



HAL
open science

High-multiplexed monitoring of protein biomarkers in the sentinel *Gammarus fossarum* by targeted scout-MRM assay, a new vision for ecotoxicoproteomics

Julien Faugère, Duarte Gouveia, Sophie Ayciriex, Arnaud Chaumot, Christine Almunia, Adeline François, Jean Armengaud, Jérôme Lemoine, Olivier Geffard, Davide Degli Esposti, et al.

► To cite this version:

Julien Faugère, Duarte Gouveia, Sophie Ayciriex, Arnaud Chaumot, Christine Almunia, et al.. High-multiplexed monitoring of protein biomarkers in the sentinel *Gammarus fossarum* by targeted scout-MRM assay, a new vision for ecotoxicoproteomics. *Journal of Proteomics*, 2020, 226, pp.103901. 10.1016/j.jprot.2020.103901 . hal-02903090

HAL Id: hal-02903090

<https://hal.science/hal-02903090v1>

Submitted on 22 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 **High-multiplexed monitoring of protein biomarkers in the**
2 **sentinel *Gammarus fossarum* by Targeted Scout-MRM assay, a**
3 **new vision for ecotoxicoproteomics.**

4 Julien Faugere^{1*}, Duarte Gouveia^{2,3*}, Sophie Ayciriex¹, Arnaud Chaumot², Christine
5 Almunia³, Adeline François², Jean Armengaud³, Jérôme Lemoine¹, Olivier Geffard², Davide
6 Degli-Esposti², Arnaud Salvador^{1#}

7 * Equal contribution

8 ¹, Université de Lyon, Université Claude Bernard Lyon 1, Institut des Sciences Analytiques,
9 CNRS UMR 5280, 5 rue de la Doua, F-69100 Villeurbanne, France.

10 ², INRAE, UR RiverLy, Laboratoire d'écotoxicologie, F-69625 Villeurbanne, France.

11 ³, Laboratoire Innovations Technologiques pour la Détection et le Diagnostic (LI2D), Service
12 de Pharmacologie et Immunoanalyse (SPI), Commissariat à l'Energie Atomique, Bagnols-sur-
13 Cèze, France.

14 #Corresponding Author; E-mail: arnaud.salvador@univ-lyon1.fr

15 **Acknowledgments**

16 The authors thank the French Ministry of Higher Education, Research and Innovation
17 (Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation) for the PhD
18 fellowship of Julien Faugere. This work was supported by the French National Agency
19 Research (Young investigator Grant ANR-18-CE34-0008 and Grant ANR-18-CE34-0013),
20 the "Office Français de la Biodiversité", and the Ecophyto National Action Plan. The authors
21 also benefitted from the French GDR "Aquatic Ecotoxicology" framework which aims at
22 fostering stimulating scientific discussions and collaborations for more integrative
23 approaches. The authors are grateful to Véronique Gouy (INRAE), Matthieu Le Dréau
24 (INRAE), Christine Combe (Water Department of the City of Lons Le Saunier) for their help
25 in the definition of the field deployment study, and to Anaïs Valada-Mennumi for her help on
26 the sample preparation and data analysis.

27 **Declaration of Competing Interest**

28 The authors declare the following financial interests: Sciex company and J.L. filled a patent
29 describing Scout-MRM concept for targeted analysis by mass spectrometry.

30 **Keywords:** Mass spectrometry, Protein quantification, Scout-MRM, ecotoxicology,
31 biomarker, multiplex approach.

1 **Abstract**

2 Ecotoxicoproteomics employs mass spectrometry-based approaches centered on proteins of
3 sentinel organisms to assess for instance, chemical toxicity in fresh water. In this study, we
4 combined proteogenomics experiments and a novel targeted proteomics approach free from
5 retention time scheduling called Scout-MRM. This methodology will enable the measurement
6 of simultaneously changes in the relative abundance of multiple proteins involved in key
7 physiological processes and potentially impacted by contaminants in the freshwater sentinel
8 *Gammarus fossarum*. The development and validation of the assay were performed to target
9 157 protein biomarkers of this non-model organism. We carefully chose and validated the
10 transitions to monitor using conventional parameters (linearity, repeatability, LOD, LOQ).
11 Finally, the potential of the methodology is illustrated by measuring 277-peptide-plex assay
12 (831 transitions) in sentinel animals exposed *in natura* to different agricultural sites
13 potentially exposed to pesticide contamination. Multivariate data analyses highlighted the
14 modulation of several key proteins involved in feeding and molting. This multiplex-targeted
15 proteomics assay paves the way for the discovery and the use of a large panel of novel protein
16 biomarkers in emergent ecotoxicological models for environmental monitoring in the future.

17 **Biological significance**

18 The study contributed to the development of Scout-MRM for the high-throughput quantitation
19 of a large panel of proteins in the *Gammarus fossarum* freshwater sentinel. Increasing the
20 number of markers in ecotoxicoproteomics is of utmost interest to assess the impact of
21 pollutants in freshwater organisms. The development and validation of the assay enabled the
22 monitoring of a large panel of reporter peptides of exposed gammarids. To illustrate the
23 applicability of the methodology, animals from different agricultural sites were analysed. The
24 application of the assay highlighted the modulation of some biomarker proteins involved in
25 key physiological pathways, such as molting, feeding and general stress response. Increasing
26 multiplexing capabilities and field test will provide the development of diagnostic protein
27 biomarkers for emergent ecotoxicological models in future environmental biomonitoring
28 programs.

29 **1 Introduction**

30 Ecotoxicoproteomics is being increasingly used in environmental hazard identification,
31 through the monitoring of protein expression in sentinel organisms exposed to environmental
32 pollutants in both laboratory and field studies [1-7]. Dynamic changes in the molecular
33 machinery of an organism subjected to a toxic stress are the starting points of its physiological
34 response [8]. Molecular biomarkers are therefore able to provide us with early diagnostics of
35 adverse effects in comparison with other higher-level endpoints such as reproductive
36 impairments or other physiological biomarkers.

37 Among the different mass spectrometry (MS) based strategies available for protein analysis,
38 shotgun proteomics using data-dependent acquisition (DDA) remains the most popular for
39 proteome discovery in ecotoxicological models. In this acquisition mode, a fixed number of
40 precursor ions is selected in the MS1 survey scan, followed by a sequential isolation and
41 fragmentation of the N most intense precursors. This acquisition mode is adapted for high-
42 throughput studies for protein discovery but presents some limitations regarding its low
43 reproducibility and inability to identify and quantify low-abundant proteins. DDA-based
44 studies have recently proved to be extremely useful for assessing the impact of several model
45 pollutants in the molecular machinery of sentinel organisms from aquatic ecosystems,
46 identifying exposure fingerprints that inform about the pollutants modes of action, and
47 highlighting potential toxicity biomarkers [1-3, 7, 9-13]. Nevertheless, despite the increasing
48 number of publications proposing new toxicity biomarkers, few are being considered for use
49 in routine environmental biomonitoring. This is mainly due to the lack of high-throughput
50 quantitative assays available for their verification and validation before its implementation in
51 operational monitoring programs.

52 Alternatively, targeted acquisition methods such as Multiple Reaction Monitoring (MRM) or
53 Parallel Reaction Monitoring (PRM), avoid the lack of reproducibility and limited
54 quantitative power of DDA methods, by focusing the MS/MS scans on a subset of
55 predetermined target peptides. These methods require *a priori* knowledge of the elution time
56 windows of the targeted peptides and the precursor-product ion transitions obtained from a
57 spectral library (subsequent to DDA analysis). MRM assays were proposed in recent years as
58 a promising tool for specific multi-biomarker measurements in environmental biomonitoring
59 [14-16]. However, targeted data acquisition experiments such as MRM or PRM exhibit some
60 restraints. One of the limitations of this approach is the number of transitions monitored per

61 peptide restricted by the duty cycle to keep an acceptable signal-to-noise ratio. In addition, the
62 development of large multiplexed assays becomes rather complex because of the RT reliance
63 or unwanted retention time shift due to sample matrix effects. More recently, a new MRM-
64 based targeted method, namely Scout-MRM has been proposed to increase the multiplexing
65 capability and the robustness of classic targeted approaches [17-19]. Briefly, Scout-MRM is
66 based on the successive monitoring of complex transition groups triggered by Scout peptide
67 signals distributed along the chromatogram. This method is completely independent from RT
68 and consequently of time scheduling, thereby increasing the multiplexing capability and
69 facilitating the analytical transfer between laboratories.

70 Herein, we develop for the first time in aquatic ecotoxicology Scout-MRM for the high-
71 throughput quantitation of a list of key proteins resulting from a proteogenomics study in
72 *Gammarus fossarum*, an aquatic model organism used as sentinel to assess freshwater
73 pollution [20, 21]. Increasing multiplexing capabilities is of great importance for the
74 development of biomarkers in ecotoxicology since it allows monitoring a broader list of
75 candidate biomarkers and validating a higher number of reliable surrogate peptide biomarkers
76 for developing absolute quantification assays. If robust enough, the simultaneous
77 quantification of hundreds of peptides and proteins also allows performing shotgun-like
78 protein network and/or co-expression analysis, with the advantage of targeting a sub-
79 proteome covering only the functions of interest. The development and validation of Scout-
80 MRM assay to monitor a large panel of protein's reported peptides from the emergent
81 ecotoxicological model *G. fossarum* is presented. We also demonstrate the interest of the
82 methodology in ecotoxicological studies through an application with active biomonitoring in
83 an agricultural pollution context, followed by a concise discussion around the advantages and
84 innovations of this methodology for environmental monitoring.

85 **2 Materials and Methods**

86 **2.1 Reagents and chemicals**

87 Water, acetonitrile (ACN) and methanol (MeOH) were obtained from Fisher Scientific (LC-
88 MS grade, Strasbourg, France). EDTA, triton X-100, iodoacetamide (IMA), dithiothreitol
89 (DTT), formic acid (FA), sodium chloride, aprotinin, ammonium bicarbonate (AMBIC),
90 leupeptin, trypsin (treated TCPK from bovine pancreas), ethylic ether and absolute ethanol
91 were obtained from Sigma Aldrich (St Quentin-Fallavier, France). Isotopically labelled
92 peptides containing either a C-terminal [¹⁵N₂ and ¹³C₆] lysine or arginine were synthesized by
93 Thermo Fisher Scientific (purity > 97 %) and stored at -20°C until use.

94 **2.2 Collection and maintenance of *G. fossarum* organisms**

95 Gammarids were collected by kick sampling from the Pollon river in France, and acclimatized
96 to laboratory conditions, as previously described [15, 16, 22, 23]. This sampling site contains
97 a gammarid population frequently used as a source of organisms for active biomonitoring
98 studies by the laboratory of ecotoxicology in INRAE [23, 24]. Before experiments, organisms
99 were kept for two weeks in 30L tanks continuously supplied with drilled groundwater, which
100 was adjusted with osmotic water to the same conductivity and pH values as the sampling site.
101 The temperature of the water was maintained at 12 ± 1°C, with 16/8h light/dark photoperiods.
102 Organisms were fed *ad libitum* with water-conditioned alder leaves (*Alnus glutinosa*) and
103 tubifex were added once a week.

104 **2.3 Sample preparation for shotgun proteomics**

105 Whole-body of five male and five female gammarids were disrupted in extraction buffer (50
106 mM Tris-Base, 100 mM NaCl, 1 mM EDTA, 0.1% v/v Triton X-100, 6 M Urea, protease
107 inhibitor cocktail) with a Tissue ruptor device (Qiagen). Homogenates were centrifuged at
108 10,000g for 7 minutes, and the supernatant transferred to a new tube. 30 µL of homogenate
109 were subjected to a short SDS-PAGE migration at 200 V (roughly 3 minutes). The whole
110 protein content from each sample in the gel was cut and processed with Trypsin Gold
111 (Promega) and 0.011% ProteaseMax surfactant (Promega) as previously described [25].

112 2.4 Sample preparation for Scout-MRM assay

113 Total protein content from gammarids was extracted according to previous published works
114 [14, 26]. 10 μ L of heavy peptides at 4 μ g/mL and 10 μ L of Scout peptides at 5 μ g/mL were
115 then added. After a SPE clean-up and evaporation, sample was resuspended with 90 μ L of a
116 H₂O/ACN mixture (90:10, v/v) with 0.5% FA, and centrifuged at 12 000 rpm for 5 min at RT,
117 before the LC-MS/MS analysis [14, 26].

118 2.5 Standards (SIL peptides, Scout peptides) preparation

119 Isotopically enriched peptide stock solution was prepared from lyophilized peptides dissolved
120 in water/ACN mixture (50:50, v/v) with 0.5% FA to obtain a 400 μ g/mL stock solution. Scout
121 peptides used in this study are stable isotope labelled peptides, selected to trigger the analysis
122 of MRM transition groups with Scout-MRM. Solutions containing isotopically enriched
123 peptides were prepared from stock solutions and diluted to obtain the desired concentrations.
124 The list of the labelled peptides and Scout peptides are shown in supplementary information
125 Table S1.

126 2.6 Nano LC-MS/MS analysis on high resolution mass spectrometry

127 Extracted and digested proteins were analysed through data-dependent acquisition mode on a
128 Q-Exactive HF mass spectrometer (Thermo Fisher) coupled to an Ultimate 3000 LC system
129 (Dionex-LC Packings). Tryptic peptides were first desalted on a reversed-phase PepMap 100
130 C₁₈ μ -precolumn, and separated on a nanoscale PepMap 100 C₁₈ nanoLC column (3 mm, 100
131 \AA , 75 μ m i.d. 50 cm, Thermo Fisher Scientific) with a 90 minute gradient of ACN with 0.1%
132 formic acid, and a flow rate of 0.3 μ L/min. Full MS were acquired from 350 to 1 800 m/z and
133 the 20 most abundant precursor ions were selected for fragmentation with 10 s dynamic
134 exclusion time. Only ions with +2 or +3 charges were selected for MS/MS analysis.

135 2.7 Liquid chromatography and targeted mass spectrometry

136 LC-MS/MS analysis was performed on an 1290 series HPLC device (Agilent Technologies,
137 Waldbronn, Germany) coupled to a QTRAP® 5500 LC/MS/MS System hybrid triple
138 quadrupole/linear ion trap mass spectrometer (Applied Biosystems/ MDS Analytical
139 Technologies, Foster City, CA, USA) equipped with a Turbo V™ ion source. The LC
140 separation of the 20 μ L injected sample was carried out on an Xbridge C₁₈ column (100 mm \times

141 2.1 mm, particle size 3.5 μm) with a symmetry C_{18} guard column (2.1 mm x 10 mm, particle
142 size 3.5 μm) from Waters (Milford, MA, USA). Elution was performed at a flow rate of 300
143 $\mu\text{L}/\text{min}$ with water containing 0.1% formic acid as eluent A and acetonitrile containing 0.1%
144 formic acid as eluent B, employing an isocratic gradient from the beginning at 5% B to for 2
145 min, followed by a linear gradient from 5% B to 35% B in 36 min. Column was washed at
146 100% B for 5 min and re-equilibrated at 5% B for 5 min. The injection duty cycle was 48 min,
147 considering the column equilibration time. Instrument control, data acquisition, and
148 processing were performed using a modified Analyst 1.6.2 software[®]. For academic research,
149 Scout-MRM provisional software patch is available on request from Sciex company (contact:
150 yves.leblanc@sciex.com). MS analysis was performed in positive ionization mode using an
151 ion spray voltage of 5500 V. The curtain gas flows and the nebulizer were set at 50 psi using
152 nitrogen. The TurboV[™] ion source was set at 550 °C with the auxiliary gas flow (nitrogen)
153 set at 40 psi. The software Skyline v4.1 (MacCoss Lab Software, USA) was used to product a
154 list of suitable SRM transitions. The mass in Q1 and Q3 as well as the collision energy (CE)
155 and the declustering potential (DP) values were predicted using Skyline for endogenous
156 peptides when the associated heavy peptides were not available. The duty cycle is set at 1.52 s
157 in order to attain fifteen points per chromatographic peak for each MRM transition in one of
158 the thirteen Scout groups. For peptides with corresponding associated heavy peptides, their
159 mass parameters have been optimized by direct infusion into the mass spectrometer. From the
160 initial set of candidates SRM transitions, three transitions by peptides were selected for the
161 final assay. Details of parameters are reported in Table S2.

162 2.8 MS/MS spectra interpretation from GFOSS database

163 MS/MS spectra were assigned to peptide sequences by searching against a customized
164 RNAseq-derived database GFOSS, which was previously published [27]. This database
165 contains 1,311,444 putative protein sequences and 289,084,257 residues. The algorithm from
166 the MASCOT Daemon v2.3.2 search engine (Matrix Science) was used for database search
167 and spectral matching. The following parameters were used: 5 ppm peptide tolerance and 0.02
168 Da MS/MS fragment tolerance, +2 or +3 peptide charge, a maximum of two missed
169 cleavages, carbamidomethylation of cysteine as fixed modification, oxidation of methionine
170 as variable modification, and trypsin as proteolytic enzyme. Mascot results were parsed with
171 IRMa 1.31.1c software [28]. Peptide-spectrum matches presenting a MASCOT peptide score
172 with a p-value of <0.05 were filtered and assigned to a protein according to the parsimony

173 principle. The data have been uploaded to the ProteomeXchange Consortium via the PRIDE
174 partner repository with the dataset identifier PXD017017 and 10.6019/PXD017017 [29].

175 2.9 Candidate list for Scout-MRM assay development

176 From the peptide and protein lists obtained in the DDA mode, peptides that did not match the
177 requirements needed for the targeted approach were filtered out of the final list of candidates.
178 At the precursor level, the parameters chosen for selection were the following: idotp >0.92
179 (isotope dot product), mass error ± 4 ppm, ≥ 4 peptides per protein. Only the four most intense
180 precursors (based on MS peak area) per protein were kept. At the MS/MS level, all product
181 ions with m/z higher than 1 050 were eliminated, and a maximum of 4 product ions per
182 precursor were kept (the most intense ones). The list was further reduced by removing
183 proteins with whose functional annotations suggested little interest to be used as toxicity
184 biomarkers (mostly housekeeping and orphan proteins). A final list of 2,942 transitions, 919
185 peptides, and 263 proteins was obtained (Table S3).

186 2.10 Method validation for Scout-MRM assay

187 Verification of MRM transitions free from interferences and the correlation of peptide ratios
188 were performed using a correlation study of transition ratios per peptide. The correlation
189 study was carried out on 46 male and 24 female gammarids at different stages of
190 reproduction. A peptide measured with three MRM transitions is considered as non-interfered
191 if at least one of the transition ratios has a RSD < 20%. The remaining analytical parameters
192 were performed from a mixture of male and female *G. fossarum* in order to have samples
193 containing the same amount of protein digest as background matrix. The repeatability of the
194 method including SPE, evaporation under nitrogen flow and MS measurement were evaluated
195 by performing analyses of independent *G. fossarum* samples (n=3) spiked with 500 ng/mL of
196 heavy peptides. The repeatability of the entire analytical protocol including the digestion step
197 for endogenous peptides was evaluated by analysing independent samples (n=10) from the
198 same gammarid pool. Matrix effects, SPE recoveries and evaporation recoveries were
199 evaluated by preparing *G. fossarum* sample spiked with heavy peptides at different protocol
200 steps at a concentration of 500 ng/mL (n=3). The different recoveries (after SPE and
201 evaporation) and matrix effect are calculated according to the following formulas:

202 SPE recovery (%) = (Heavy peptide area add before SPE / Heavy peptide area add after SPE)
203 x 100

204 Evaporation recovery (%) = (Heavy peptide area add after SPE / Heavy peptide area add after
205 evaporation) x 100

206 Matrix effect (%) = ((Heavy peptide area in matrix / Heavy peptide area add in solvent)-1) x
207 100

208 In order to carry out the linearity study of heavy peptides, as well as the evaluation of
209 detection limits (LOD) and quantification limits (LOQ), 9 pools of *G. fossarum* peptide
210 solutions were spiked with an increasing quantity of heavy peptides from 1 to 10 000 ng/mL.
211 Each calibration curve has at least 5 levels of concentration (between 1 and 10000 ng/mL, 10
212 levels of concentrations were investigated) with an accuracy between 80 and 120%. In
213 addition, a weighted least-square linear regression was used ($1/x^2$). To determine the LOD, we
214 measured the signal-to-noise ratio in the detection area of the chromatographic peak and
215 considered the 3:1 ratio as acceptable to define it. The LOQ is defined as being equal to three
216 times the LOD. Protein extraction and digestion kinetics in *G. fossarum* have already been
217 optimized and described in a previous publication [14].

218 2.11 Data analysis

219 The integration of chromatographic peaks and data reprocessing was performed with
220 MultiQuant™ software (version 2.1.1, Sciex). Peak areas for each peptide were log₂
221 transformed to perform multivariate analyses. In order to identify possible outliers, sample
222 clustering was performed using the hierarchical clustering function implemented in the lumi R
223 package [30]. Principal Component Analysis (PCA) was used to analyse and identify the
224 variables explaining the maximum variance associated to the proteomic data in male
225 gammarids caged in the contaminated or the reference sites. Differential protein abundance
226 analysis was performed using linear models and empirical Bayes methods implemented in the
227 limma R package [31, 32]. Peptides showing differences with a BH adjusted p-value (FDR)
228 <0.1 are considered significantly different. Targeted proteomics data have been uploaded to
229 the PeptideAtlas SRM Experiment Library (PASSEL) under dataset identifier PASS01501.

230 2.12 Application of the assay

231 2.12.1 Selection and exposure of organisms

232 For each experiment, organisms were collected from the water tanks at specific reproductive
233 stages. For method validation, mature male and female organisms were sampled based on

234 visual observation of couples having a female in an advanced stage of the reproductive cycle
235 (well-developed embryos in the brood pouch as described in [20]). For the field studies,
236 couples of *G. fossarum* were sorted in with female in the final molting stages D2 and placed
237 by 7 in punctured polypropylene cylinders with alder leaves as food supply. The day after,
238 four cages of 7 couples were deployed in each study site during three weeks following the
239 reproductive bioassay protocol described in [33]. The organisms were then brought back to
240 the laboratory in water from the study site. Seven males by sites were weighted and directly
241 frozen in liquid nitrogen and stored at -80°C for the proteomic analysis.

242 2.12.2 Study sites

243 Organisms were caged in four distinct study sites in the “Jura” region in France for a three
244 week exposure. Three sites (A, I, Av) were localized along the river « La Madeleine »
245 (N 46°42'15"; E 5°31'9") which drains a mixed agricultural watershed (wine, corn and wheat
246 crops) before joining drinking water catchments of the City of Lons-le-Saunier. According to
247 the pesticide program commissioned by the water authorities of Lons-le-Saunier between
248 2011 and 2017, this watershed is at risk for pesticide contamination. A fourth study site
249 situated upstream of the “Seille” river (N 46°42'58"; E 5°37'54") was used as a reference site.
250 This station is located at less than 10 km of the east of the “Madeleine” watershed and is in a
251 good chemical and ecological status among the National Reference Network implemented for
252 the European Water Framework Directive.

253 **3 Results and discussion**

254 **3.1 Selection of proteins of interest from proteogenomics and DDA** 255 **analysis**

256 In the shotgun experiment, an average of 46,068 MS/MS spectra were recorded per sample,
257 from which 27-29% were attributed to peptide sequences after querying the GFOSS
258 proteogenomics database. These peptide-to-spectrum matches (PSMs) allowed the
259 identification of 5,298 protein sequences, which were functionally annotated through BLAST
260 searches and used as the starting point for the selection of the candidates for the targeted
261 assay. The *in silico* mining of this large list of proteins allowed to select a panel of 263
262 proteins according to the criteria described in section 2.9. These proteins are involved in key
263 physiological functions such as reproduction, immunity, homeostasis, and detoxification /
264 defence mechanisms, and are therefore susceptible to be disrupted by exposure to toxic
265 compounds. The list of the proteins is shown in Table S3.

266 **3.2 Reporter peptide identification and MS optimization in targeted MS** 267 **approach**

268 The list of 263 proteins was used to develop the Scout-MRM method on a triple quadrupole
269 mass spectrometer. This type of acquisition method is more suitable for quantification
270 because it provides more robustness. The specificity and sensitivity of targeted-based assays
271 depend on the suitable selection of the proteotypic peptides. For the unambiguous peptide
272 identification in low resolution targeted MS, four MRM transitions corresponding to the most
273 intense MS/MS fragment ion peptides were selected from proteogenomic experiments, which
274 correspond to 3, 676 transitions among 27 MRM methods. The chromatographic conditions of
275 the proteogenomic study and the Scout-MRM method are different. Indeed, nano-LC
276 configuration was not used for the development of Scout-MRM method. It was therefore
277 necessary to redefine the retention times of each peptide. RT information was used as a
278 reference for peptide detection in our system. A retention time was assigned to a peptide when
279 the monitored transitions were perfectly aligned.

280 As shown in Fig. 1A, the extracted ion chromatograms showed four overlapped and identical
281 MRM transitions for GTLAVIPVQNR and for their corresponding heavy peptide
282 GTLAVIPVQNR*. Fig. 1B shows extracted ion chromatogram of LQEQVADYK at 8.1

283 min with the 4 overlapped transitions that confirms the elution time of this peptide. In certain
284 cases, difficulties can be encountered for the determination of the RT of a proxy peptide that
285 exhibits 4 aligned transitions when an interference occurred and the corresponding labelled
286 peptide is not available. As shown in Fig. 1C, the extracted ion chromatogram exhibits several
287 peaks for which the 4 transitions for TDLSITLAER are perfectly aligned. To overcome this
288 problem *i.e.* to correctly identify the RT area of the target peptides, we established a
289 correlation curve between peptide RT obtained in targeted-MRM based assay and high-
290 resolution-MS² based experiment (Fig. 2 and Table S4). Even if peptide separation conditions
291 were very different (nanoLC *versus* microLC configuration, different stationary phase), the
292 RT of the target peptides have been easily identified from other interfering isobaric
293 compounds by using the formula of the regression curve. For 187 peptides, we had also to
294 corroborate our results by performing additional MRM runs by adding supplemental
295 transitions, *i.e.* 1879 transitions. Finally, after the optimization of the precursor and transition
296 selection, 341 peptides proxy of 182 proteins have been perfectly identified on the Q-Trap
297 system.

298 **3.3 Improvement of multiplexing capacity**

299 In a second step, we evaluated the capacity to configure a 182-plex MRM protein assay. This
300 assay requires following 1,155 transitions including 44 heavy labelled peptides (up to 3
301 peptides per protein, at least 3 transitions per peptides). Considering the large number of
302 transitions, it was not possible to develop a single MRM method. To solve this problem, an
303 acquisition method capable of segmenting the chromatogram is needed to focus the mass
304 spectrometer on specific areas associated with retention time of compounds of interest. For
305 this purpose, we can use a single time fragmentation method (called Scheduled MRM[®], Time
306 MRM[®] or Dynamic MRM[®] according to the MS supplier).

307 3.3.1 Principle and limitations of MRM methods with time window.

308 To guarantee a constant acquisition setting of 10-15 points per chromatographic peak in time
309 window scheduling experiment, MRM transitions of each peptide are centered on their
310 expected retention times. A constant cycle time is maintained with the simultaneous
311 adjustment of the dwell time according to the number of peptides co-eluted in the same
312 window. To ensure highly multiplexed analysis, RT scheduling methods must monitor the
313 transitions across small time segments. However, any unexpected RT shift out of the

314 scheduled RT window (modification in chromatographic conditions, change of volume delay
315 with distinct pumping systems during a method transfer, injection overload of the column due
316 to samples with different protein contents) can compromise assay robustness and its
317 implementation for large-scale analysis. As shown in Fig. 3A, an artificial modification in
318 HPLC conditions, *i.e.* decrease of the isocratic step at the beginning of the gradient, leads as
319 expected to a RT shift of QFQYTWR peptide during retention time scheduling method. The
320 chromatographic peak is truncated, and the quantification is difficult to achieve. In this case,
321 reducing the RT window detection proves to be detrimental to the multiplexed analysis. In
322 some cases, peptides may no longer be detected because they are outside the detection
323 windows.

324 3.3.2 Principle and advantages of Scout-MRM

325 To avoid partial or no peak detection, we introduced a new mode of targeted data acquisition
326 called Scout-MRM in order to rationalize the development of targeted proteomics assay and
327 to facilitate dissemination of ready-to-use methods [18, 34]. Scout-MRM relies on the
328 monitoring of complex transitions successively triggered under the detection of Scout
329 peptides. Compared with scheduled methods, the acquisition was triggered by Scout peptides
330 for each segment, instead of by pre-defined scheduled time windows (here in this case,
331 segment of 1.5 min). No extra adjustment for acquisition windows is needed. As we observed,
332 the peptide is correctly eluted during chromatographic variations (Fig. 3B). In practice, the
333 Scout peptides chosen in this study are 12 stable isotope labelled peptides that can be used for
334 relative quantification (Table S1). Scout peptides triggering groups of transitions (13 groups),
335 are regularly dispatched all along the chromatogram (Fig. 4). When the intensity of the MRM
336 transition of the first Scout exceeds a defined threshold, the monitoring of a transition group is
337 triggered. The follow-up of the group stops when the second Scout peptide is detected and
338 consequently the second group is triggered and so on [18, 34] (Fig. 4). Any incidental RT
339 shift is completely without consequence on the target detection of peptides. To ensure highly
340 multiplexed analysis, RT scheduling method monitors the transitions across large segments
341 generally set to less than 1 min that ensures a compromise between multiplexing capacity and
342 robustness of the method [34]. In Scout-MRM it is sufficient to distribute Scouts with an
343 interval of less than one minute to have a higher gain in multiplexing capacity without loss of
344 robustness. To increase multiplexing capacity for future analysis, it is therefore enough to
345 increase the number of Scout groups in the method.

346 **3.4 Scout-MRM method optimisation**

347 3.4.1 Biological validation of MRM transitions

348 To ensure that the selected MRM transitions were not interfered in the different biological
349 conditions, we verified that transition ratios from a specific peptide are constant. In order to
350 be most relevant several samples need to be analyzed. From the samples available at the
351 laboratory at the time of the study, forty-six adult male sampled from different locations and
352 24 adult female organisms at distinct reproductive stages were used to include as much
353 variability as possible. MRM transitions were considered as non-interfered when the ratio
354 between 3 transitions remains constant regardless the sample. From our data we have
355 performed transition area ratio calculation for all samples and estimation of relative standard
356 deviation values for each peptide. A threshold above 20% of MRM ratio is considered as
357 interfered. As shown in Fig. 5, the peptide CQLFNDPSDR exhibits different transition ratios
358 between y8/y7, ions with a RSD > 20%, in two different biological conditions (different male
359 gammarids from different sites). On the contrary, the tryptic peptide SLVNLGDVQEGK
360 conserves the transition ratio between the different ions y9/y8, y8/y7 and y9/y7 with a RSD <
361 20% [35]. After evaluation of all the MRM ratio transitions (341 peptides proxy for 182
362 proteins), we kept 277 peptides proxy for 157 proteins (Table S2). As a result, some proteins
363 are only reported by one or two peptides in the method. More precisely, 67 proteins with 1
364 peptide, 42 proteins with 2 peptides and 30 proteins with 3 peptides were followed in male
365 gammarids. For female gammarids, 76 proteins with 1 peptide, 47 protein with 2 peptides and
366 25 proteins with 3 peptides per protein were monitored.

367 3.4.2 Biological validation of peptides

368 A second study between the different peptides of the same protein was carried out in male and
369 female *G. fossarum*. Indeed, peptides from the same protein must have constant area ratios
370 between the different samples analysed, as shown in Fig. 6. This ratio is calculated from the
371 most intense and least interfering transition areas determined in the first correlation study. If
372 the ratios are not respected, this may be due for example to the presence of a modification on
373 one of the peptides of the protein or different matrix effects. Peptide ratios for the same
374 protein are more uncertain than transition ratios of a peptide. Indeed, the different transitions
375 of the same peptide will undergo the same matrix effects because they are eluted at the same
376 retention time while the peptides of the same protein will have different retention times and

377 can therefore undergo different matrix effects between different samples. This criterion was
378 therefore not used to remove a transition from the method. Rather, it has been used as an
379 indicative value to know which peptides within the protein correlate best with each other. The
380 results obtained for our correlation study between the different peptides of the same protein
381 are summarized in Fig. 7. These results show that for most proteins we have a good
382 correlation between the different peptides.

383 **3.5 Analytical performance evaluation of the assay for quantification**

384 As the multiplexed protein assay will be used to support future ecotoxicology studies, the
385 analytical performances of the assay for quantification were assessed. With our scout-MRM
386 method, an absolute or relative quantification can be considered. Absolute quantitation is
387 achieved with incorporation for each peptide of isotopically labelled synthetic peptide internal
388 standards. Disposing of all the peptides labelled with high purity results in substantial costs
389 that are often incompatible with the financial means of ecotoxicological studies. In a first step,
390 a relative quantification is performed. Only 44 labelled peptides have been synthesized and
391 will be used for performance evaluation. Since relative quantification method is based on
392 sample comparison, it must first be ensured that the analytical protocol is repeatable.
393 Therefore, a repeatability study was carried out.

394 3.5.1 Repeatability

395 Intraday repeatability was assessed by spiking heavy peptides to a pool of gammarids at a
396 concentration of 500 ng/mL. All the results are presented in Table S1 as supplementary data.
397 The relative standard deviations (RSD) for this repeatability study is between 1 and 12% that
398 shows good repeatability. However, since we do not have the counterparts marked for each
399 peptide, we have carried out another repeatability study by analysing endogenous peptides
400 from ten independent extractions to consider all the peptides. Furthermore, this study also
401 takes into account the reproducibility of protein digestion into peptides. All the results are
402 presented in Table S5 as supplementary data. Fig. 8 shows that 90.2% of MRM transitions
403 have an RSD of less than 20%. This threshold represents the limit values that have been set to
404 define whether a transition is repeatable. The RSD obtained for all previously selected
405 qualifying MRM transitions are less than or equal to 20%. When the RSD is greater than
406 20%, this corresponds to the lowest MRM transitions that generate more difficult integrations
407 and are more easily interfered. These results indicate that the digestion step and extraction
408 step are controlled and reproducible.

409 3.5.2 Extraction recoveries

410 It was shown in a previous paper that peptides can be lost during evaporation [26]. Therefore,
411 this step was also specifically evaluated in addition to the SPE extraction recovery. To
412 estimate extraction recoveries, digested protein extracts from gammarids were spiked with
413 labelled peptides before and after SPE, and after evaporation at a concentration of 500 ng/mL
414 (n=3) and compared. The results are presented in Table S1. It can be observed that SPE
415 recoveries are between 72 and 122% with an average of 96% for the 44 heavy peptides.
416 Evaporation recovery range from 67 to 106% with an average of 94% for the 44 heavy
417 peptides. These results showed a good extraction recovery.

418 3.5.3 Matrix effect evaluation

419 Matrices effects in MS correspond to the influence of coeluted compounds during their
420 simultaneous introduction into the source due to ionization competition. This can result in an
421 increase or a decrease of their intensity. To measure these matrix effects, sample extracts were
422 spiked with heavy peptides added just before injection at a concentration of 500 ng/mL (n=3)
423 which were compared to samples containing the same quantity of heavy peptides spiked in
424 reconstitution solvent (n=3). The results are presented in Table S1. The 44 heavy peptides
425 undergo matrix effects that induce a loss or gain of the signal between -67% and 51% with an
426 average loss of 28%. As expected, matrix effects occur. However, the case presented here
427 illustrates the most extreme case where a very complex mixture is compared to compounds in
428 solution. The analysis of heavy peptides spiked at the same concentration in biologically
429 different *G. fossarum* extracts showed that even if matrix effects affect the peptides analysed,
430 these effects are comparable from one sample to another. Indeed, the results presented in
431 Fig.S1. obtained from the data in table S6 show that for 93% of heavy peptides the RSD are
432 less than 20%. To correct different matrix effects between different samples, internal
433 standards (heavy peptides) similar to endogenous peptides must be used, which will undergo
434 the same matrix effects as the compounds of interest. In our case, we do not have all the
435 heavy peptides corresponding to the 277 peptides selected in the method. Our study aims to
436 highlight trends by identifying potential biomarkers. Once the candidates have been found, it
437 will be necessary to have their counterparts marked to carry out the most accurate
438 quantification possible, considering the matrix effects. This correction of matrix effects will
439 also allow to get peptide ratios for the same protein as precise as possible.

440 3.5.4 Linearity and limit of quantification (LOQ)

441 Although this study is based on a relative quantification, we have determined the linearity
442 domains and LOQ of our heavy peptides in order to evaluate the performance of the method.
443 A weighted least-square linear regression ($1/x^2$) was used for the calibration. All calibration
444 curves corresponding to heavy peptides spiked into the *G. fossarum* matrix at different levels
445 of concentration confirmed the high degree of linearity ($0.991 < r^2 < 0.999$) (Table S1 and Fig.
446 S2). The LOD was determined to be 3 times greater than the background noise ($S/N=3$) and
447 the LOQ is determined as 3 times the LOD. The LOQ are between 1.12 pmol/mL and 113
448 pmol/mL (Table S1). Finally, the method developed with Scout-MRM enabled to monitor 157
449 proteins in male and female species of *G. fossarum* following the validation criteria (157
450 proteins, 277 peptides, and 831 MRM transitions) among the targeted 263 proteins considered
451 initially.

452 3.6 Application of Scout-MRM assay: comparative proteomic analysis 453 using active biomonitoring in an agriculture watershed

454 We applied SCOUT-MRM method to assess its interest and feasibility in the context of active
455 biomonitoring in freshwater streams. Twenty-eight calibrated male gammarids (7 from a non-
456 contaminated reference site, 7 from each of the three contaminated sites) were used for the
457 proteomics analysis. Two hundred sixty-five peptides were detected for each male organism.

458 To compare the global proteomes of the organisms caged in the contaminated sites or the
459 reference site, we performed clustering and principal component analyses. The cluster
460 analysis identified three organisms as potential outliers, based on their distance with the rest
461 of the samples (Figure 9A). These outliers (an organism in the reference site, R3; and two
462 organisms in two different contaminated sites, A CE3 and AV CE10) were excluded from the
463 following analyses. The PCA analysis showed that organisms are quite dispersed, even
464 though there is a tendency of the organisms caged in the reference site to cluster more closely
465 than those clustered in the contaminated sites (Figure 9B).

466 In order to identify peptides whose abundance is able to distinguish the organisms caged in
467 the different type of sites, and thus could be identified as potential biomarkers of exposure to
468 agriculture pollutants, a differential analysis was performed. The differential analysis
469 identified 10 peptides as differently expressed ($FDR < 0.1$) (Figure 9C). The peptide
470 DIDAAFLVGAMPR, annotated as cytosolic malate dehydrogenase or malic enzyme (ME)

471 was the most upregulated in organisms exposed in the contaminated sites (logFC=1.38,
472 FDR=0.038) (Figure 9D). The *Drosophila* ME activity can be induced by the juvenile
473 hormone (JH) by both a direct effect on the enzyme in the short term and the activation of its
474 gene (Men) transcription [36]. Since JH analogs are among the most used insecticides in
475 agriculture, the increased expression of a homolog of ME in *G. fossarum* might suggest a
476 response of these non-target organisms to JH analogs contamination in the observed sites.
477 Among the downregulated peptides, the most significant was QIDNPDYK (logFC=-0.86,
478 FDR=7.9*10⁻⁵) (Figure 9D), belonging to a homolog of calreticulin. Calreticulin is a highly
479 conserved endoplasmic reticulum protein of the lectin family. It is involved in
480 osmoregulation, molecular chaperoning and immune response in crustaceans, with most data
481 available from decapods models [37-39]. Calreticulin has been reported as a general stress
482 biomarker in *in vitro* and vertebrate models [40] and its downregulation may thus suggest a
483 stress effects in amphipods caged in the contaminated sites.

484

485 Among the downregulated peptides found in organisms caged in the contaminated site, we
486 found 2 peptides belonging to an homolog of endochitinase (EAFDTVGR, logFC=-0.61,
487 FDR=0.059; LVLGIPFYGR, logFC=-0.55, FDR=0.072) and 2 peptides belonging to a
488 protein annotated as endoglucanase (or cellulase A). (SAMNVALIAFK logFC=-0.76,
489 FDR=0.063; AADLGLDSTNNR, logFC=-1.28, FDR=0.06). Chitinases are key enzymes for
490 successfully complete molting cycle in arthropods [41]. We have previously observed its
491 decreased detection in *G. fossarum* exposed to contaminated sites [15]. Similarly, other
492 crustacean chitinases have been reported as sensitive biomarker to insecticides and fungicides
493 [41]. Cellulases are essential for digestion in *G. fossarum* and their decreased activity and/or
494 abundances may affect its reproductive capacity [39]. Moreover, cellulase activity were
495 reported to be very sensitive to insecticides and fungicides [39]. Overall, these results suggest
496 that multiple protein biomarkers in *G. fossarum* are useful to detect a biological response to a
497 contamination of agricultural origin. The limited number of modulated peptides in the
498 contaminated sites might suggest a certain specificity of the impact of the chemicals present
499 in the aquatic environment to which the gammarids were exposed. However, due to the
500 limited number of investigated sites, it will be crucial to extend the use of these multiplexed
501 protein biomarker approach to a larger spatial scale with contrasted contamination profiles.

502 **4 Conclusion**

503 A robust LC-MS/MS method for the simultaneous analysis of many potential protein
504 biomarkers in the sentinel species *G. fossarum* has been successfully developed and validated.
505 We applied a new acquisition mode called Scout-MRM recently developed in our laboratory
506 to significantly increase the multiplexing capacity through the implementation of a 157-
507 protein multiplex (277 peptides, 831 MRM transitions) in adult gammarids. Scout-MRM
508 provides a more robust method than acquisition mode using time window scheduling. Scout-
509 MRM is free from retention times, thus limiting the loss of information due to potential RT
510 shifts. The RT independency of Scout-MRM opens up the perspective to build large multiplex
511 by adding more scout peptides and MRM transitions if one wants to follow more proteins of
512 interest in the future. Indeed, since *G. fossarum* genome is still not fully characterized, new
513 discovered proteins can be implemented.

514 The first application of our method for biomonitoring key proteins to assess freshwater
515 pollution from different agricultural sites demonstrated the potential value of this
516 methodology in ecotoxicology studies. Indeed, the detection of protein reporter peptides with
517 modulation in response to stress shows that the Scout-MRM method is a relevant method to
518 detect biological responses due to contamination. Further studies will allow to identify
519 reference values for the investigated peptides and eventually fine-tuning the choice of more
520 specific reporter peptides of proteins related with different mode of actions involved in the
521 adverse outcome observed in the ecotoxicological bioassays used in the field [42, 43]. In
522 conclusion, Scout-MRM streamlines the development of targeted proteomics method in an
523 ecotoxicology study and simplifies dissemination of ready-to-use assays as they are easily
524 transferable from one laboratory to another.

525 In conclusion, this Scout-MRM method streamlines the development of targeted proteomics
526 method in an ecotoxicology study and simplifies dissemination of ready-to-use assays as they
527 are easily transferable from one laboratory to another.

528

529 **Figure captions**

530 **Fig.1. Identification of endogenous peptides from *G. fossarum* by MRM and selection of**
531 **transitions.** (A) Detection of endogenous peptides when labelled peptides are available.
532 Retention times and transition ratios must be identical. (B) Detection of endogenous peptides
533 when isotopically enriched peptides are not available. The four MRM transitions must be
534 aligned. (C) Detection and selection of peptide with an isobaric interference also having an
535 alignment of the four MRM transitions. The RT area in MRM mode was predicted based on
536 proteogenomics data.

537 **Fig. 2. Correlation curve plotting the retention time of peptides identified by**
538 **proteogenomics to the peptide retention time obtained in MRM.** Retention times of 66
539 peptides obtained in nanoLC-MS/MS (reversed-phase PepMap 100 C18 μ -precolumn,
540 nanoscale PepMap 100 C18 nanoLC column, 3 mm, 100 \AA , 75 μm i.d. 50 cm and Q-Exactive
541 HF) and LC-MS/MS (symmetry C18 guard column, 2.1 mm x 10 mm, particle size 3.5 μm ,
542 Xbridge C18 LC column, 2.1 mm, 10 cm, particle size 3.5 μm and Qtrap 5500) have been
543 reported. (Table S4).

544 **Fig. 3. Effects of RT shifts for targeted peptides monitored by Scheduled-MRM® versus**
545 **Scout-MRM method.** (A) When the chromatographic conditions change (decrease in the
546 isocratic level at the beginning of the gradient) a RT shift occurs and induces an offsetting of
547 the chromatographic peak from the detection window. (B) If any RT shift occurs in Scout-
548 MRM method, scout peptides automatically realigns the detection window.

549 **Fig. 4. Scout-MRM concept.** This acquisition mode consists in the detection of a first
550 compound called Scout. When the intensity of the Scout MRM transition exceeds a threshold
551 defined by the operator, the monitoring of an unordered transition group as well as the
552 monitoring of a new scout transition and so on until the last group is triggered by scout n. The
553 triggering of a new group ends the follow-up of the previous one.

554 **Fig. 5. Correlation study between 3 MRM transitions of the same peptide.** (A) Non-
555 constant transition ratio obtained for the 3 MRM transitions of CQLFNDPSDR peptide under
556 two different biological conditions. (B) Constant transition ratio obtained for 3 MRM
557 transitions the peptide SLVNLGDVQEGK under two different biological conditions. (C)
558 RSD obtained for the mean transition ratios of CQLFNDPSDR and SLVNLGDVQEGK

559 peptides in 46 different male *G.fossarum* samples. Only the 3 transitions of the peptide
560 SLVNLGDVQEGK are considered as non-interfered (RSD < 20%)

561 **Fig. 6. Area ratio between 3 peptides of the same protein.** Peptide area ratio between
562 different peptides of the same protein, aspartate aminotransferase and cytosolic malate
563 dehydrogenase, in 46 different male *G. fossarum* samples. For the aspartate aminotransferase,
564 the correlation of the 3 peptides shows good results (RSD < 20%) and for the cytosolic malate
565 dehydrogenase protein, the peptide area ratios show a lower correlation (RSD > 40%).

566 **Fig. 7. Assessment of the correlation study of peptide ratios for the same protein.** For
567 each gender are represented the percentages of protein that have at least one peptide ratio with
568 the indicated correlation coefficient. The red and blue curves represent the cumulative protein
569 percentages in function of peptide area ratios

570 **Fig. 8. Repeatability of endogenous peptides.** Average intensity of the MRM transitions
571 presents in the Scout-MRM method according to their repeatability. Results obtained after
572 independent extraction from the same pool of male and female *G. fossarum* (n=10). The
573 repeatability validation criterion is set at 20% which corresponds to 90.2% of the MRM
574 transitions of the method. The entire protocol method is therefore well repeatable. When the
575 criterion is not satisfied, this corresponds to the least intense MRM transition not used as a
576 qualifying transition (more easily interferable and more difficult to integrate) (Table S5)

577 **Fig. 9. Comparative proteomic analysis using active biomonitoring in an agriculture**
578 **watershed.** Analyses were performed using log₂ transformed peak areas for each peptide
579 identified in caged organisms. (A) Hierarchical clustering Samples over the red line show a
580 higher distance compared with the rest of the organisms, thus they were excluded from the
581 datasets. (B) Principal Component Analysis. Samples caged in the reference site (green) tend
582 to cluster together. (C) Heatmap of differentially expressed proteins in male gammarids caged
583 for three weeks in a reference site (R) and three sites under agriculture pressure (A, I, Av).
584 Colours represent abundance fold changes (orange over detection and blue down
585 regulation).(D) Boxplot representing the differential expression of the most deregulated
586 peptides (DIDAAFLVGAMPR; QIDNPDYK).

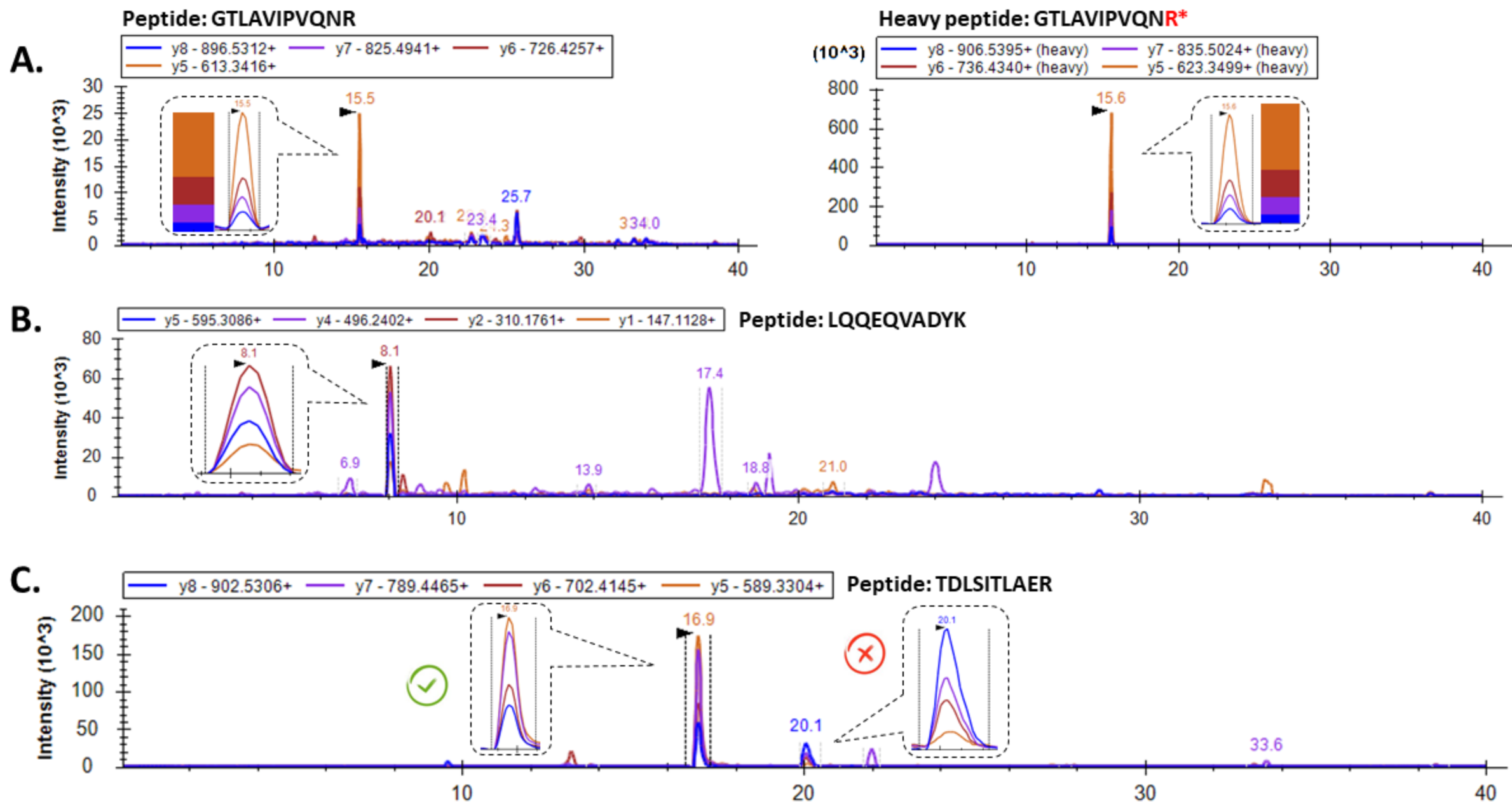
References

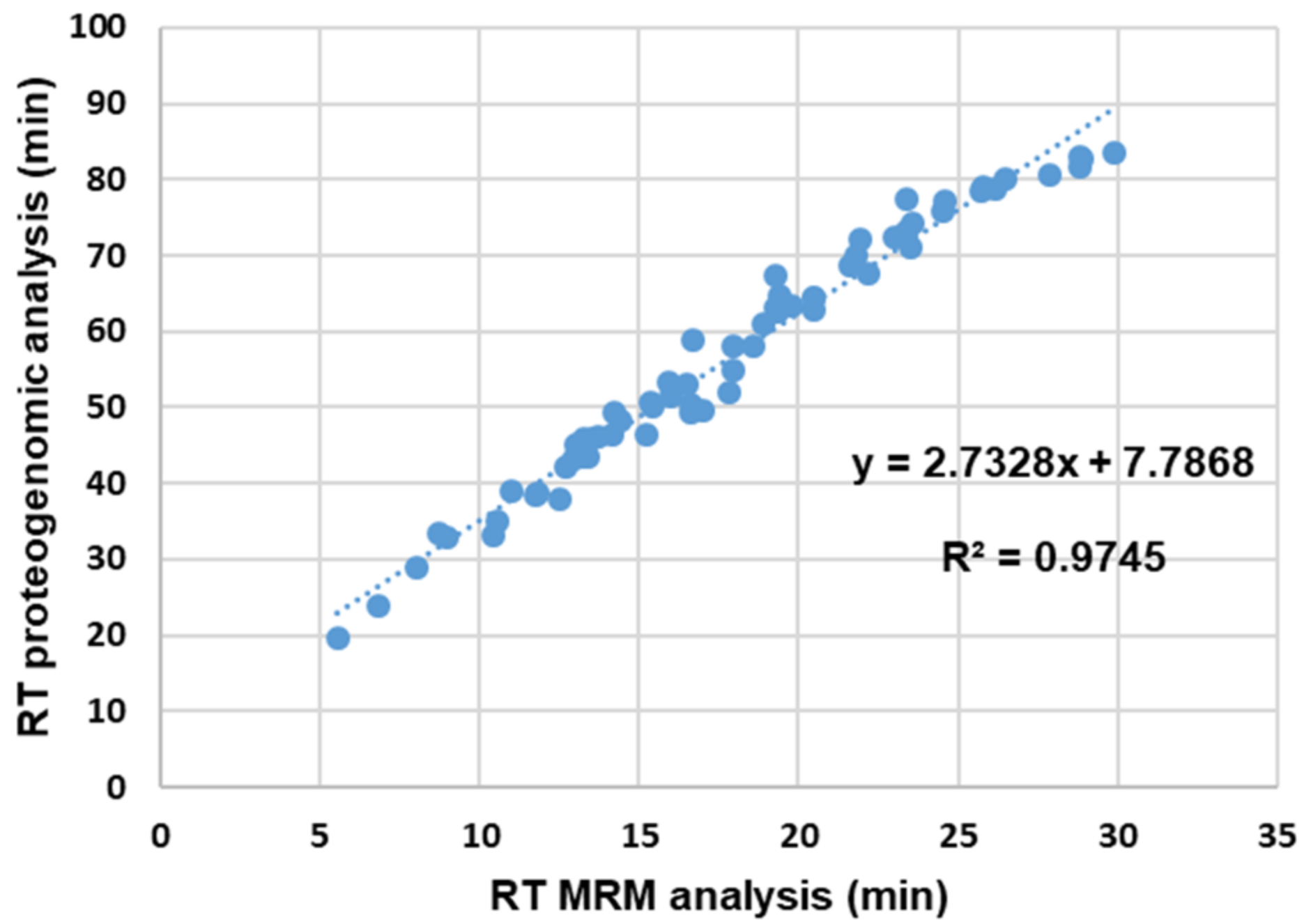
- [1] A. Lafontaine, D. Baiwir, C. Joaquim-Justo, E. De Pauw, S. Lemoine, C. Boulange-Lecomte, J. Forget-Leray, J.P. Thome, E. Gismondi, Proteomic response of *Macrobrachium rosenbergii* hepatopancreas exposed to chlordecone: Identification of endocrine disruption biomarkers?, *Ecotoxicol Environ Saf* 141 (2017) 306-314.
- [2] W. Schmidt, L.C. Rainville, G. McEneff, D. Sheehan, B. Quinn, A proteomic evaluation of the effects of the pharmaceuticals diclofenac and gemfibrozil on marine mussels (*Mytilus* spp.): evidence for chronic sublethal effects on stress-response proteins, *Drug Test Anal* 6(3) (2014) 210-9.
- [3] J. Trapp, J. Armengaud, O. Pible, J.C. Gaillard, K. Abbaci, Y. Habtoul, A. Chaumot, O. Geffard, Proteomic investigation of male *Gammarus fossarum*, a freshwater crustacean, in response to endocrine disruptors, *J Proteome Res* 14(1) (2015) 292-303.
- [4] V.A. Dowling, D. Sheehan, Proteomics as a route to identification of toxicity targets in environmental toxicology, *Proteomics* 6(20) (2006) 5597-604.
- [5] M.F. Lemos, A.C. Esteves, B. Samyn, I. Timperman, J. van Beeumen, A. Correia, C.A. van Gestel, A.M. Soares, Protein differential expression induced by endocrine disrupting compounds in a terrestrial isopod, *Chemosphere* 79(5) (2010) 570-6.
- [6] D. Gouveia, C. Almunia, Y. Cogne, O. Pible, D. Degli-Esposti, A. Salvador, S. Cristobal, D. Sheehan, A. Chaumot, O. Geffard, J. Armengaud, Ecotoxicoproteomics: A decade of progress in our understanding of anthropogenic impact on the environment, *J Proteomics* 198 (2019) 66-77.
- [7] F. Silvestre, J.F. Dierick, V. Dumont, M. Dieu, M. Raes, P. Devos, Differential protein expression profiles in anterior gills of *Eriocheir sinensis* during acclimation to cadmium, *Aquat Toxicol* 76(1) (2006) 46-58.
- [8] T.E. Allen, J.M. Goodman, S. Gutsell, P.J. Russell, Defining molecular initiating events in the adverse outcome pathway framework for risk assessment, *Chem Res Toxicol* 27(12) (2014) 2100-12.
- [9] M. Borgatta, C. Hernandez, L.A. Decosterd, N. Chevre, P. Waridel, Shotgun ecotoxicoproteomics of *Daphnia pulex*: biochemical effects of the anticancer drug tamoxifen, *J Proteome Res* 14(1) (2015) 279-91.
- [10] E. Gismondi, J.P. Thome, N. Urien, E. Uher, D. Baiwir, G. Mazzucchelli, E. De Pauw, L.C. Fechner, J.D. Lebrun, Ecotoxicoproteomic assessment of the functional alterations caused by chronic metallic exposures in gammarids, *Environ Pollut* 225 (2017) 428-438.
- [11] W. Hu, S. Tedesco, B. McDonagh, D. Sheehan, Shotgun redox proteomics in sub-proteomes trapped on functionalised beads: Identification of proteins targeted by oxidative stress, *Mar Environ Res* 69 Suppl (2010) S25-7.
- [12] M. Lepretre, C. Almunia, J. Armengaud, A. Salvador, A. Geffard, M. Palos-Ladeiro, The immune system of the freshwater zebra mussel, *Dreissena polymorpha*, decrypted by proteogenomics of hemocytes and plasma compartments, *J Proteomics* 202 (2019) 103366.

- [13] C.J. Martyniuk, K.J. Kroll, N.J. Doperalski, D.S. Barber, N.D. Denslow, Environmentally relevant exposure to 17 α -ethinylestradiol affects the telencephalic proteome of male fathead minnows, *Aquat Toxicol* 98(4) (2010) 344-53.
- [14] A. Charnot, D. Gouveia, J. Armengaud, C. Almunia, A. Chaumot, J. Lemoine, O. Geffard, A. Salvador, Multiplexed assay for protein quantitation in the invertebrate *Gammarus fossarum* by liquid chromatography coupled to tandem mass spectrometry, *Anal Bioanal Chem* 409(16) (2017) 3969-3991.
- [15] D. Gouveia, A. Chaumot, A. Charnot, C. Almunia, A. Francois, L. Navarro, J. Armengaud, A. Salvador, O. Geffard, Ecotoxic-Proteomics for Aquatic Environmental Monitoring: First in Situ Application of a New Proteomics-Based Multibiomarker Assay Using Caged Amphipods, *Environ Sci Technol* 51(22) (2017) 13417-13426.
- [16] G. Jubeaux, R. Simon, A. Salvador, C. Lopes, E. Lacaze, H. Queau, A. Chaumot, O. Geffard, Vitellogenin-like protein measurement in caged *Gammarus fossarum* males as a biomarker of endocrine disruptor exposure: inconclusive experience, *Aquat Toxicol* 122-123 (2012) 9-18.
- [17] S. Ayciriex, R. Carriere, C. Bardet, J. Blanc, A. Salvador, T. Fortin, J. Lemoine, Streamlined Development of Targeted Mass Spectrometry-Based Method Combining Scout-MRM and a Web-Based Tool Indexed with Scout Peptides, *Proteomics* 20(2) (2020) e1900254.
- [18] B. Rougemont, S. Bontemps Gallo, S. Ayciriex, R. Carriere, H. Hondermarck, J.M. Lacroix, J.C. Le Blanc, J. Lemoine, Scout-MRM: Multiplexed Targeted Mass Spectrometry-Based Assay without Retention Time Scheduling Exemplified by *Dickeya dadantii* Proteomic Analysis during Plant Infection, *Anal Chem* 89(3) (2017) 1421-1426.
- [19] A. Salvador, R. Carriere, S. Ayciriex, C. Margoum, Y. Leblanc, J. Lemoine, Scout-multiple reaction monitoring: A liquid chromatography tandem mass spectrometry approach for multi-residue pesticide analysis without time scheduling, *J Chromatogr A* 1621 (2020) 461046.
- [20] O. Geffard, B. Xuereb, A. Chaumot, A. Geffard, S. Biagiante, C. Noel, K. Abbaci, J. Garric, G. Charmantier, M. Charmantier-Daures, Ovarian cycle and embryonic development in *Gammarus fossarum*: application for reproductive toxicity assessment, *Environ Toxicol Chem* 29(10) (2010) 2249-59.
- [21] P.Y. Kunz, C. Kienle, A. Gerhardt, *Gammarus* spp. in aquatic ecotoxicology and water quality assessment: toward integrated multilevel tests, *Rev Environ Contam Toxicol* 205 (2010) 1-76.
- [22] J.P. Besse, M. Coquery, C. Lopes, A. Chaumot, H. Budzinski, P. Labadie, O. Geffard, Caged *Gammarus fossarum* (Crustacea) as a robust tool for the characterization of bioavailable contamination levels in continental waters: towards the determination of threshold values, *Water Res* 47(2) (2013) 650-60.
- [23] A. Ciliberti, A. Chaumot, R. Recoura-Massaquant, A. Chandris, A. Francois, M. Coquery, M. Ferreol, O. Geffard, Caged *Gammarus* as biomonitors identifying thresholds of toxic metal bioavailability that affect gammarid densities at the French national scale, *Water Res* 118 (2017) 131-140.
- [24] B. Alric, O. Geffard, A. Chandris, M. Ferreol, A. Francois, O. Perceval, J. Piffady, B. Villeneuve, A. Chaumot, Multisubstance Indicators Based on Caged *Gammarus* Bioaccumulation Reveal the Influence of Chemical Contamination on Stream Macroinvertebrate Abundances across France, *Environ Sci Technol* 53(10) (2019) 5906-5915.

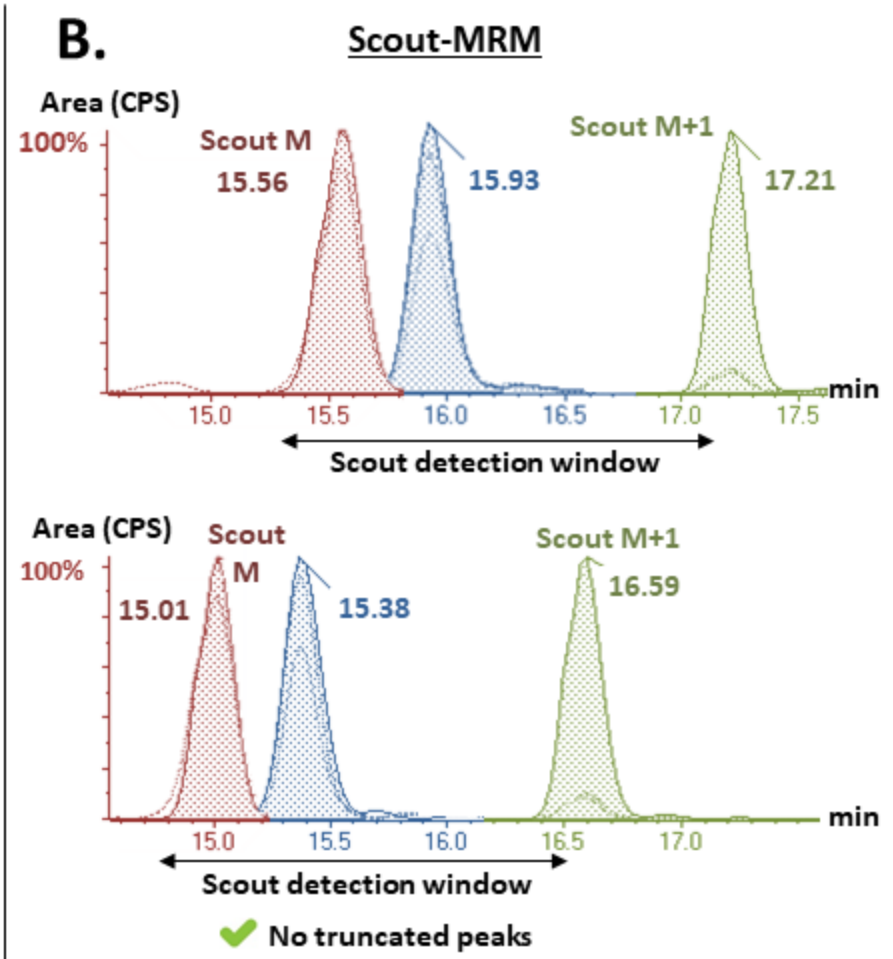
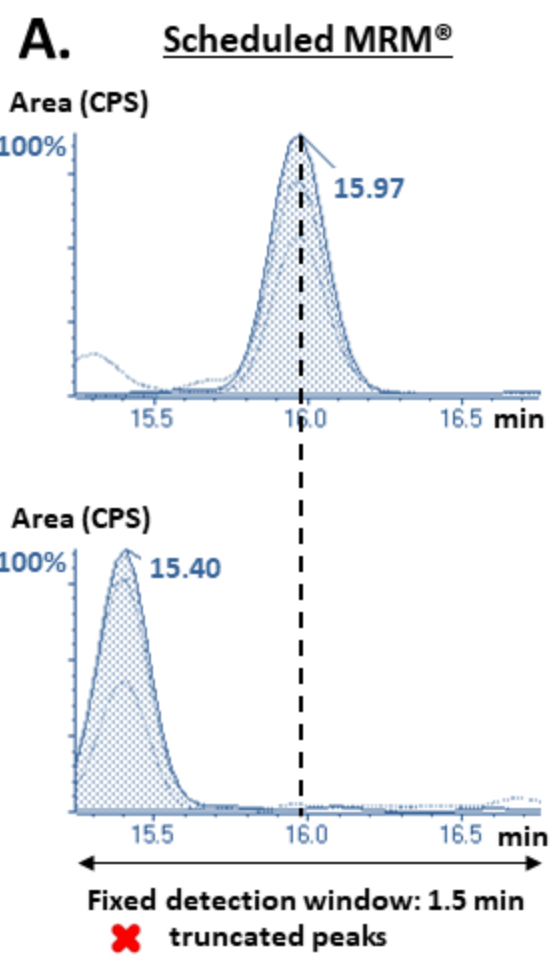
- [25] E.M. Hartmann, F. Allain, J.C. Gaillard, O. Pible, J. Armengaud, Taking the shortcut for high-throughput shotgun proteomic analysis of bacteria, *Methods Mol Biol* 1197 (2014) 275-85.
- [26] D.G. Aurore Charnot, Sophie Ayciriex, Jérôme Lemoine, Jean Armengaud, Christine Almunia, Arnaud Chaumot, Olivier Geffard, Arnaud Salvador, On-Line Solid Phase Extraction Liquid Chromatography-Mass Spectrometry Method for Multiplexed Proteins Quantitation in an Ecotoxicology Test Specie: *Gammarus fossarum*, *Journal of Applied Bioanalysis* 4(3) (2018).
- [27] J. Trapp, O. Geffard, G. Imbert, J.C. Gaillard, A.H. Davin, A. Chaumot, J. Armengaud, Proteogenomics of *Gammarus fossarum* to document the reproductive system of amphipods, *Mol Cell Proteomics* 13(12) (2014) 3612-25.
- [28] V. Dupierris, C. Masselon, M. Court, S. Kieffer-Jaquinod, C. Bruley, A toolbox for validation of mass spectrometry peptides identification and generation of database: IRMa, *Bioinformatics* 25(15) (2009) 1980-1.
- [29] J.A. Vizcaino, R.G. Cote, A. Csordas, J.A. Dianes, A. Fabregat, J.M. Foster, J. Griss, E. Alpi, M. Birim, J. Contell, G. O'Kelly, A. Schoenegger, D. Ovelheiro, Y. Perez-Riverol, F. Reisinger, D. Rios, R. Wang, H. Hermjakob, The PRoteomics IDentifications (PRIDE) database and associated tools: status in 2013, *Nucleic Acids Res* 41(Database issue) (2013) D1063-9.
- [30] D. Degli Esposti, C. Almunia, M.A. Guery, N. Koenig, J. Armengaud, A. Chaumot, O. Geffard, Co-expression network analysis identifies gonad- and embryo-associated protein modules in the sentinel species *Gammarus fossarum*, *Sci Rep* 9(1) (2019) 7862.
- [31] K. Kammers, R.N. Cole, C. Tiengwe, I. Ruczinski, Detecting Significant Changes in Protein Abundance, *EuPA Open Proteom* 7 (2015) 11-19.
- [32] G.K. Smyth, Linear models and empirical bayes methods for assessing differential expression in microarray experiments, *Stat Appl Genet Mol Biol* 3 (2004) Article3.
- [33] C. Lopes, A. Chaumot, B. Xuereb, R. Coulaud, G. Jubeaux, H. Queau, A. Francois, O. Geffard, In situ reproductive bioassay with caged *Gammarus fossarum* (Crustacea): Part 2 – evaluating the relevance of using a molt cycle temperature-dependent model as reference to assess toxicity in freshwater monitoring *Environmental Toxicology and Chemistry* (2020).
- [34] S. Ayciriex, R. Carriere, C. Bardet, J.C.Y. Blanc, A. Salvador, T. Fortin, J. Lemoine, Streamlined Development of Targeted Mass Spectrometry-Based Method Combining Scout-MRM and a Web-Based Tool Indexed with Scout Peptides, *Proteomics* 20(2) (2020) e1900254.
- [35] A.J. Percy, A.G. Chambers, J. Yang, C.H. Borchers, Multiplexed MRM-based quantitation of candidate cancer biomarker proteins in undepleted and non-enriched human plasma, *Proteomics* 13(14) (2013) 2202-15.
- [36] R. Farkas, P. Danis, L. Medved'ova, B.M. Mechler, J. Knopp, Regulation of cytosolic malate dehydrogenase by juvenile hormone in *Drosophila melanogaster*, *Cell Biochem Biophys* 37(1) (2002) 37-52.
- [37] Y. Duan, P. Liu, J. Li, Y. Wang, J. Li, P. Chen, Molecular responses of calreticulin gene to *Vibrio anguillarum* and WSSV challenge in the ridgetail white prawn *Exopalaemon carinicauda*, *Fish Shellfish Immunol* 36(1) (2014) 164-71.

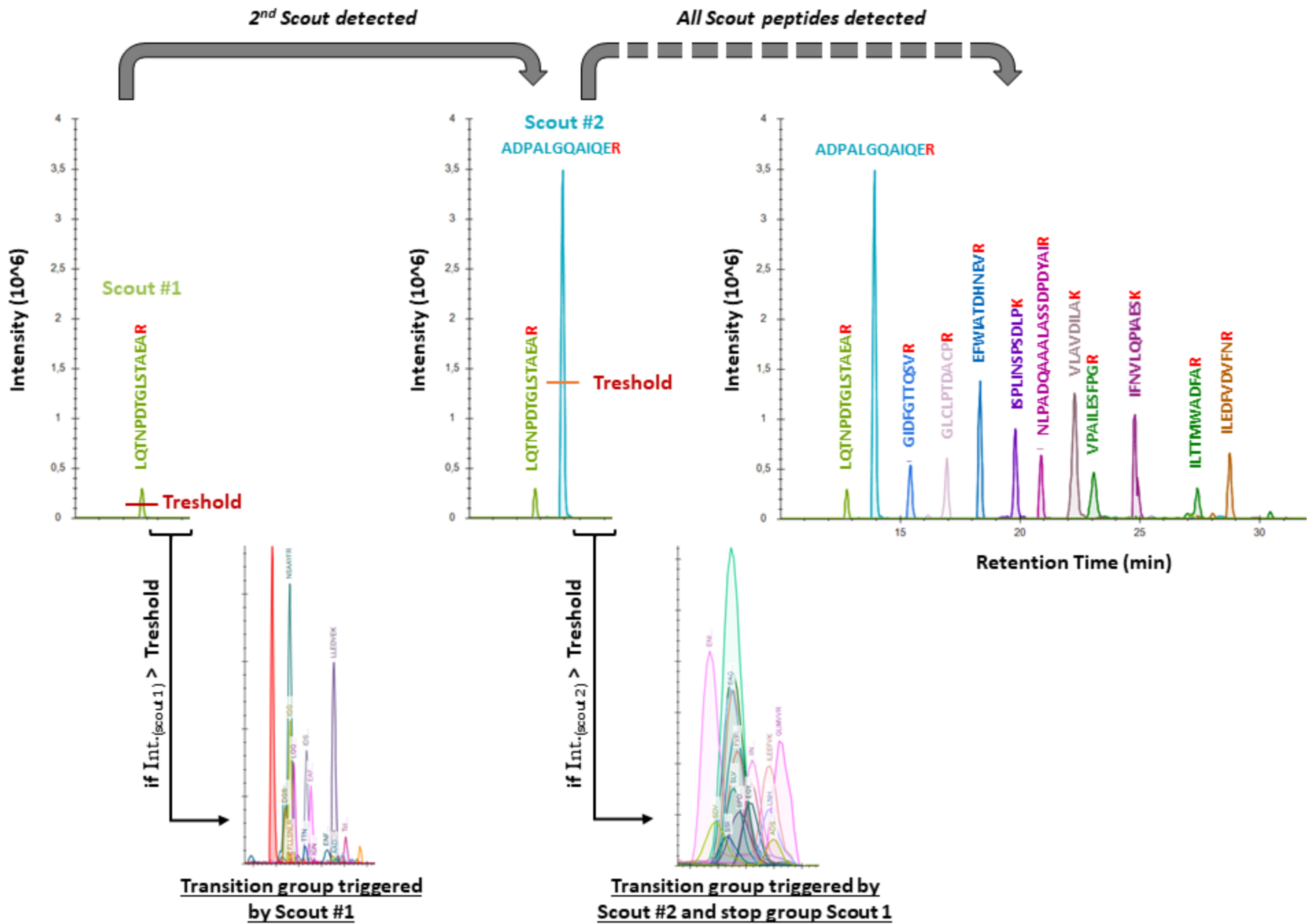
- [38] Y. Huang, K. Hui, M. Jin, S. Yin, W. Wang, Q. Ren, Two endoplasmic reticulum proteins (calnexin and calreticulin) are involved in innate immunity in Chinese mitten crab (*Eriocheir sinensis*), *Sci Rep* 6 (2016) 27578.
- [39] A. Moshtaghi, M.L. Rahi, P.B. Mather, D.A. Hurwood, Understanding the Genomic Basis of Adaptive Response to Variable Osmotic Niches in Freshwater Prawns: A Comparative Intraspecific RNA-Seq Analysis of *Macrobrachium australiense*, *J Hered* 108(5) (2017) 544-552.
- [40] K. Groebe, K. Hayess, M. Klemm-Manns, G. Schwall, W. Wozny, M. Steemans, A.K. Peters, C. Sastri, P. Jaeckel, W. Stegmann, H. Zengerling, R. Schopf, S. Poznanovic, T.C. Stummann, A. Seiler, H. Spielmann, A. Schrattenholz, Protein biomarkers for in vitro testing of embryotoxicity, *J Proteome Res* 9(11) (2010) 5727-38.
- [41] S. Muthukrishnan, H. Merzendorfer, Y. Arakane, Q. Yang, Chitin Organizing and Modifying Enzymes and Proteins Involved In Remodeling of the Insect Cuticle, *Adv Exp Med Biol* 1142 (2019) 83-114.
- [42] R. Coulaud, O. Geffard, B. Xuereb, E. Lacaze, H. Queau, J. Garric, S. Charles, A. Chaumot, In situ feeding assay with *Gammarus fossarum* (Crustacea): Modelling the influence of confounding factors to improve water quality biomonitoring, *Water Res* 45(19) (2011) 6417-29.
- [43] C. Lopes, A. Chaumot, B. Xuereb, R. Coulaud, G. Jubeaux, H. Queau, A. Francois, O. Geffard, 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) (2020) 678-691.





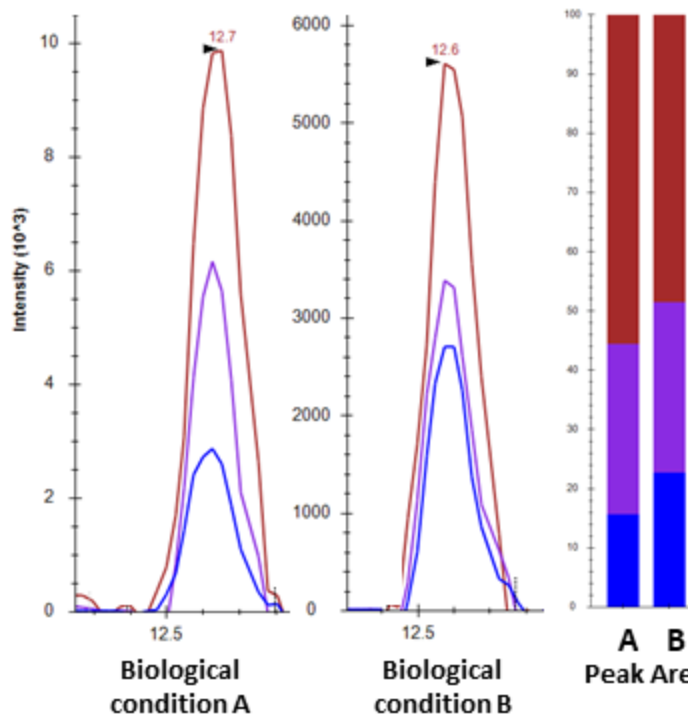
Variation of chromatographic conditions





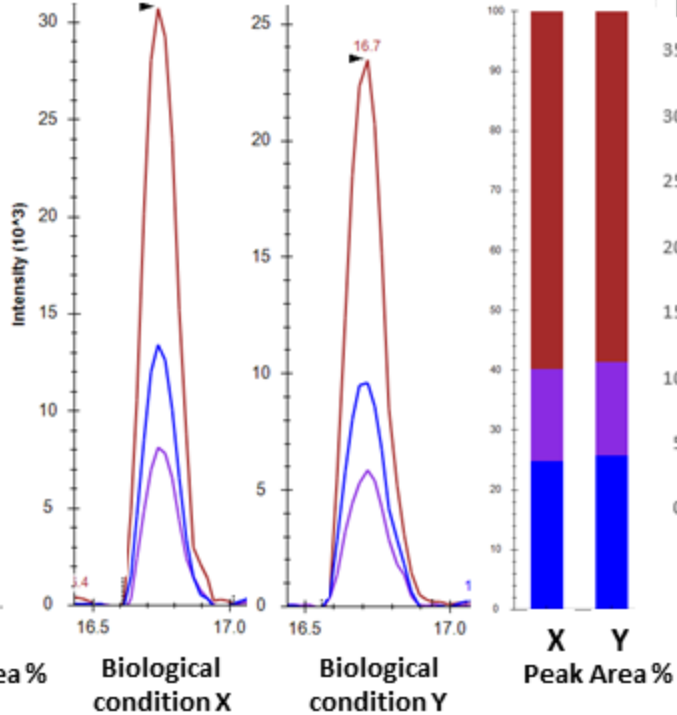
A. CQLFNDPSDR → non-constant transition ratio

— y8 - 963.4530+ — y7 - 850.3690+ — y4 - 474.2307+

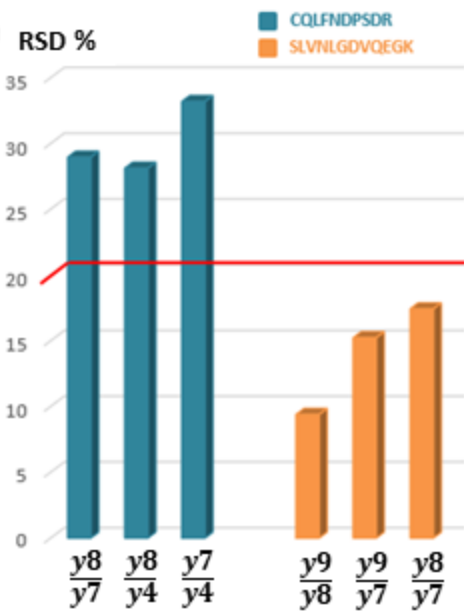


B. SLVNLGDVQEGK → constant transition ratio

— y9 - 959.4793+ — y8 - 845.4363+ — y7 - 732.3523+



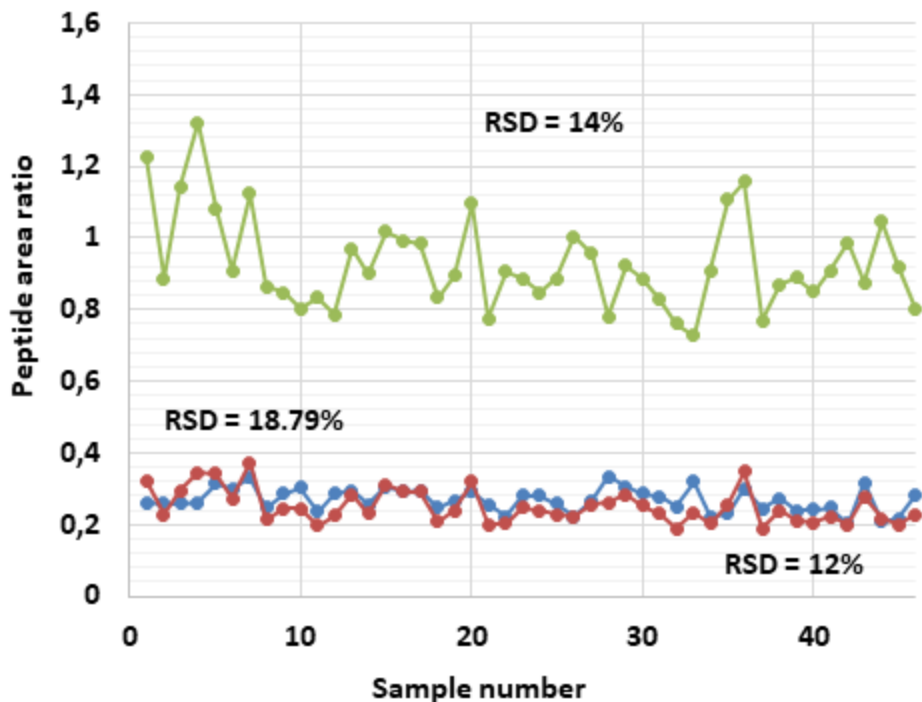
C.



Correlation between the different transitions of the same peptide for 46 *G.fossarum*

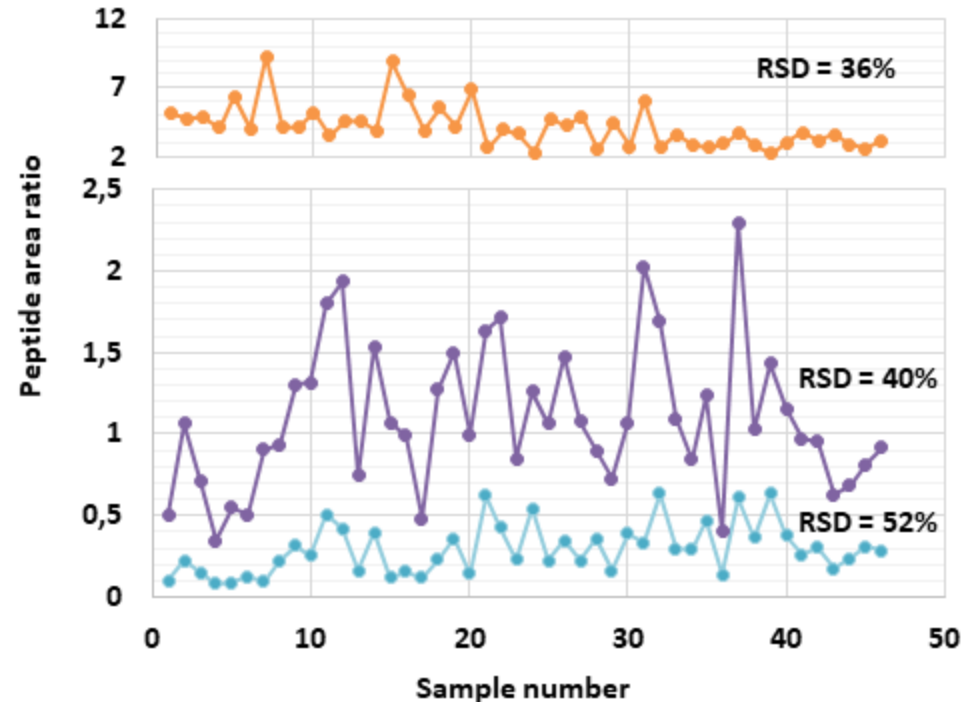
Aspartate aminotransferase

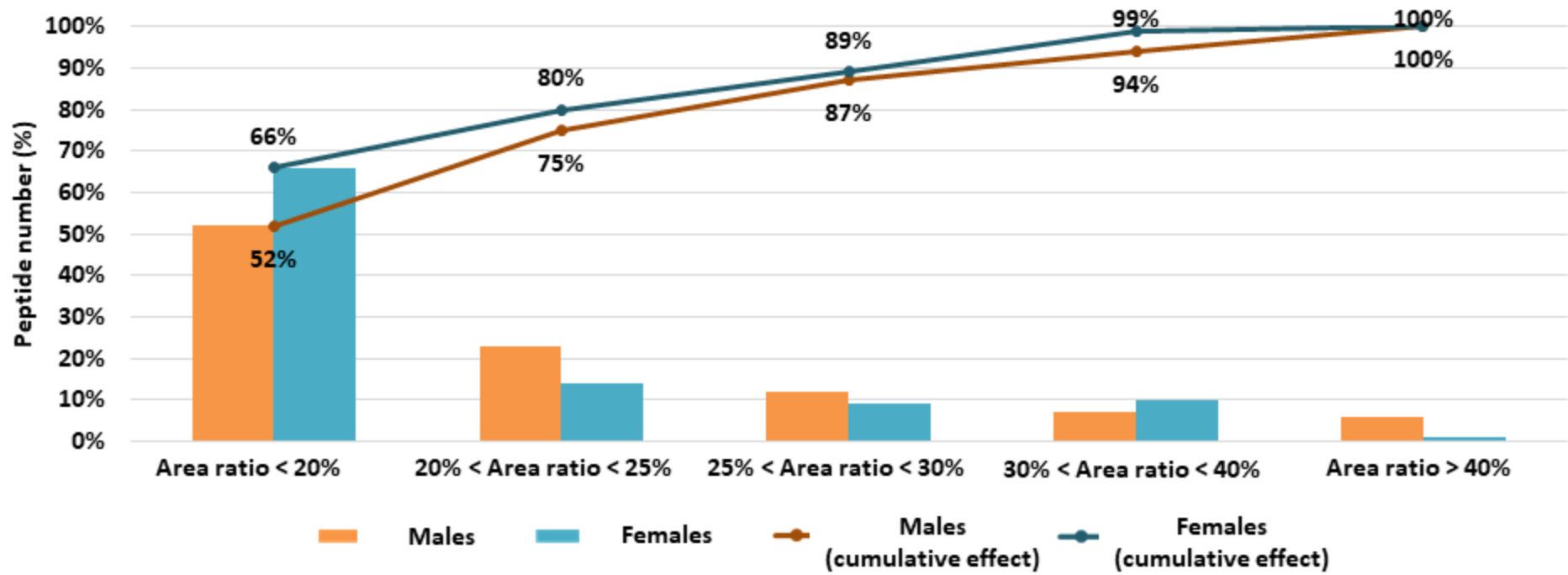
- INICGLTTK.+2y6.light / IVSTVLNDPK.+2y8.light
- INICGLTTK.+2y6.light / SQVTLLVR.+2y6.light
- IVSTVLNDPK.+2y8.light / SQVTLLVR.+2y6.light



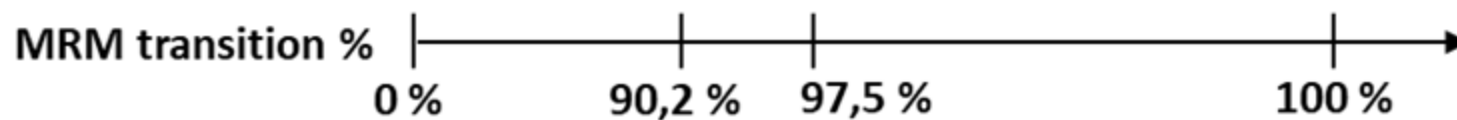
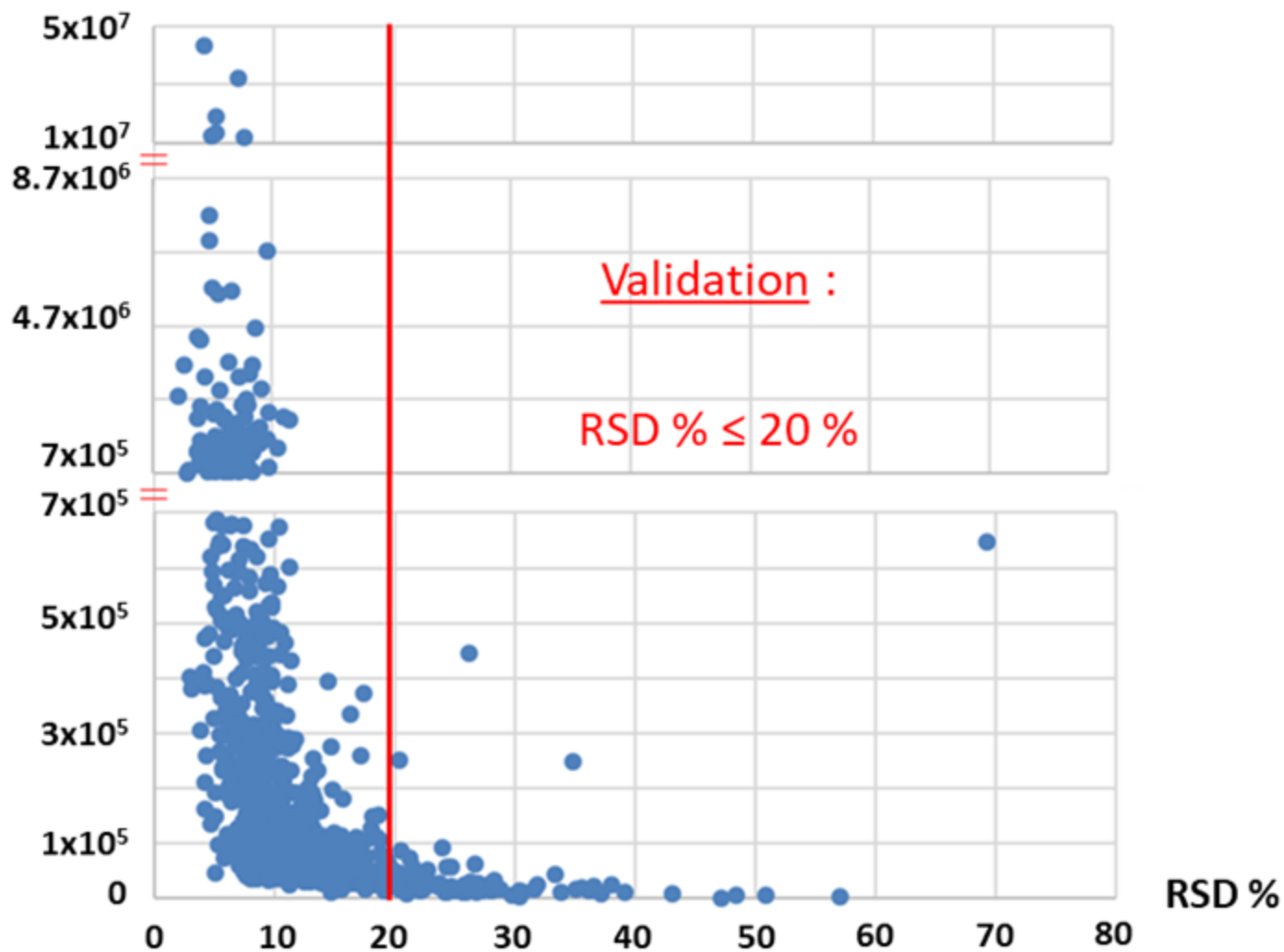
Cytosolic malate dehydrogenase

- DIDAAFLVGAMPR.+2y5.light / DVVATDDPAVAFK.+2y10.light
- DIDAAFLVGAMPR.+2y5.light / WSIVQGLPIDDFSK.+2y7.light
- DVVATDDPAVAFK.+2y10.light / WSIVQGLPIDDFSK.+2y7.light

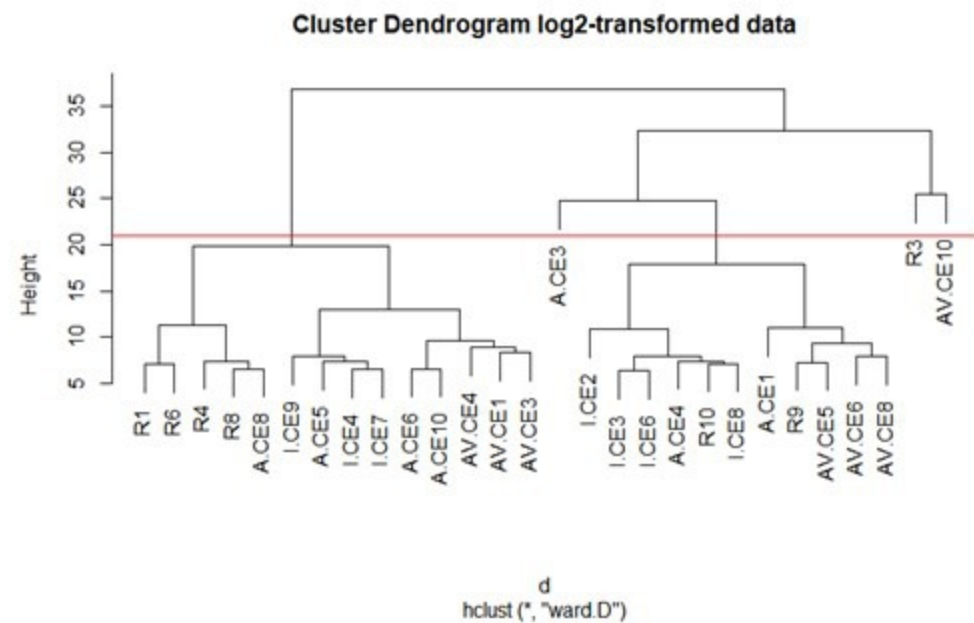




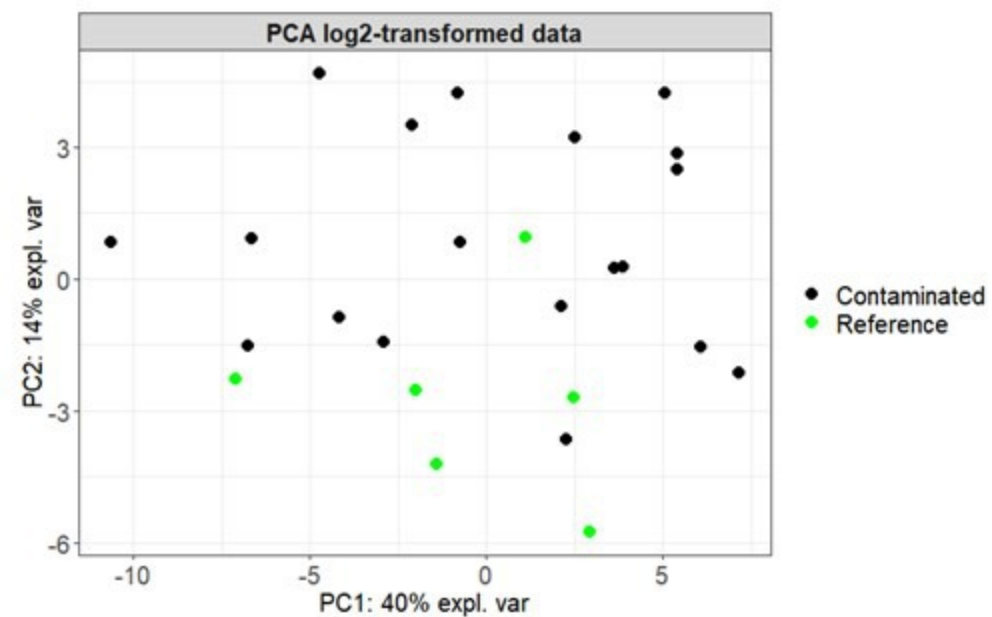
Area (Endogenous peptide)



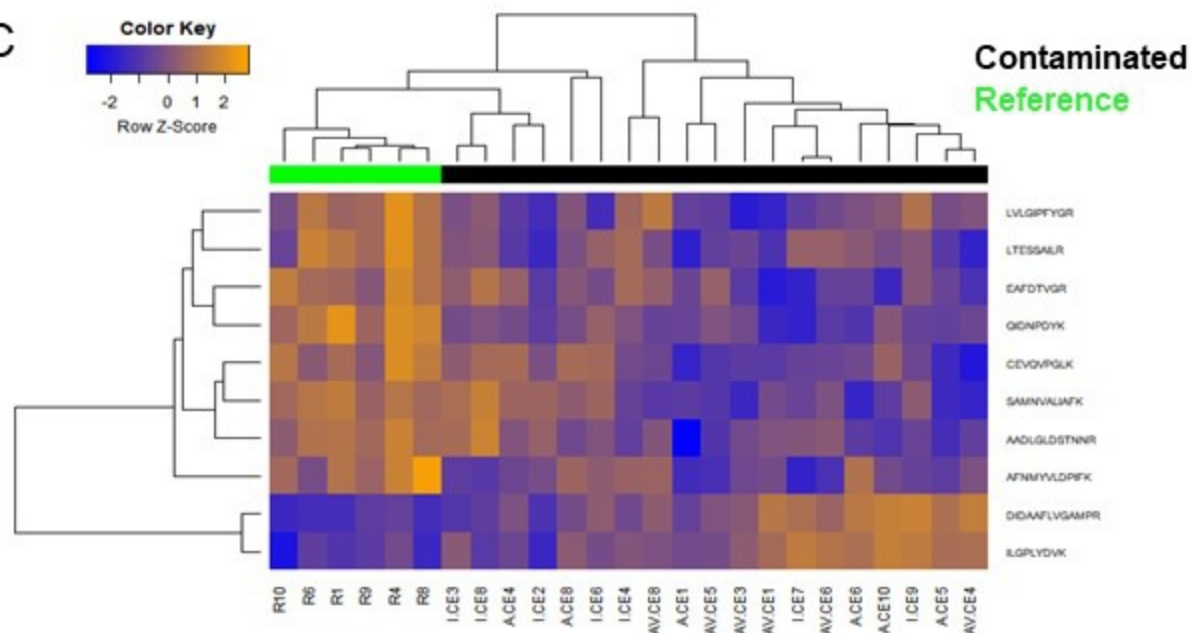
A



B



C



D

