteomics, proteogenomic analysis are now routinely conducted to characterize the proteomes of sentinel species without sequenced-genomes (Armengaud et al., 2014). Shotgun analyses also provide the quantitative values of thousands of proteins and are therefore conventionally used to study protein modulations in organisms exposed to pollutants and identify potential molecular biomarkers (Gouveia et al., 2019; Trapp et al., 2018; Teng et al., 2021). For example, shotgun investigations performed on the gammarid *Gammarus fossarum* exposed in the laboratory to different concentrations of xenobiotics resulted in the identification of several candidate proteins related to endocrine perturbation (Trapp et al., 2015; Koenig et al., 2021). However, shotgun analysis remains costly and often suffers from poor reproducibility over time, making it unsuitable for the routine monitoring of biomarkers (Bian et al., 2020). The multiple reaction monitoring mode (MRM) of MS appears to be a promising tool for the quantification and application of protein markers in sentinel species, as already demonstrated in clinical practice (Shi et al., 2016). Targeted proteomic assays offer a highly selective, multiplexed, and accurate quantitation of peptides derived from extracted proteins (Shi et al., 2016; Schmidt et al., 2012). Compared to shotgun proteomics, MRM technologies are cost- and time-effective in terms of sample preparation, and data processing (White, 2011; Percy et al., 2014). These tools are therefore particularly suitable for the protein measurement in large number of samples, such as those collected within biomonitoring survey (Shi et al., 2016; Vidova and Spacil, 2017). In aquatic ecotoxicology, MRM assays have been developed for the quantification of protein markers in *G. fossarum* and *Dreisseina polymorpha* exposed to pollutants in laboratory or field condition. Recent technical and methodological advances in MRM provide rapid detection of hundreds of peptides and related proteins (Faugere et al., 2020; Leprêtre et al., 2020). For example, Faugere et al. (2020) have recently developed an MRM assay that provides the monitoring of 277 reporter peptides of 157 proteins in the whole body of *G. fossarum,* offering the possibility to conduct comprehensive and accurate investigations of the gammarid's proteome and to discover proteins of interest without *a priori* preselection.

As evidenced in several studies, proteomic data acquired by MS are often subject to measurement variability due to variations in sample preparation or in analytical performance, potentially introducing biases in the comparison of protein levels measured in samples. To address this issue, peptides targeted by MRM are generally quantified using heavy isotope-labeled peptides introduced into samples before LC-MS analyses. These internal standards (ISTDs) are spiked into samples after protein digestion steps and are therefore particularly efficient for correcting analytical biases generated by the variations of LC-MS instruments but do not address biases introduced during sample extraction and digestion (Chambers et al., 2021). In the context of exhaustive MRM assays, the use of hundreds of spiked ISTDs appears not a suitable option given the high cost of ISTDs. Thus, normalizing MRM data by the signal of stably expressed endogenous peptides, called housekeeping peptides (HKPs), may be an attractive solution to compare peptide levels measured in samples. Stably expressed in cells, regardless of physiological conditions of organisms, housekeeping genes and proteins have been used for many years as internal control to compare gene and protein expression data obtain by transcriptomic (qPCR, RNAseq, DNA Microarray) and Western blot (WB) analyses (Lee et al., 2016; Wang et al., 2019; Tao et al., 2020). To our knowledge, few studies have proposed HKPs to normalize proteomic data obtained from MS proteomic analyses. Moreover, several reports indicate that common selected housekeeping proteins differ in expression between species and organs (McCurley and Callard, 2008; Joshi et al., 2021), highlighting the need to identify and select HKPs unique to each sentinel species.

As a proof of concept, this study aims to demonstrate the relevance of an original application of targeted proteomic approaches in field studies for the identification of modulated proteins under realistic environmental conditions. Using the advantages of active biomonitoring strategies (*i.e.* calibrated organisms and reliable inter-site comparison) and MRM technologies, this study benefited from a large proteomic dataset, which included the monitoring of 249 peptides derived from 147 proteins in 273 male gammarids caged in 56 environmental sites. The sites belong to the monitoring networks of French water agencies, and they are representative of the national diversity of water bodies. From this unique environmental proteomic dataset, a strategy was first developed to identify and select HKPs for normalizing MRM data and make comparable peptide levels between caged gammarids. Subsequently, a comparative proteomic analysis was conducted to reveal highly modulated peptides under field conditions and define their biological functions.

2. Material and methods

2.1. Caging of gammarids

Field exposure of male gammarids was performed according to a procedure adapted from the standardized caging protocol (AFNOR XP T90-721) as described in Alric et al. (2019). Briefly, male gammarids were collected from well-established population of *G. fossarum* leaving in a watercress farm located in Saint-Maurice-de-Rements (France). Organisms were acclimatized to the laboratory for 10 days in tanks supplied with drilled groundwater, under a constant temperature (12 °C), controlled aeration and a photoperiod of 16/8 h light/dark. During the whole acclimatization period, organisms were fed *ad libitum* with alder leaves (*Alnus glutinosa*). Twenty-four hours before field exposures, male gammarids were calibrated (same size) and placed in punctured polypropylene cylinders ready for caging experiments. Field exposure experiments were conducted from September to October 2017.

For each site, the river typology (ranked from very small to very big), chemical status (either good or bad) and ecological status (ranked from good to bad) were retrieved from the French Water Agencies at the corresponding dates of caging (Table S1). Chemical and ecological statuses are defined following the Water Framework Directive (2000/60/EC). The chemical status is an assessment of the amount of priority substances in the aquatic environment. Good chemical status is achieved when no priority substances exceed agreed standards. The ecological status assessment falls into one of five status classes (bad, poor, moderate, good or high). The chemical status was unknown for only two sites and the ecological status was unknown for only one site. Physicochemical parameters (pH, temperature, dissolved oxygen and conductivity) of water were monitored throughout the exposure (Table S1). Gammarids were caged in 56 environmental sites distributed throughout the French territory (Fig. 1). For each environmental site, four cylinders containing 20 male gammarids, fed *ad libitum* with alder leaves, were placed at the site for 7 days. After 7 days of exposure, gammarids were brought back to the laboratory in water of their respective caging sites. For each environmental condition, 5 gammarids were randomly sampled, weighed, frozen in liquid nitrogen, and stored at −80°c for proteomic investigations.

2.2. Sample preparation for proteomic investigations

Total protein content of gammarids was extracted following the protocol described in Charnot et al. (2017), without modifications. To optimize the reproducibility of sample preparation, digestion and solid phase extraction (SPE) steps were conducted using an automated liquid handler (Biomek NXP, Beckman Coulter). For protein digestion, a volume of 3 mL of ammonium bicarbonate (50 mM) and 130 µL of dithiothreitol at a final concentration of 15 mM was added to 250 µl of protein extracts resuspended in a Tris-HCL buffer (50 mM, pH 7.8). After incubation of 40 min at 60 °C, 200 µL of iodoacetamide was added to a final concentration of 15 mM and samples were placed in the dark at room temperature for 40 min. Then, protein digestion was performed by adding 300 µg of treated TCPK trypsin into samples (Fisher Scientist, France). Samples were incubated for 1 h at 37 °C and the digestion was stopped by the addition of 40 µL of formic acid (FA). Before the SPE step, 10 µL of a mixing solution containing 34 internal standards (ISTD, purity>97%), consisting of peptides isotopically labeled with [$^{15}N_2$ and $^{13}C_6$] on the C-ter of lysine or arginine (Scientific Thermo Fisher, Scotland), were spiked at a concentration of 4 µg/mL in samples. SPE was performed using 60 mg hydrophiliclipophilic balance (HLB) cartridge Oasis from Waters (Millford, MA). HLB-cartridges were conditioned by 1 mL of methanol and 1 mL of acidified water (0.5% FA). After loading 3 mL of sample solution, cartridges were rinsed with 1 mL of water/-methanol (95/5 v/v) solution acidified to 0.5% FA. Peptides were then eluted in an Eppendorf tube by the addition of 1 mL of methanol acidified to 0.5%. After evaporation under nitrogen flow, peptides were re-suspended in 90 µL of water/acetonitrile (90/10) with 0.5% FA. Finally, samples were centrifuged at room temperature during 5 min and 15,000×*g* and supernatants were transferred to vials before LC-MS/MS

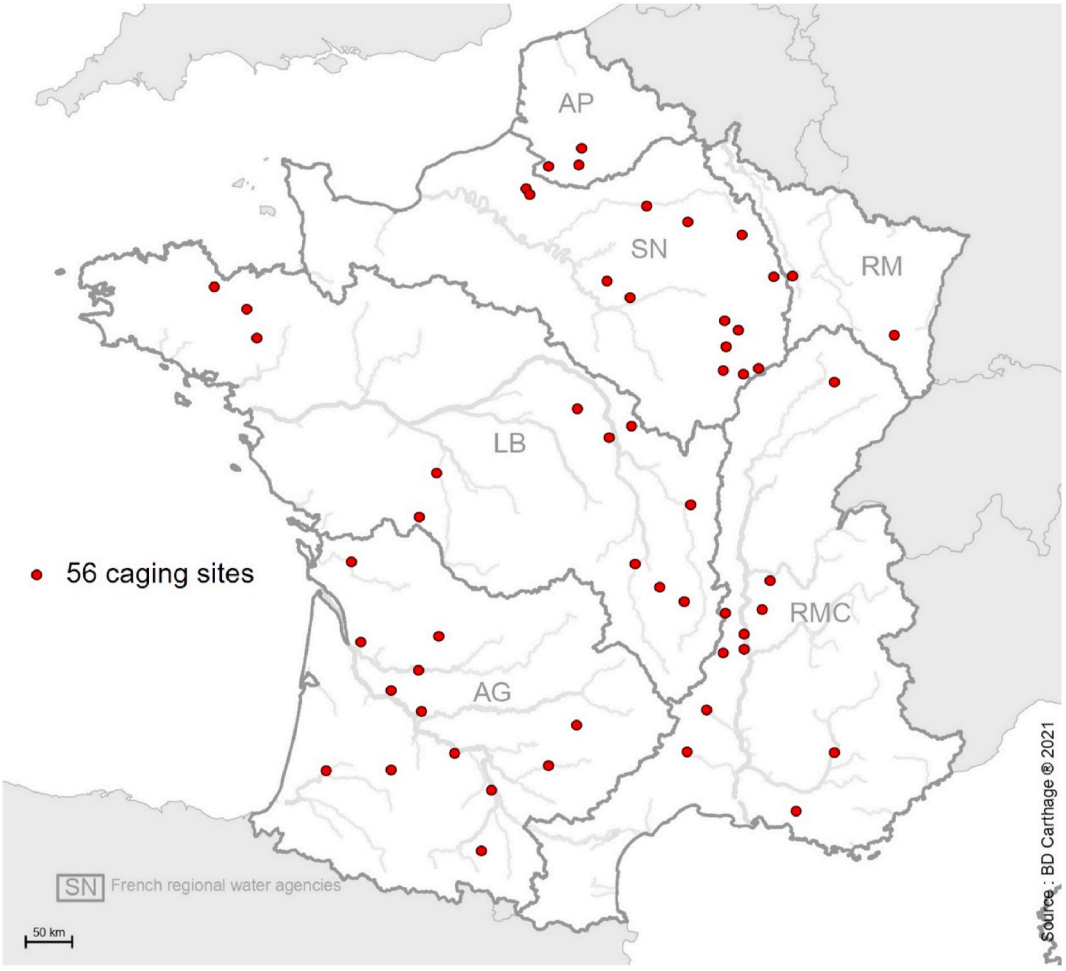

Fig. 1. Distribution of caging sites over the French territory. The red dots represent caging sites of gammarids. The geographical boundaries on the map represent the boundaries of the six French regional environmental agencies: Adour Garonne (AG), Artois-Picardie (AP), Loire-Bretagne (LB), Seine-Normandie (SN), Rhin-Meuse (RM), Rhône Méditerranée-Corse (RMC).

analysis.

2.3. Liquid chromatography and mass spectrometry analysis

Peptide solutions were analyzed with an Agilent 1290 Infinity II High Performance Liquid chromatography (HPLC) system coupled to Agilent 6495B triple quadrupole mass spectrometer (Agilent Technologies, Waldbronn, Germany). LC separation of the 20 μL sample injected was carried out on an XBridge BEH C18 column (100 mm × 2.1 mm, particle size 3.5 μm) coupled to a symmetry XBridge BEH C18 guard column (2.1 mm × 5 mm, particle size 3.5 μm) purchased from Waters (Milford, MA, USA). Elution was performed at a flow rate of 300 μL/min with water containing 0.1% (v/v) FA as eluent A and acetonitrile containing 0.1% (v/v) FA as eluent B, using the following gradient (time in minutes, %B): (0, 5%), (2, 5%), (38, 35%), (38.1, 100%), (43, 100%), (43.1, 5%), (48, 5%) for an injection duty cycle of 48 min considering the column equilibration time. The temperature of the auto-sampler and column were kept at room temperature and 40 °C, respectively.

Instrument control, data acquisition, and processing were performed using MassHunter LC/MS Data Acquisition B.09.00 software. The mass spectrometer was initially tuned and calibrated using Agilent ESI-L Low Concentration Tuning Mix, and checked every week after cleaning protocol. The mass spectrometer was interfaced with an Agilent ESI JetStream (AJS) electrospray ion source operating in positive ionization mode with an ion spray voltage of 4500 V. The nebulizer gas flow rate was set to 12 L/min with a pressure of 40 psi using nitrogen (N2). The ion source was set at 250 °C at the nebulizer and 210 °C at the transfer capillary. A gas flow rate (N2) of 17 L/min was applied in the transfer capillary. The collision cell accelerator voltage was set to a value of 4 V. Ion funnel parameters were set at 145 V for High Pressure RF et 115 V for Low Pressure RF.

The software Skyline 21.1.0.146 (MacCoss Lab Software, USA) was used to produce the list of appropriate MRM transitions. Q1 and Q3 masses for endogenous and heavy labeled peptides as well as collision energy (CE) values were predicted by Skyline. A total of 850 MRM transitions were monitored in a dynamic MRM mode (dMRM), with scheduled windows ranging from 2 to 10 min (Table S2). All MRM transitions were monitored at a unit resolution (0.7 Da) in Q1 and Q3, with a minimum dwell time of 5 ms and a cycle time set to 1500 ms in order to ensure a minimum of 10 points per chromatographic peak. With an average of three MRM transitions per peptide, the dMRM method included the monitoring of 283 peptides. Specifically, 249 peptides derived from 147 *G. fossarum* proteins as well as 34 labeled peptides (ISTDs) of the same sequence as *G. fossarum* peptides were targeted by dMRM (Table S3). As detailed in Faugere et al. (2020), these peptides were selected from a proteogenomic database acquired from *G. fossarum* males and were successfully detected by MRM.

2.4. Multiple reaction monitoring data processing

Quantitative data of targeted peptides were estimated after automatic and manual integration of chromatographic peak areas using MassHunter Quantitative Analysis software (version B.09.00; Agilent Technologies). Peptide levels were assessed using the most intense and the least interfered MRM transitions among the 3 transitions monitored per peptide (Faugere et al., 2020). Only transitions detected with a signal-to-noise ratio (S/N) greater than 5 were considered for further analysis. To minimize the constraints of missing values in differential proteomics (Jin et al., 2021), peptides detected below this limit of quantification (LOQ) were replaced by the minimum values accepted for the same peptides in other samples. Raw data are available in Table S4.

2.5. Diagnosis of batch effects

Batch effects were defined as variations in the overall peptide levels observed across different analytical runs, ranging from sample preparation to MRM analyses. The presence of batch effects in raw data and normalized datasets was assessed visually using different graphical representations as provided by the R package proBatch (Čuklina et al., 2021). In particular, the distribution of log-2 transformed peptide intensities measured in the samples were visually compared using hierarchical clustering heatmap and box plots to reveal the presence of batch effects among the different analytical runs conducted over time.

2.6. Normalizations of raw data

The normalization of raw data was performed on log2 transformed peptide intensities using the dataProcess function of the R package MSstats (Choi et al., 2014). Raw data underwent four different normalization procedures, two were based on the global signal of MRM data distributions and two others were based on the signal acquired from specific peptides. Among global normalization methodologies, quantile and median adjustment methods were applied to raw data by assuming that the distribution of quantiles from measured peptide levels are equally distributed across all samples. The normalizations based on specific peptides were performed using the 'globalstandards' option of the dataProcess function, which assumes that median levels of a set of peptides defined as global standards are the same across samples. Using this procedure, raw data were normalized using (1) the signal acquired from ISTDs introduced into samples with the same quantity before the SPE step, or (2) using the signal of HKPs, identified by the methodology detailed below.

2.7. Identification and selection of housekeeping peptides

Investigation of HKPs was performed on peptides detected in all samples above the LOQ (S/N > 5). The strategy developed for the mapping of HKPs relies on the identification of the least variable and the most correlated peptides measured in different individual samples analyzed by MRM. First, the GeNorm algorithm (Vandesompele et al., 2002) was applied to evaluate the most stable peptides in quantile – and median – normalized datasets as well as the data normalized by ISTDs. In each normalized dataset, peptides were ranked with respect to their stability values (M-values) calculated by GeNorm, where peptides with the lowest M-values were considered as the most stable peptides. Then, an overall stability ranking was assigned to each peptide by summing their ranks in the different normalized datasets. Based on this overall stability ranking, the most stable peptides were selected and further investigated as candidate HKPs. To validate and select a set of stable peptides as HKPs, Pearson's correlation tests were performed on raw intensities of the most stable peptides with respect to the GeNorm analysis. Based on these results, a set of peptides correlated with a Pearson's correlation coefficient higher than 90% were selected as HKPs.

2.8. Identification of differentially expressed peptides

Differential proteomic analysis was performed on the HKPs-normalized dataset to evidence differentially expressed peptides (DEPs) in gammarids exposed to different environmental conditions. To avoid misinterpretation, comparisons of peptide levels were conducted exclusively on well detected peptides measured above the LOQ in at least 80% of samples. First, a student t-test analysis was conducted using the web application Biostatflow (v.2.9.2; http://biostatflow.org) to highlight peptides whose abundances differed statistically between at least two environmental sites. Pairwise comparisons were performed between each environmental site, and multiple test p-values were corrected by the Benjamini-Hochberg method. Peptides levels were considered different between two sites with a false discovery rate (FDR) lower than 0.05. Then, based on median (med) levels and quantile (q) distributions of peptides, induction factors (IFs) were assigned to each peptide to evaluate the degree of their modulation in specific

environmental sites compared to overall sites. IFs specific to each peptide (a) and environmental condition (b) were calculated as follows: IF_{ab}= (med_{ab}-med_a)/(q75(med_a)−q25(med_a)). Based on these calculated IFs, peptides found statistically differentially expressed with IFs greater than 2 were considered as DEPs in specific sites. For each protein, reporter peptides showing the highest IFs were retained and further analyzed.

2.9. Functional analysis

Protein sequences related to targeted peptides were annotated to different functional categories to evaluate the biological functions (BP) and molecular functions (MF) of HKPs and DEPs. BLAST annotations of protein sequences were performed using the BLASTp module of OmicsBox software (https://www.biobam.com/omicsbox) (Götz et al., 2008). The search for sequence similarity was conducted against the SWISSPROT database, without taxonomic restriction. Still using OmicsBox, slim gene ontology (GO) slim annotations of blasted protein sequences were retrieved using the GO mapping module. All functional annotations were validated with an expected value (E-value) threshold set at 1×10^{-3}. To avoid redundancy of GO-slims, MF or BP annotations mapped to DEPs were clustered using the affinity propagation clustering algorithm (Frey and Dueck, 2007). Graphical visualizations of MF and BP GO-slims annotated to DEPs were done with the R package ggplot2.

3. Result and discussion

MRM analyses were performed to monitor a large set of peptides in male gammarids caged in sites covered by the monitoring network of French water agencies, as part of a collaborative project with the French Biodiversity Agency (OFB). In the present study, this large environmental proteomic dataset was processed to investigate the ability of MRM as a promising and original approach for the identification of modulated proteins in realistic environmental conditions.

3.1. Batch effects

The large proteomic dataset included the analysis of 249 reporter peptides of 147 proteins measured in 273 gammarid samples caged in 56 environmental sites. To acquire this dataset, 23 analytical series, each of twelve samples ranging from sample preparation to the MRM analysis, were conducted over a period of about two months (Fig. 2A). Given the large number of samples and analytical series, the presence of potential batch effects related to variations in sample preparation and mass spectrometer responses should be taken into account to properly normalize peptide levels across samples. Boxplots and hierarchical clustered heatmap representations were used to visually inspect data for the presence of batch effects (Fig. 2). Peptide levels measured in samples varied in intensity within and across the 23 analytical series conducted over time, highlighting the presence of batch effects on raw data (Fig. 2A). In particular, the 23 analytical series could be classified into 3 batches, characterized by different patterns of global peptide intensities (Fig. 2A). First, peptide signal intensities showed constant levels in samples analyzed until 6th of August (Batch 1). Then, peptide signal intensities increased significantly in the 3 analytical series performed from the 6th to the 9th of August (batch 2). Finally, in the batch 3, the global peptide intensities returned to levels similar to those observed in batch 1. These different patterns of signal intensities between batches were confirmed by the clustered heatmap (Fig. 2B). Indeed, samples were grouped into three clusters characterized by different peptide intensities. A first cluster (cl 1) was composed of about twenty samples analyzed from all the three batches. These samples differed from the others due to the occurrence of peptides measured with low intensities, probably below the LOQ. The second cluster (cl 2) comprised samples analyzed in batches 1 and 3, which displayed relatively similar peptide-level profiles. The third cluster (cl 3) was composed exclusively of samples analyzed from the batch 2.

The increased peptide levels observed in the samples analyzed from the 6th to the 9th of August are probably associated with a change in the response of the LC-MS instruments, since signal increases were also

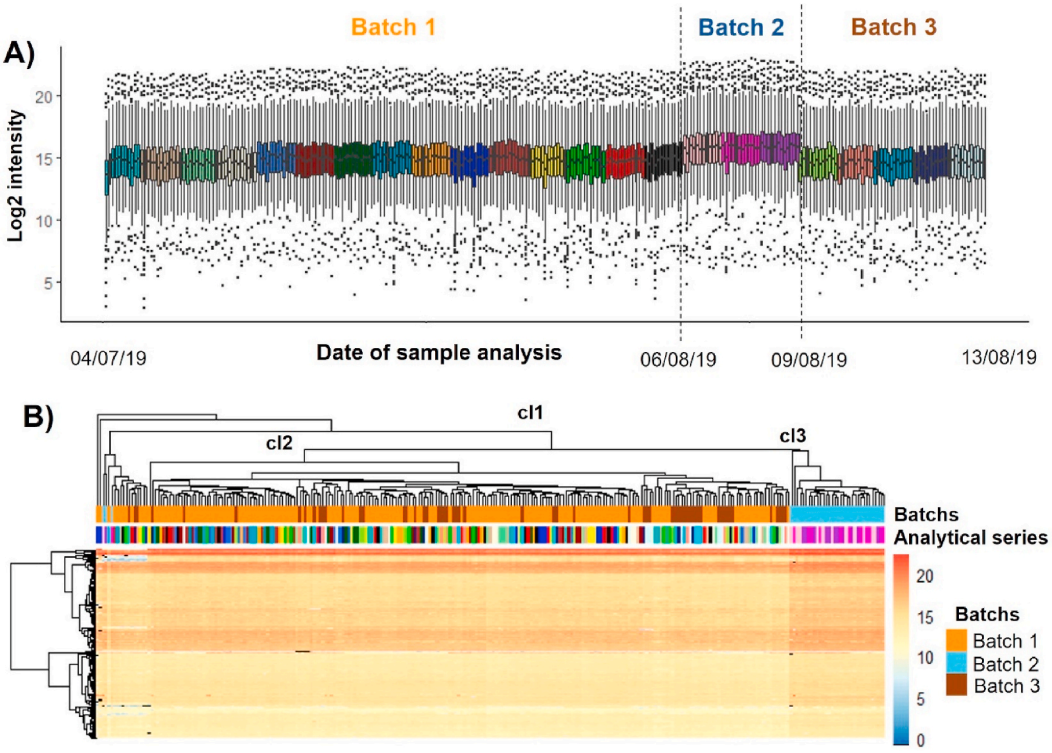

Fig. 2. (A) Boxplots representing the distribution of peptide raw intensities in each sample. Boxplot colors represent the 23 different analytical series, starting from sample preparation to MRM analyses. (B) Clustered heatmap of peptide raw intensities in samples classified by analytical series and batches.

observed for the 34 ISTDs (Fig. S1). Stability of LC-MS instruments can be impacted by several parameters, such as clogging of columns or MS components, causing the drift of LC-MS analytical performances and contributing to the introduction of batch effects (Fernández-Albert et al., 2014). To solve this issue, MRM data are traditionally normalized using the signal of ISTDs (Leprêtre et al., 2020; Charnot et al., 2017; Addona et al., 2009; Liebler and Zimmerman, 2013; Villanueva et al., 2014). Spiked in samples in the same quantities before LC-MS analyses, ISTDs are effective for adjusting biases of quantitation caused by the instability of LC-MS instruments. However, because they are generally spiked in samples after the protein digestion step, normalizations based on ISTDs cannot consider the potentially non-negligible part of variability introduced during sampling and protein extractions (Piehowski et al., 2013). Indeed, subtle random variations in conditions during sample preparation may also significantly contribute to the peptide-level variability observed within the 273 gammarid samples assayed by MRM (Fig. 2). In particular, Piehowski and coworkers (Piehowski et al., 2013) showed that the protein extraction step explain the main part of variability in LC-MS analysis. In this study, sample preparations were performed on calibrated gammarids using the same materials and chemical reagent solutions. In addition, variability was controlled by performing all protein digestion and SPE steps with an automated liquid handler. Nevertheless, different extracted protein yields per samples could explain disparities of peptides levels observed within analytical series and throughout the two-month analysis period. Global normalization methods, such as those based on the median or quantile distribution of peptide levels, have been typically proposed to address technical biases in large-scale proteomic datasets, like in shotgun experiments (Callister et al., 2006; Välikangas et al., 2018). These normalization procedures assume that global signals have the same distribution across samples and are therefore justified in datasets where thousands of peptides are measured and only a minority of proteins are expected to be modulated. However, when only a few peptides are targeted or when they are expected to be modulated, global normalizations may conceal the true biological variability and introduce biases in the comparison of peptide levels between samples (Nakayasu et al., 2021). In this context, normalization with HKPs may be the most appropriate method to minimize biases introduced during both sample preparation and LC-MS analysis steps, ensuring a reliable comparison of peptide levels measured by MRM. However, such a normalization remains challenging since it first implies the identification of ubiquitously and constitutively expressed HKPs in tested samples.

3.2. Identification of housekeeping peptides and raw data normalization

The proteomic dataset was acquired from 273 gammarids exposed *in situ* to 56 watershed sites, representing the variability of realistic and complex environmental conditions (Table S1), making it particularly suitable for the identification of stable HKPs. In fact, previous studies that aimed to propose HKPs derived them from known proteins used as external controls in Western blot or gene expression analysis from model organisms (Hartung et al., 2021; Whiteaker et al., 2011; Groh et al., 2013). Indeed, to be effective in data normalization, HKPs must be ubiquitously and stably expressed in the tissues of the analyzed organisms, regardless of whether their physiological status has been potentially altered by environmental changes (Zhang et al., 2015). To identify and select appropriate HKPs in *G. fossarum*, our methodology relied on the identification of well-detected, stable and correlated peptides among the 273 gammarids. Among the 277 peptides measured by MRM in male gammarids, 126 were detected in all samples above the LOQ (S/N $>$ 5) and were further investigated as candidate HKPs. Then a GeNorm analysis was conducted on the median, quantile and ISTD-normalized datasets to evidence the most stable peptides measured in samples. Based on the mean pairwise variation between individual peptide and other tested peptides, GeNorm algorithm estimates a stability measure for each peptide (M-value) (Vandesompele et al., 2002). According to M-values, peptides were ranked from the most to the least stably expressed in the three normalized datasets. Then, a global ranking was assigned to each peptide by summing their ranks obtained from the three normalized datasets and the most stable peptides were further investigated as candidate HKPs (Table 1). Except for the peptides LFEVGGPPSCTK, GNLANVIR and DVNAAIAAIK, all candidate HKPs were ranked among the most stably expressed peptides in median-, quantile- and ISTD-normalized databases. In particular, the three most stable peptides evidenced by the global ranking (AVIDSGEGLIR, LPTVAAIIYR and LGFLTFCPTNLGTTIR) were ranked in the top 5 of all normalized datasets. To validate and select a set of stable peptides as HKPs, correlations of candidate HKPs were evaluated using Pearson's correlation tests. All candidate HKPs were positively correlated with Pearson's correlation coefficient (R^2) greater than 80% (Fig. S2). Among them, five peptides AVIDSGEGLIR, LPTVAAIIYR, LGFLTFCPTNLGTTIR, LFEVGGPPSCTK, and IINEGAALLR were highly correlated with an R^2 greater than 90% and were therefore selected as HKPs to normalize proteomic data (Fig. S2). According to the functional analysis, these five selected HKPs were derived from proteins that play a central role in the maintenance of cell function and are therefore likely to be constitutively expressed in gammarid cells. In particular, the peptides 'AVIDSGEGLIR' and 'LPTVAAIIYR' were derived from the structural protein alpha actinin and the mitochondrial protein citrate synthase, respectively. Structural proteins are essential to maintain cell shape and mitochondrial proteins are crucial for cell survival since mitochondria are the main organelles that ensure the production of cell energy. In this context, both structural and mitochondrial proteins are considered as valuable housekeeping proteins and have been extensively used as internal reference proteins to normalize protein levels measured by western blots (Eaton et al., 2013). The peptide 'LGFLTFCPTNLGTTIR' was annotated to an Arginine kinase (AK), a key enzyme involved in energy metabolism homeostasis (Yao et al., 2009). Found stably expressed in several invertebrate species, the gene coding for the AK has been considered as a valuable reference gene to normalize gene expression data obtained by quantitative PCR (Lu et al., 2013; Li et al., 2019). Finally, the two other selected peptides (IQVDPAK and FGFYEVFK) were linked to a phosphatase enzyme (PP2). Involved in the regulation of many vital cell-signaling pathways, PP2 proteins have been found stably expressed in several organisms and the gene coding for PP2 is frequently used as a reference gene in qPCR analyses (Ruediger et al., 1991; Seshacharyulu et al., 2013; Tang et al., 2019). The normalization of raw data using selected HKPs eliminated the batch effect previously observed in raw data. As illustrated in Fig. 3, the distribution of peptide levels measured in samples over time were more homogeneous after normalization with HKPs. The clustered heatmap also confirms the correction of batch effects, revealing two clusters independent of analytical series (Fig. 3B). Like in the raw data, a first cluster (cl 1) included samples that contain missing data or peptides measured below the LOQ. The second cluster (cl 2) included other samples analyzed from the 23 analytical series, confirming the batch effect correction.

By providing the measure of hundreds peptides in a considerable number of samples and environmental conditions, MRM tools are offering new opportunities for the identification of customized HKPs in sentinel species. This study proposes, for the first time, a methodology to identify and select a set of HKPs specific to tested biological samples. This strategy led to the identification of co-expressed HKPs derived from key proteins involved in basal cell maintenance, that are stably expressed in males of *G. fossarum*. These selected HKPs have proven to be particularly effective to correct batch effects by minimizing technical variability introduced during sample preparations and LC-MS analysis, confirming their value to normalize large scale proteomic data acquired by MRM. Successfully applied to the whole body of *G. fossarum*, this methodology could be easily extended to find HKPs in other biological tissues (i.e. gills, caecum) or sentinel species, provided that a significant number of peptides are measured in samples with contrasting biological states. The use of these HKPs may open new opportunities for the

Table 1
Descriptions of the ten most stable peptides according to their global ranks. In each normalized dataset, peptides were ranked from the most to least stable based on their M-values. Global ranks were then assigned to each peptide by summing their ranks obtained from the different normalized datasets.

Candidate HKPs	Blast annotations	Global Rank	Normalized datasets					
			Median ISTDs		Median		Quantile	
			M-value	Rank	M-value	Rank	M-value	Rank
AVIDSGEGLIR	Alpha-actinin-4	**1**	0.316	1	0.336	1	0.342	2
LPTVAAIIYR	Citrate synthase, mitochondrial	**2**	0.317	2	0.338	2	0.346	3
LGFLTFCPTNLGTTIR	Arginine kinase	**3**	0.321	4	0.339	3	0.353	5
LFEVGGPPSCTK	protein phosphatase 2B	**4**	0.318	3	0.343	4	0.361	11
IQVDPAK	Phosphate carrier protein, mitochondrial	**5**	0.334	10	0.344	5	0.352	4
IINEGAALLR	protein phosphatase 2B	**6**	0.328	7	0.351	9	0.354	6
FGFYEVFK	Phosphate carrier protein, mitochondrial	**7**	0.322	5	0.348	8	0.365	10
ASHSDILASIAK	unknown	**8**	0.333	8	0.347	7	0.361	9
GNLANVIR	ADP/ATP translocase 2	**9**	0.323	6	0.346	6	0.373	18
DVNAAIAAIK	Tubulin alpha-1C chain	**10**	0.338	16	0.351	10	0.356	7

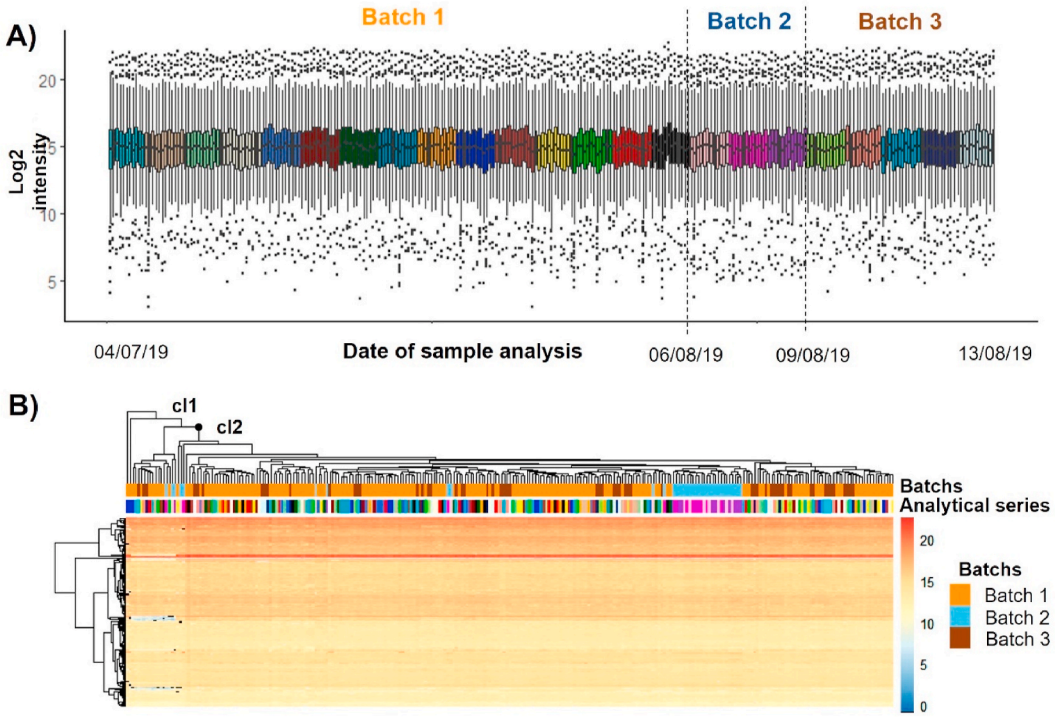

Fig. 3. (A) Boxplots representing the distribution of peptide HKPs-normalized intensities in each sample. Boxplot colors represent the 23 different analytical series, starting from sample preparation to MRM analyses. (B) Clustered heatmap of peptide HKPs-normalized intensities in samples classified by analytical series and batches.

quantification and assessment of protein markers in male gammarids. Indeed, by comparing peptide intensities to HKPs intensities, it may become possible to quantify and properly compare biomarker levels in gammarids collected and analyzed on different dates, as required within large-scale biomonitoring surveys.

3.3. Discovery of proteins highly modulated in the environment

A comparative proteomic analysis was performed on HKPs-normalized data to identify peptides and related proteins modulated in gammarids exposed *in situ* to the different environmental conditions (Table S1). The comparative analysis was performed exclusively on the 220 peptides detected above the LOQ (S/N > 5) in more than 80% of samples. First, a *t*-test analysis was conducted to highlight peptides whose expression differed statistically (FDR<0.05) between at least two environmental sites, resulting in the identification of 75 peptides (related to 47 proteins). Then, induction factors (IFs) of peptides were calculated to estimate their levels of modulation in specific sites. For each protein, reporter peptides showing the highest IFs were retained and further analyzed (Fig. 4). Overall, out of the 75 peptides and 47 proteins that satisfied statistical criteria, 40 DEPs related to 30 proteins displayed IFs greater than 2 and were considered modulated in the environment (Fig. 4). A functional analysis was performed on modulated proteins to identify the biological functions impacted in the environment and evidence the implication of these proteins in response to environmental stress. The purpose of this functional analysis was to find whether general biological trends were present in the list of proteins generated by the mass spectrometry techniques and bioinformatics analyses. The Gene Ontology (GO) analysis reveals that more than 50% of the modulated proteins were annotated to the BP 'stress response' (Fig. 5). More than 30% of proteins were annotated to the BP 'homeostatic process', a BP essential for maintaining internal steady states and often impaired by pollutants, such as heavy metals (Liu and Wang, 2016; Mezzelani et al., 2021). In addition, about 15% of proteins were

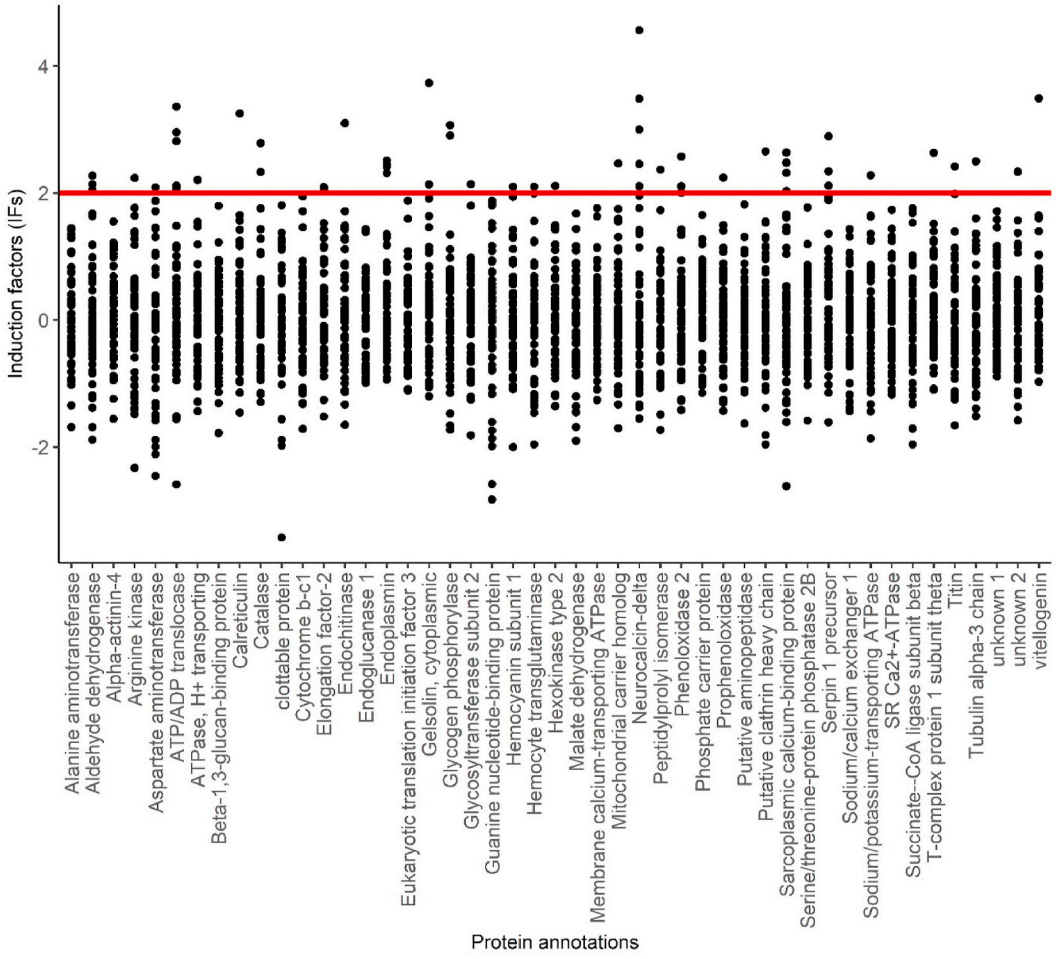

Fig. 4. Induction factors (IFs) of the most induced peptides reporting for the proteins shown in the x axis. Each dot corresponds to one of the 56 environmental sites investigated. Only proteins related to peptides that satisfied statistical criteria (FDR<0.05) are presented. The red line indicates a threshold of IF equal 2.

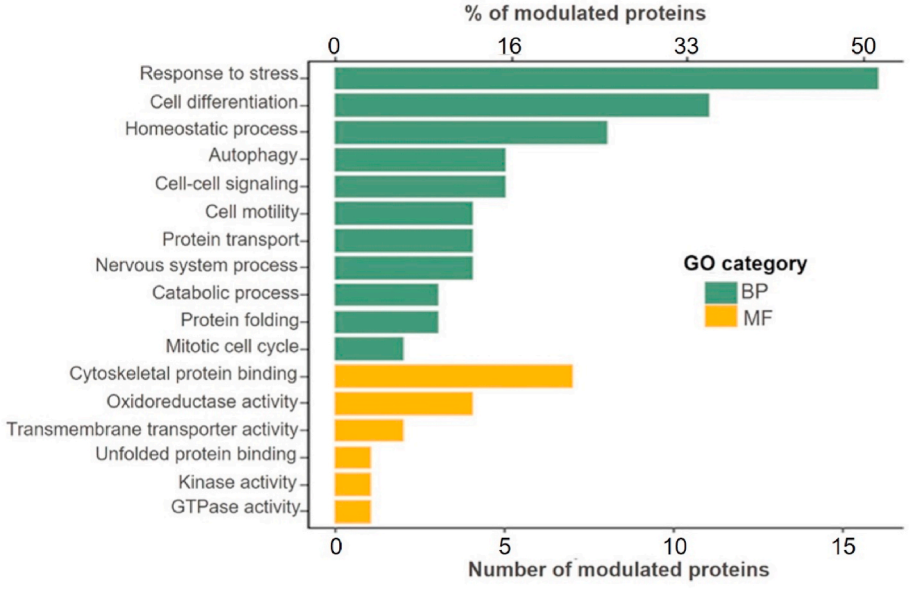

Fig. 5. Biological process (BP) and molecular function (MF) GO slim annotations related to modulated proteins. To avoid redundancy, GO-slims were clustered using affinity propagation clustering algorithm.

assigned to the BPs 'Autophagy' and 'Nervous system' (Fig. 5). Autophagy processes promote the recycling of dysfunctional or damaged cellular elements under conditions of cellular stress (Kroemer et al., 2010). In aquatic organisms, several studies have demonstrated the implication of autophagy processes in response to environmental pollutants and many authors reviewed the value of monitoring autophagic responses to probe the environmental pollution (Tang, 2016). It is also not surprising to identify modulated proteins related to the nervous system since thousands of chemical released in the environment exert neurotoxic effect (Iqubal et al., 2020). For example, several studies have shown the nervous system is targeted by different classes of anthropogenic pollutants, such as pesticides and microplastics (Prokić et al., 2019; Deidda et al., 2021). Regarding the molecular functions, 4 modulated proteins were annotated to the GO-slim 'oxidoreductase activity', a key function involved in the oxidative stress response induced by a wide range of environmental pollutants (Ahmad, 1995; Valavanidis et al., 2006). Out of these 30 modulated proteins, 8 proteins showed IFs higher than 3 (Fig. 4). For example, a gelsolin-like protein was observed with an IF higher than 3.5 in a specific site. Involved in the regulation of actin filament assembly, gelsolins are considered as biomarkers of inflammation in humans (DiNubile, 2008) and their modulation is recurrently observed in invertebrate species exposed in the laboratory to different classes of pollutants (Trapp et al., 2018; Khondee et al., 2016; Lafontaine et al., 2017). Others proteins that were found modulated in the environment (Fig. 4) are already considered as valuable biomarkers to track the impact of pollution in freshwater sentinel species, including *G. fossarum*. In particular, catalase and phenoloxidase enzyme activities have been monitored for several years to assess the oxidative stress or immune modulations of sentinel species exposed to pollutants in laboratory or environmental conditions (Catteau et al., 2021; Luna-Acosta et al., 2017; Xiao et al., 2021).

This study proposes a new methodology for the discovery of protein markers that integrate the environmental complexity, overcoming some of the limits of the traditional single-molecule approach under controlled laboratory conditions. As a proof of concept, this work illustrates the potential of highly multiplexed MRM tools to robustly investigate and compare proteomes of gammarid caged at different environmental sites, such as in large-scale biomonitoring surveys. Given their high modulation patterns and their biological functions related to the stress response, most of the proteins identified in this study may represent potential molecular markers to track the impact of pollution in waterbodies. To validate these proteins as environmental biomarkers, their sensitivities to pollutants and confounding factors must be further investigated. Indeed, several environmental factors, other than pollutants, may also influence the biology of gammarids and lead to protein modulations. For example, the temperature is known to influence several physiological processes in crustacean species, including energy metabolism, reproduction and immune defense (Ren et al., 2021). Finally, protein markers developed at the organ level may respond more sensitively and specifically to pollutants compared to those developed from whole bodies. Depending on their physicochemical properties, pollutants have different modes of action and different accumulation profiles in organs. For instance, Gestin et al. (2021) have recently revealed that heavy metals accumulate differently in the organs of *G. fossarum*. Thus, investigating protein modulations in *G. fossarum* organs targeted by metals (i.e. caecum, gills) would enhance the discovery of metal-sensitive biomarkers.

4. Conclusion

This study highlights the strength of combining active biomonitoring strategies and highly multiplexed MRM mass spectrometry to investigate the proteomes of sentinel organisms under real field conditions. The high-throughput measurement of hundreds of peptides in gammarids caged across the French territory provides a unique dataset of environmental proteomic data. Based on this dataset, a novel strategy was designed to identify customized HKPs, resulting in the identification of stable peptides effective for normalizing MRM data and compare properly peptides levels measured in male gammarids. Then, the comparative proteomic analysis performed on HKPs-normalized data resulted in the identification of highly modulated proteins in the environment, which have biological functions related to the environmental stress response. Overall, this innovative approach offers new perspectives for the discovery of protein biomarkers in sentinel species by providing the identification of modulated proteins that integrate the environmental complexity.

Author statement

Maxime Lepretre: Conceptualization, Formal analysis, Data Curation, Visualization, Writing - Original Draft, Olivier Geffard: Writing - Review & Editing, Funding acquisition, Anabelle. Espeyte: Investigation, Resources, Formal analysis, Julien. Faugere: Methodology, Sophie. Ayciriex: Methodology, Arnaud. Salvador: Methodology, Nicolas. Delorme: Investigation, Resources, Arnaud. Chaumot: Writing - Review & Editing, Funding acquisition, Davide Degli-Esposti: Validation, Writing - Review & Editing, Supervision, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work has been supported by the APPROve project funded by the French National Research Agency (ANR-18-CE34-0013-01) and the French Biodiversity Agency (OFB, Office Français de la Biodiversité). This work benefitted from the French GDR 'Aquatic Ecotoxicology' framework which aims at fostering stimulating scientific discussions and collaborations for more integrative approaches. The authors also thank the French Office for Biodiversity for financial support and Rémi Recoura-Massaquant for his help in mapping the caging sites.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envpol.2022.120393.

References

Adams, S.M., 2001. Biomarker/bioindicator response profiles of organisms can help differentiate between sources of anthropogenic stressors in aquatic ecosystems. Biomarkers 6 (1), 33–44. https://doi.org/10.1080/135475001452779.

Addona, T.A., Abbatiello, S.E., Schilling, B., Skates, S.J., Mani, D.R., Bunk, D.M., Spiegelman, C.H., Zimmerman, L.J., Ham, A.-J.L., Keshishian, H., Hall, S.C., Allen, S., Blackman, R.K., Borchers, C.H., Buck, C., Cardasis, H.L., Cusack, M.P., Dodder, N.G., Gibson, B.W., Held, J.M., Hiltke, T., Jackson, A., Johansen, E.B., Kinsinger, C.R., Li, J., Mesri, M., Neubert, T.A., Niles, R.K., Pulsipher, T.C., Ransohoff, D., Rodriguez, H., Rudnick, P.A., Smith, D., Tabb, D.L., Tegeler, T.J., Variyath, A.M., Vega-Montoto, L.J., Wahlander, Å., Waldemarson, S., Wang, M., Whiteaker, J.R., Zhao, L., Anderson, N.L., Fisher, S.J., Liebler, D.C., Paulovich, A.G., Regnier, F.E., Tempst, P., Carr, S.A., 2009. Multi-site assessment of the precision and reproducibility of multiple reaction monitoring–based measurements of proteins in plasma. Nat. Biotechnol. 27 (7), 633–641. https://doi.org/10.1038/nbt.1546.

Aguirre-Martínez, G.V., DelValls, T.A., Martín-Díaz, M.L., 2016. General stress, detoxification pathways, neurotoxicity and genotoxicity evaluated in ruditapes philippinarum exposed to human pharmaceuticals. Ecotoxicol. Environ. Saf. 124, 18–31. https://doi.org/10.1016/j.ecoenv.2015.09.031.

Ahmad, S., 1995. Oxidative stress from environmental pollutants. Arch. Insect Biochem. Physiol. 29 (2), 135–157. https://doi.org/10.1002/arch.940290205.

Alric, B., Geffard, O., Chandesris, A., Ferréol, M., François, A., Perceval, O., Piffady, J., Villeneuve, B., Chaumot, A., 2019. Multisubstance indicators based on caged *Gammarus* bioaccumulation reveal the influence of chemical contamination on stream macroinvertebrate abundances across France. Environ. Sci. Technol. 53 (10), 5906–5915. https://doi.org/10.1021/acs.est.9b01271.

Armengaud, J., Trapp, J., Pible, O., Geffard, O., Chaumot, A., Hartmann, E.M., 2014. Non-model organisms, a species endangered by proteogenomics. J. Proteonomics 105, 5–18. https://doi.org/10.1016/j.jprot.2014.01.007.

Babut, M., Ferrari, B.J.D., Jame, P., Assoumani, A., Lestremau, F., Hette-Tronquart, N., Miege, C., Perceval, O., 2020. Monitoring priority substances in biota under the water framework directive: how effective is a tiered approach based on caged invertebrates? A proof-of-concept study targeting PFOS in French rivers. Environ. Sci. Eur. 32 (1), 131. https://doi.org/10.1186/s12302-020-00416-4.

Barjhoux, I., Fechner, L.C., Lebrun, J.D., Anzil, A., Ayrault, S., Budzinski, H., Cachot, J., Charron, L., Chaumot, A., Clérandeau, C., Dedourge-Geffard, O., Faburé, J., François, A., Geffard, O., George, I., Labadie, P., Lévi, Y., Munoz, G., Noury, P., Oziol, L., Quéau, H., Servais, P., Uher, E., Urien, N., Geffard, A., 2018. Application of a multidisciplinary and integrative weight-of-evidence approach to a 1-year monitoring survey of the seine river. Environ. Sci. Pollut. Res. Int. 25 (24), 23404–23429. https://doi.org/10.1007/s11356-016-6993-6.

Besse, J.-P., Geffard, O., Coquery, M., 2012. Relevance and applicability of active biomonitoring in continental waters under the water framework directive. TrAC, Trends Anal. Chem. 36, 113–127. https://doi.org/10.1016/j.trac.2012.04.004.

Beyer, J., Petersen, K., Song, Y., Ruus, A., Grung, M., Bakke, T., Tollefsen, K.E., 2014. Environmental risk assessment of combined effects in aquatic ecotoxicology: a discussion paper. Mar. Environ. Res. 96, 81–91. https://doi.org/10.1016/j.marenvres.2013.10.008.

Bian, Y., Zheng, R., Bayer, F.P., Wong, C., Chang, Y.-C., Meng, C., Zolg, D.P., Reinecke, M., Zecha, J., Wiechmann, S., Heinzlmeir, S., Scherr, J., Hemmer, B., Baynham, M., Gingras, A.-C., Boychenko, O., Kuster, B., 2020. Robust, reproducible and quantitative analysis of thousands of proteomes by micro-flow LC–MS/MS. Nat. Commun. 11 (1), 157. https://doi.org/10.1038/s41467-019-13973-x.

Butcherine, P., Kelaher, B.P., Benkendorff, K., 2022. Assessment of acetylcholinesterase, catalase, and glutathione S-transferase as biomarkers for imidacloprid exposure in penaeid shrimp. Aquat. Toxicol. 242, 106050 https://doi.org/10.1016/j.aquatox.2021.106050.

Callister, S.J., Barry, R.C., Adkins, J.N., Johnson, E.T., Qian, W., Webb-Robertson, B.-J. M., Smith, R.D., Lipton, M.S., 2006. Normalization approaches for removing systematic biases associated with mass spectrometry and label-free proteomics. J. Proteome Res. 5 (2), 277–286. https://doi.org/10.1021/pr050300l.

Catteau, A., Porcher, J.-M., Bado-Nilles, A., Bonnard, I., Bonnard, M., Chaumot, A., David, E., Dedourge-Geffard, O., Delahaut, L., Delorme, N., François, A., Garnero, L., Lopes, C., Nott, K., Noury, P., Palluel, O., Palos-Ladeiro, M., Quéau, H., Ronkart, S., Sossey-Alaoui, K., Turiès, C., Tychon, B., Geffard, O., Geffard, A., 2021. Interest of a multispecies approach in active biomonitoring: application in the meuse watershed. Sci. Total Environ., 152148 https://doi.org/10.1016/j.scitotenv.2021.152148.

Chambers, A.G., Sweet, S.M.M., Chain, D., Kim, Y.J., 2021. Global extraction from parallel reaction monitoring to quantify background peptides for improved normalization and quality control in targeted proteomics. Anal. Chem. 93 (40), 13434–13440. https://doi.org/10.1021/acs.analchem.1c01375.

Charnot, A., Gouveia, D., Armengaud, J., Almunia, C., Chaumot, A., Lemoine, J., Geffard, O., Salvador, A., 2017. Multiplexed assay for protein quantitation in the invertebrate Gammarus fossarum by liquid chromatography coupled to tandem mass spectrometry. Anal. Bioanal. Chem. 409 (16), 3969–3991. https://doi.org/10.1007/s00216-017-0348-0.

Choi, M., Chang, C.-Y., Clough, T., Broudy, D., Killeen, T., MacLean, B., Vitek, O., 2014. MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. Bioinformatics 30 (17), 2524–2526. https://doi.org/10.1093/bioinformatics/btu305.

Ciliberti, A., Chaumot, A., Recoura-Massaquant, R., Chandesris, A., François, A., Coquery, M., Ferréol, M., Geffard, O., 2017. Caged Gammarus as biomonitors identifying thresholds of toxic metal bioavailability that affect gammarid densities at the French national scale. Water Res. 118, 131–140. https://doi.org/10.1016/j.watres.2017.04.031.

Coulaud, R., Geffard, O., Xuereb, B., Lacaze, E., Quéau, H., Garric, J., Charles, S., Chaumot, A., 2011. In situ feeding assay with Gammarus fossarum (Crustacea): modelling the influence of confounding factors to improve water quality biomonitoring. Water Res. 45 (19), 6417–6429. https://doi.org/10.1016/j.watres.2011.09.035.

Coulaud, R., Geffard, O., Vigneron, A., Quéau, H., François, A., Chaumot, A., 2015. Linking feeding inhibition with reproductive impairment in Gammarus confirms the ecological relevance of feeding assays in environmental monitoring. Environ. Toxicol. Chem. 34 (5), 1031–1038. https://doi.org/10.1002/etc.2886.

Čuklina, J., Lee, C.H., Williams, E.G., Sajic, T., Collins, B.C., Rodríguez Martínez, M., Sharma, V.S., Wendt, F., Goetze, S., Keele, G.R., Wollscheid, B., Aebersold, R., Pedrioli, P.G.A., 2021. Diagnostics and correction of batch effects in large-scale proteomic studies: a tutorial. Mol. Syst. Biol. 17 (8), e10240 https://doi.org/10.15252/msb.202110240.

Deidda, I., Russo, R., Bonaventura, R., Costa, C., Zito, F., Lampiasi, N., 2021. Neurotoxicity in marine invertebrates: an update. Biology 10 (2), 161. https://doi.org/10.3390/biology10020161.

DiNubile, M.J., 2008. Plasma gelsolin as a biomarker of inflammation. Arthritis Res. Ther. 10 (6), 124. https://doi.org/10.1186/ar2547.

Faugere, J., Gouveia, D., Ayciriex, S., Chaumot, A., Almunia, C., François, A., Armengaud, J., Lemoine, J., Geffard, O., Degli-Esposti, D., Salvador, A., 2020. High-multiplexed monitoring of protein biomarkers in the sentinel Gammarus fossarum by targeted scout-MRM assay, a new vision for ecotoxicoproteomics. J. Proteonomics 226, 103901. https://doi.org/10.1016/j.jprot.2020.103901.

Fernández-Albert, F., Llorach, R., Garcia-Aloy, M., Ziyatdinov, A., Andres-Lacueva, C., Perera, A., 2014. Intensity drift removal in LC/MS metabolomics by common variance compensation. Bioinforma. Oxf. Engl. 30 (20), 2899–2905. https://doi.org/10.1093/bioinformatics/btu423.

Frey, B.J., Dueck, D., 2007. Clustering by passing messages between data points. Science 315 (5814), 972–976. https://doi.org/10.1126/science.1136800.

Gestin, O., Lacoue-Labarthe, T., Coquery, M., Delorme, N., Garnero, L., Dherret, L., Ciccia, T., Geffard, O., Lopes, C., 2021. One and multi-compartments toxico-kinetic modeling to understand metals' organotropism and fate in Gammarus fossarum. Environ. Int. 156, 106625 https://doi.org/10.1016/j.envint.2021.106625.

Giarratano, E., Gil, M.N., Malanga, G., 2014. Biomarkers of environmental stress in gills of ribbed mussel Aulacomya atra atra (nuevo gulf, northern patagonia). Ecotoxicol. Environ. Saf. 107, 111–119. https://doi.org/10.1016/j.ecoenv.2014.05.003.

Götz, S., García-Gómez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., Robles, M., Talón, M., Dopazo, J., Conesa, A., 2008. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res. 36 (10), 3420–3435. https://doi.org/10.1093/nar/gkn176.

Gouveia, D., Almunia, C., Cogne, Y., Pible, O., Degli-Esposti, D., Salvador, A., Cristobal, S., Sheehan, D., Chaumot, A., Geffard, O., Armengaud, J., 2019. Ecotoxicoproteomics: a decade of progress in our understanding of anthropogenic impact on the environment. J. Proteonomics. https://doi.org/10.1016/j.jprot.2018.12.001.

Groh, K.J., Schönenberger, R., Eggen, R.I.L., Segner, H., Suter, M.J.-F., 2013. Analysis of protein expression in zebrafish during gonad differentiation by targeted proteomics. Gen. Comp. Endocrinol. 193, 210–220. https://doi.org/10.1016/j.ygcen.2013.07.020.

Hartung, N.M., Ostermann, A.I., Immenschuh, S., Schebb, N.H., 2021. Combined targeted proteomics and oxylipin metabolomics for monitoring of the COX-2 pathway. Proteomics 21 (3–4), 1900058. https://doi.org/10.1002/pmic.201900058.

Hödl, E., 2018. Legislative framework for river ecosystem management on international and European level. In: Schmutz, S., Sendzimir, J. (Eds.), Riverine Ecosystem Management: Science for Governing towards a Sustainable Future, Aquatic Ecology Series. Springer International Publishing, Cham, pp. 325–345. https://doi.org/10.1007/978-3-319-73250-3_17.

Iqubal, A., Ahmed, M., Ahmad, S., Sahoo, C.R., Iqubal, M.K., Haque, S.E., 2020. Environmental neurotoxic pollutants: Review. Environ. Sci. Pollut. Res. Int. 27 (33), 41175–41198. https://doi.org/10.1007/s11356-020-10539-z.

Ji, Y., Lu, G., Wang, C., Song, W., Wu, H., 2010. Fish transplantation and stress-related biomarkers as useful tools for assessing water quality. J. Environ. Sci. 22 (11), 1826–1832. https://doi.org/10.1016/S1001-0742(09)60326-0.

Jin, L., Bi, Y., Hu, C., Qu, J., Shen, S., Wang, X., Tian, Y., 2021. A comparative study of evaluating missing value imputation methods in label-free proteomics. Sci. Rep. 11 (1), 1760. https://doi.org/10.1038/s41598-021-81279-4.

Joshi, C.J., Ke, W., Drangowska-Way, A., O'Rourke, E.J., Lewis, N.E., 2021. What are housekeeping genes? August bioRxiv 3. https://doi.org/10.1101/2021.02.14.431149, 2021.02.14.431149.

Khondee, P., Srisomsap, C., Chokchaichamnankit, D., Svasti, J., Simpson, R.J., Kingtong, S., 2016. Histopathological effect and stress response of mantle proteome following TBT exposure in the hooded oyster saccostrea cucullata. Environ. Pollut. 218, 855–862. https://doi.org/10.1016/j.envpol.2016.08.011.

Koenig, N., Almunia, C., Bonnal-Conduzorgues, A., Armengaud, J., Chaumot, A., Geffard, O., Esposti, D.D., 2021. Co-expression network analysis identifies novel molecular pathways associated with cadmium and pyriproxyfen testicular toxicity in Gammarus fossarum. Aquat. Toxicol. 235, 105816 https://doi.org/10.1016/j.aquatox.2021.105816.

Kroemer, G., Mariño, G., Levine, B., 2010. Autophagy and the integrated stress response. Mol. Cell. 40 (2), 280–293. https://doi.org/10.1016/j.molcel.2010.09.023.

Lafontaine, A., Baiwir, D., Joaquim-Justo, C., De Pauw, E., Lemoine, S., Boulangé-Lecomte, C., Forget-Leray, J., Thomé, J.-P., Gismondi, E., 2017. Proteomic response of Macrobrachium rosenbergii hepatopancreas exposed to chlordecone: identification of endocrine disruption biomarkers? Ecotoxicol. Environ. Saf. 141, 306–314. https://doi.org/10.1016/j.ecoenv.2017.03.043.

Lee, H.-G., Jo, J., Hong, H.-H., Kim, K.K., Park, J.-K., Cho, S.-J., Park, C., 2016. State-of-the-Art housekeeping proteins for quantitative western blotting: revisiting the first Draft of the human proteome. Proteomics 16 (13), 1863–1867. https://doi.org/10.1002/pmic.201500344.

Leprêtre, M., Palos Ladeiro, M., Faugere, J., Almunia, C., Lemoine, J., Armengaud, J., Geffard, A., Salvador, A., 2020. From shotgun to targeted proteomics: rapid scout-MRM assay development for monitoring potential immunomarkers in Dreissena polymorpha. Anal. Bioanal. Chem.

Li, Q.-Y., Li, Z.-L., Lu, M.-X., Cao, S.-S., Du, Y.-Z., 2019. Selection of valid reference genes for quantitative real-time PCR in cotesia chilonis (hymenoptera: braconidae) exposed to different temperatures. PLoS One 14 (12), e0226139. https://doi.org/10.1371/journal.pone.0226139.

Liebler, D.C., Zimmerman, L.J., 2013. Targeted quantitation of proteins by mass spectrometry. Biochemistry 52 (22), 3797–3806. https://doi.org/10.1021/bi400110b.

Liu, X., Wang, W.-X., 2016. Physiological and cellular responses of oysters (Crassostrea hongkongensis) in a multimetal-contaminated estuary. Environ. Toxicol. Chem. 35 (10), 2577–2586. https://doi.org/10.1002/etc.3426.

Lopes, C., Chaumot, A., Xuereb, B., Coulaud, R., Jubeaux, G., Quéau, H., François, A., Geffard, O., 2020. In situ reproductive bioassay with caged Gammarus fossarum (Crustacea): Part 2-evaluating the relevance of using a molt cycle temperature-

dependent model as a reference to assess toxicity in freshwater monitoring. Environ. Toxicol. Chem. 39 (3), 678–691. https://doi.org/10.1002/etc.4656.
Lu, Y., Yuan, M., Gao, X., Kang, T., Zhan, S., Wan, H., Li, J., 2013. Identification and validation of reference genes for gene expression analysis using quantitative PCR in spodoptera litura (Lepidoptera: noctuidae). PLoS One 8 (7), e68059. https://doi.org/10.1371/journal.pone.0068059.
Luna-Acosta, A., Breitwieser, M., Renault, T., Thomas-Guyon, H., 2017. Recent findings on phenoloxidases in bivalves. Mar. Pollut. Bull. 122 (1–2), 5–16. https://doi.org/10.1016/j.marpolbul.2017.06.031.
McCurley, A.T., Callard, G.V., 2008. Characterization of housekeeping genes in zebrafish: male-female differences and effects of tissue type, developmental stage and chemical treatment. BMC Mol. Biol. 9 (1), 102. https://doi.org/10.1186/1471-2199-9-102.
Mezzelani, M., Nardi, A., Bernardini, I., Milan, M., Peruzza, L., d'Errico, G., Fattorini, D., Gorbi, S., Patarnello, T., Regoli, F., 2021. Environmental pharmaceuticals and climate change: the case study of carbamazepine in M. Galloprovincialis under ocean acidification scenario. Environ. Int. 146, 106269 https://doi.org/10.1016/j.envint.2020.106269.
Milinkovitch, T., Geffard, O., Geffard, A., Mouneyrac, C., Chaumot, A., Xuereb, B., Fisson, C., Minier, C., Auffret, M., Perceval, O., Egea, E., Sanchez, W., 2019. Biomarkers as tools for monitoring within the water framework directive context: concept, opinions and advancement of expertise. Environ. Sci. Pollut. Res. 26 (32), 32759–32763. https://doi.org/10.1007/s11356-019-06434-x.
Nakayasu, E.S., Gritsenko, M., Piehowski, P.D., Gao, Y., Orton, D.J., Schepmoes, A.A., Fillmore, T.L., Frohnert, B.I., Rewers, M., Krischer, J.P., Ansong, C., Suchy-Dicey, A. M., Evans-Molina, C., Qian, W.-J., Webb-Robertson, B.-J.M., Metz, T.O., 2021. Tutorial: best practices and considerations for mass-spectrometry-based protein biomarker discovery and validation. Nat. Protoc. 16 (8), 3737–3760. https://doi.org/10.1038/s41596-021-00566-6.
Oikari, A., 2006. Caging techniques for field exposures of fish to chemical contaminants. Aquat. Toxicol. Amst. Neth. 78 (4), 370–381. https://doi.org/10.1016/j.aquatox.2006.03.010.
Palais, F., Geffard, O.D., Beaudon, A., Devin, S.P., Trapp, J., Olivier, G., Noury, P., Francé, C.G., Uher, E., Mouneyrac, C., Risbourg, S.B., Geffard, A., 2012. One-year monitoring of core biomarker and digestive enzyme responses in transplanted zebra mussels (Dreissena polymorpha). Ecotoxicology 21, 888.
Palos Ladeiro, M.P., Barjhoux, I., Bigot-Clivot, A., Bonnard, M., David, E., Dedourge-Geffard, O., Geba, E., Lance, E., Lepretre, M., Magniez, G., Rioult, D., Aubert, D., Villena, I., Daniele, G., Salvador, A., Vulliet, E., Armengaud, J., Geffard, A., 2017. Mussel as a tool to define continental watershed quality. In: Ray, S. (Ed.), Organismal and Molecular Malacology. InTech. https://doi.org/10.5772/67995.
Percy, A.J., Chambers, A.G., Yang, J., Hardie, D.B., Borchers, C.H., 2014. Advances in multiplexed MRM-based protein biomarker quantitation toward clinical utility. Biochim. Biophys. Acta BBA - Proteins Proteomics 1844 (5), 917–926. https://doi.org/10.1016/j.bbapap.2013.06.008.
Piedade, F., Bio, S., Nunes, B., 2020. Effects of common pharmaceutical drugs (paracetamol and acetylsalicylic acid) short term exposure on biomarkers of the mussel Mytilus spp. Environ. Toxicol. Pharmacol. 73, 103276 https://doi.org/10.1016/j.etap.2019.103276.
Piehowski, P.D., Petyuk, V.A., Orton, D.J., Xie, F., Ramirez-Restrepo, M., Engel, A., Lieberman, A.P., Albin, R.L., Camp, D.G., Smith, R.D., Myers, A.J., 2013. Sources of technical variability in quantitative LC-MS proteomics: human brain tissue sample analysis. J. Proteome Res. 12 (5), 2128–2137. https://doi.org/10.1021/pr301146m.
Prokić, M.D., Radovanović, T.B., Gavrić, J.P., Faggio, C., 2019. Ecotoxicological effects of microplastics: examination of biomarkers, current state and future perspectives. TrAC, Trends Anal. Chem. 111, 37–46. https://doi.org/10.1016/j.trac.2018.12.001.
Ren, X., Wang, Q., Shao, H., Xu, Y., Liu, P., Li, J., 2021. Effects of low temperature on shrimp and crab physiology, behavior, and growth: a Review. Front. Mar. Sci. 8.
Ruediger, R., Hood, J.E.V.W., Mumby, M., Walter, G., 1991. Constant expression and activity of protein phosphatase 2A in synchronized cells. Mol. Cell Biol. https://doi.org/10.1128/mcb.11.8.4282-4285.1991.
Schmidt, C., Urlaub, H., 2012. Absolute quantification of proteins using standard peptides and multiple reaction monitoring. In: Marcus, K. (Ed.), Quantitative Methods in Proteomics, Methods in Molecular Biology. Humana Press, Totowa, NJ, pp. 249–265. https://doi.org/10.1007/978-1-61779-885-6_17.
Schwarzenbach, R.P., Escher, B.I., Fenner, K., Hofstetter, T.B., Johnson, C.A., von Gunten, U., Wehrli, B., 2006. The challenge of micropollutants in aquatic systems. Science 313 (5790), 1072–1077. https://doi.org/10.1126/science.1127291.
Seshacharyulu, P., Pandey, P., Datta, K., Batra, S.K., 2013. Phosphatase: PP2A structural importance, regulation and its aberrant expression in cancer. Cancer Lett. 335 (1), 9–18. https://doi.org/10.1016/j.canlet.2013.02.036.
Shi, T., Song, E., Nie, S., Rodland, K.D., Liu, T., Qian, W.-J., Smith, R.D., 2016. Advances in targeted proteomics and applications to biomedical research. Proteomics 16 (15–16), 2160–2182. https://doi.org/10.1002/pmic.201500449.
Tang, B.L., 2016. Autophagy in response to environmental stresses: new monitoring perspectives. Ecol. Indicat. 60, 453–459. https://doi.org/10.1016/j.ecolind.2015.07.022.
Tang, F., Chu, L., Shu, W., He, X., Wang, L., Lu, M., 2019. Selection and validation of reference genes for quantitative expression analysis of MiRNAs and MRNAs in poplar. Plant Methods 15 (1), 35. https://doi.org/10.1186/s13007-019-0420-1.
Tao, J., Hao, Y., Li, X., Yin, H., Nie, X., Zhang, J., Xu, B., Chen, Q., Li, B., 2020. Systematic identification of housekeeping genes possibly used as references in Caenorhabditis elegans by large-scale data integration. Cells 9 (3). https://doi.org/10.3390/cells9030786.
Teng, J., Zhao, J., Zhu, X., Shan, E., Wang, Q., 2021. Oxidative stress biomarkers, physiological responses and proteomic profiling in oyster (Crassostrea gigas) exposed to microplastics with irregular-shaped PE and PET microplastic. Sci. Total Environ. 786, 147425 https://doi.org/10.1016/j.scitotenv.2021.147425.
Total protein analysis as a reliable loading control for quantitative fluorescent western blotting. PLoS One 8 (8), 2013, e72457. https://doi.org/10.1371/journal.pone.0072457.
Trapp, J., Armengaud, J., Salvador, A., Chaumot, A., Geffard, O., 2014. Next-generation proteomics: toward customized biomarkers for environmental biomonitoring. Environ. Sci. Technol. 48 (23), 13560–13572. https://doi.org/10.1021/es501673s.
Trapp, J., Armengaud, J., Pible, O., Gaillard, J.-C., Abbaci, K., Habtoul, Y., Chaumot, A., Geffard, O., 2015. Proteomic investigation of male *Gammarus fossarum* , a freshwater Crustacean, in response to endocrine disruptors. J. Proteome Res. 14 (1), 292–303. https://doi.org/10.1021/pr500984z.
Trapp, J., Gouveia, D., Almunia, C., Pible, O., Degli Esposti, D., Gaillard, J.-C., Chaumot, A., Geffard, O., Armengaud, J., 2018. Digging deeper into the pyriproxyfen-response of the amphipod Gammarus fossarum with a next-generation ultra-high-field orbitrap analyser: new perspectives for environmental toxicoproteomics. Front. Environ. Sci. 6, 54. https://doi.org/10.3389/fenvs.2018.00054.
Valavanidis, A., Vlahogianni, T., Dassenakis, M., Scoullos, M., 2006. Molecular biomarkers of oxidative stress in aquatic organisms in relation to toxic environmental pollutants. Ecotoxicol. Environ. Saf. 64 (2), 178–189. https://doi.org/10.1016/j.ecoenv.2005.03.013.
Välikangas, T., Suomi, T., Elo, L.L., 2018. A systematic evaluation of normalization methods in quantitative label-free proteomics. Briefings Bioinf. 19 (1), 1–11. https://doi.org/10.1093/bib/bbw095.
Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., Speleman, F., 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. 3 (7), 1. https://doi.org/10.1186/gb-2002-3-7-research0034 research0034.
Vidova, V., Spacil, Z., 2017. A Review on mass spectrometry-based quantitative proteomics: targeted and data independent acquisition. Anal. Chim. Acta 964, 7–23. https://doi.org/10.1016/j.aca.2017.01.059.
Villanueva, J., Carrascal, M., Abian, J., 2014. Isotope dilution mass spectrometry for absolute quantification in proteomics: concepts and strategies. J. Proteonomics 96, 184–199. https://doi.org/10.1016/j.jprot.2013.11.004.
Wang, Z., Lyu, Z., Pan, L., Zeng, G., Randhawa, P., 2019. Defining housekeeping genes suitable for RNA-seq analysis of the human allograft kidney biopsy tissue. BMC Med. Genom. 12 (1), 86. https://doi.org/10.1186/s12920-019-0538-z.
Wattier, R., Mamos, T., Copilaş-Ciocianu, D., Jelić, M., Ollivier, A., Chaumot, A., Danger, M., Felten, V., Piscart, C., Żganec, K., Rewicz, T., Wysocka, A., Rigaud, T., Grabowski, M., 2020. Continental-scale patterns of hyper-cryptic diversity within the freshwater model taxon Gammarus fossarum (Crustacea, Amphipoda). Sci. Rep. 10 (1), 16536 https://doi.org/10.1038/s41598-020-73739-0.
Wernersson, A.-S., Carere, M., Maggi, C., Tusil, P., Soldan, P., James, A., Sanchez, W., Dulio, V., Broeg, K., Reifferscheid, G., Buchinger, S., Maas, H., Van Der Grinten, E., O'Toole, S., Ausili, A., Manfra, L., Marziali, L., Polesello, S., Lacchetti, I., Mancini, L., Lilja, K., Linderoth, M., Lundeberg, T., Fjällborg, B., Porsbring, T., Larsson, D.J., Bengtsson-Palme, J., Förlin, L., Kienle, C., Kunz, P., Vermeirssen, E., Werner, I., Robinson, C.D., Lyons, B., Katsiadaki, I., Whalley, C., den Haan, K., Messiaen, M., Clayton, H., Lettieri, T., Carvalho, R.N., Gawlik, B.M., Hollert, H., Di Paolo, C., Brack, W., Kammann, U., Kase, R., 2015. The European technical report on aquatic effect-based monitoring tools under the water framework directive. Environ. Sci. Eur. 27 (1), 7. https://doi.org/10.1186/s12302-015-0039-4.
White, F.M., 2011. The potential cost of high-throughput proteomics. Sci. Signal. 4 (160) https://doi.org/10.1126/scisignal.2001813 pe8–pe8.
Whiteaker, J.R., Lin, C., Kennedy, J., Hou, L., Trute, M., Sokal, I., Yan, P., Schoenherr, R. M., Zhao, L., Voytovich, U.J., Kelly-Spratt, K.S., Krasnoselsky, A., Gafken, P.R., Hogan, J.M., Jones, L.A., Wang, P., Amon, L., Chodosh, L.A., Nelson, P.S., McIntosh, M.W., Kemp, C.J., Paulovich, A.G., 2011. A targeted proteomics–based pipeline for verification of biomarkers in plasma. Nat. Biotechnol. 29 (7), 625–634. https://doi.org/10.1038/nbt.1900.
Xiao, C., Zhang, Y., Zhu, F., 2021. Immunotoxicity of polychlorinated biphenyls (PCBs) to the marine Crustacean species, Scylla paramamosain. Environ. Pollut. 291, 118229 https://doi.org/10.1016/j.envpol.2021.118229.
Yao, C.-L., Ji, P.-F., Kong, P., Wang, Z.-Y., Xiang, J.-H., 2009. Arginine kinase from Litopenaeus vannamei: cloning, expression and catalytic properties. Fish Shellfish Immunol. 26 (3), 553–558. https://doi.org/10.1016/j.fsi.2009.02.012.
Zhang, Y., Li, D., Sun, B., 2015. Do housekeeping genes exist? PLoS One 10 (5), e0123691. https://doi.org/10.1371/journal.pone.0123691.