Fish metabolome from sub-urban lakes of the Paris area (France) and potential influence of noxious metabolites produced by cyanobacteria

Benjamin Marie, Alison Gallet

Fish metabolome from sub-urban lakes of the Paris area (France) and potential influence of noxious metabolites produced by cyanobacteria

Benjamin MARIE[1,*], *Alison GALLET*[1]

[1] UMR 7245, CNRS/MNHN, Molécules de Communication et Adaptation des Micro-organismes (MCAM), équipe "Cyanobactéries, Cyanotoxines et Environnement", 12 rue Buffon - CP 39, 75231 Paris Cedex 05, France.

* Corresponding author: benjamin.marie@mnhn.fr

Abstract

The recent democratization of high-throughput molecular phenotyping allows the rapid expansion of promising untargeted multi-dimensional approaches (*e.g.* epigenomics, transcriptomics, proteomics, and/or metabolomics). Indeed, these emerging omics tools, processed for ecologically relevant species, may present innovative perspectives for environmental assessments, that could provide early warning of eco(toxico)logical impairments.

In a previous pilot study (Sotton et al., Chemosphere 2019), we explore by ^{1}H NMR the bio-indicative potential of metabolomics analyses on the liver of 2 sentinel fish species (*Perca fluviatilis* and *Lepomis gibbosus*) collected in 8 water bodies of the peri-urban Paris' area (France). In the present study, we further investigate on the same samples the high potential of high-throughput UHPLC-HRMS/MS analyses. We show that the LC-MS metabolome investigation allows a clear separation of individuals according to the species, but also according to their respective sampling lakes. Interestingly, similar variations of *Perca* and *Lepomis* metabolomes occur locally indicating that site-specific environmental constraints drive the metabolome variations which seem to be influenced by the production of noxious molecules by cyanobacterial blooms in certain lakes. Thus, the development of such reliable environmental metabolomics approaches appears to constitute an innovative bio-indicative tool for the assessment of ecological stress, such as toxigenic cyanobacterial blooms, and aim at being further follow up.

Keywords: Untargeted metabolomics, environmental assessment, cyanobacteria, biomarkers, sentinel species

1. INTRODUCTION

In the past decades, the use of aquatic organisms as bio-indicator has been widely adopted to monitor and to manage the quality of water bodies worldwide (Lorenz 2003). In parallel, the difficulties and the rational limits of a direct chemical monitoring for providing sufficient information to adequately assess the risks from anthropogenic chemicals in the natural environment was progressively documented (Reyjol et al., 2014). More recently, the recent democratisation of high-throughput molecular phenotyping has allowed the rapid expansion of promising untargeted multi-dimensional approaches (*e.g.* epigenomics, transcriptomics, proteomics, metabolomics), that now represent promising perspectives for environmental assessments (Bahamonde et al. 2016; Cordier et al., 2020). It appears now very likely that emerging omics-based analyses, such as metabolomics, developed for ecologically relevant aquatic species, may support the development of new ecological evaluation tools, that could provide early warning of ecotoxicological impairments (Marie 2020).

The metabolome is defined as the set of primary and secondary metabolites synthesized at a given time and thus represents its global metabolic chemical picture. This set of metabolites is assumed to represent a final endpoint of the phenotypic response of an organism that can potentially be modified when the ecological and environmental stress conditions change (Bundy et al., 2008). In this way, metabolomic studies have become a relevant approach to investigate physiological processes involved in the homeostatic responses of the organisms encountering environmental stresses from potentially multiple origins (Capello et al., 2016; Good et al., 2020). However, despite its high potential to support the understanding the molecular mechanisms implicated in the ecotoxicological responses of aquatic organisms (Sardans et al., 2011; Viant, 2008), such environmental metabolomics investigations remains still limited. Indeed, numerous studies have compared the metabolite profiles of individuals collecting from a limited set of environments, often including only one contaminated site and a unique reference location (*e.g.* Capello et al., 2016; Melvin et al., 2019; De Marco et al., 2022). As previously suggested, the selection of a unique "good" reference site may crucially orientate the results of the analyses, and larger efforts should be provided at understanding similarities and differences among various references sites (Simmons et al., 2015; Marjan et al., 2017). Such approaches would gain at investigating larger sets of environments regarding their respective positions within environmental gradients of global ecotoxicological pressures (Jeffries et al., 2010; Davis et al., 2016; Meador et al., 2020).

Among several aquatic organisms, such as molluscs or arthropods, fishes present remarkable features that make them especially relevant for bio-indication purposes. In fact, they occupy almost any aquatic habitat, being capable of experiencing different and variable environmental conditions (Harris 1995). Fishes are also responsive to numerous abiotic (temperatures, water velocities, sediment loads, hypoxia, …) or biotic (famine, predation, parasitism, …) pressures, together with anthropogenic stressors, such as contaminants, that represent additional constraints that fish may experience in disturbed ecosystems (McClanahan 2019). Environmental stressors are known to affect the specific physiological,

developmental and reproductive traits, that allow the different fish species to colonize and occupy various ecosystems (Hamilton et al., 2016).

Nowadays, cyanobacterial blooms are one of the common sources of stress encountered by organisms living in modern freshwater lentic systems around the world (Harke et al., 2016; Burford et al., 2020). Then, various efforts have been attempted to characterize the potential effect of cyanobacteria proliferation on various fish species according to experimentation performed in the lab, in mesocosms, or from field sampling (Amado et al., 2010; Ferrao-Filho et al., 2011; Sotton et al., 2017a, 2017b; Le Manach et al., 2018). To this end, we initially investigated by ^{1}H NMR technics the liver metabolome of fish collected from different lakes of the peri-urban Paris area (France), and encountering low or high cyanobacterial proliferation (Sotton et al., 2019). This pilot study leads to the description of putative metabolite concentrations, corresponding to specific NMR signals, that were significantly correlated with higher cyanobacterial occurrence.

The present study investigates the bio-indicative potential of high-throughput molecular phenotyping supported by LC-HRMS approaches on 2 sentinel fish specimens (*Perca fluviatilis* and *Lepomis gibbosus*) that were previously investigated by ^{1}H NMR (Sotton et al., 2019). We hypothesis that LC-MS-based metabolomics approach may present a higher potential for environmental diagnosis. Indeed, such investigations provide valuable multi-metric characters indicative of biological or ecological conditions, together with a prospective mechanistic under-standing of the cause-effect relationships. With regard to the cyanobacterial occurrence and the associated production of bioactive metabolites in the different lakes, we compare here the metabolome of two different sentinel fish species. Indeed, these organisms could present rather different or comparable sensitivity and responsiveness to local eco(toxico)logical constraints (*i.e.* potentially the production of noxious molecules by cyanobacteria), and may support the general interest for omics-based environmental approach for the characterization of cyanobacterial bloom threat.

2. MATERIALS AND METHODS

2.1. Fish sampling

Field sampling campaigns were performed during late summer 2015 (7-10th September), as this “end-of-the-bloom” period appears more appropriate to observe cyanobacterial effects on fishes. Eight peri-urban pounds around Paris’ area (Île-de-France region, France) were selected for their respective eutrophication levels and the presence or the absence of recurrent cyanobacterial blooms: Cergy-Pontoise (Cer), Champs-sur-Marne (Cha), Maurepas (Mau), Rueil (Rue), Verneuil (Ver), Varennes-sur-Seine (Var), Fontenay-sur-Loing (Fon) and Triel (Tri) pounds (Supplementary figure S1) (Maloufi et al., 2016; Sotton et al., 2019). These sites were sampled by the Hydrosphère company (www.hydrosphere.fr) with electric fishing device (FEG 8000, EFKO, Leutkirch, Germany) for capturing live caught fish, that were then promptly sacrificed for liver collection. The investigation of the fish guild indicates that only

the perch (*Perca fluviatilis*) and pumpkinseed sunfish (*Lepomis gibbosus*) were presents in all or almost all these pounds (supplementary table S1) and were further selected as sentinel species for further molecular phenotyping analyses by metabolomics.

Briefly, live caught fishes (n=5-10 young-of-the-year per pounds and per species) were briefly euthanized by neck dislocation, directly measured (12.0±4.8 cm), weighed (9.3±2.6 g), and then liver of each individual was shortly sampled, flash-frozen in liquid nitrogen and kept at -80°C until analyses, in accordance with European animal ethical concerns and regulations.

In every lake, sub-surface chlorophyll-*a* equivalent concentrations attributed to the four-main phytoplankton groups (Chlorophyta, Diatoms, Cyanobacteria and Cryptophyta) were measured with an *in-situ* fluorometer (Fluoroprobe II, Bbe-Moldenke, Germany) (Supplementary figure S2). Sub-surface water samples filtered on 20-μm mesh size were also collected for phytoplanktonic community analysis and further metabolomics characterisation, and then kept at -80°C until analysis as previously described (Sotton et al., 2019). Physico-chemical parameters of the water column (temperature, conductivity, pH, dissolved O_2 and nitrate concentrations) were monitored with a YSI DSS probe, as previously described (Sotton et al., 2019).

2.2. Liver metabolite extraction and metabolomics analyses

The liver extraction was performed on 132 individuals (comprising 78 *Perca* and 54 *Lepomis*) with methanol/chloroform/water (ratio 2/2/1.8 – 22 $mL.g^{-1}$ at 4°C) and the polar fraction was analyzed on a 600-MHz NMR spectrometer equipped with a 5-mm cryoprobe (Advance III HD Bruker, Germany) with a noesygppr1d pulse sequence as previously described (Sotton et al., 2019). ^{1}H-NMR spectra were treated with Batman R-package for deconvolution, peak assignment and quantification of 222 putative metabolites (Hoa et al., 2014).

The liver extracted polar phase was additionally injected (2 µL) on C_{18} column (Polar Advances II 2.5 pore - Thermo), then eluted at a 300 $\mu L.min^{-1}$ flow rate with a linear gradient of acetonitrile in 0.1% formic acid (5 to 90 % in 21 min) with an ultra-high-performance liquid chromatography (UHPLC) system (ELUTE, Bruker). Consecutively, the individual metabolite contents were analysed using an electrospray ionization hybrid quadrupole time-of-flight (ESI-Qq-TOF) high-resolution mass spectrometer (Compact, Bruker) at 2 Hz speed on positive MS mode on the 50–1500 *m/z* range. The feature peak list was generated from recalibrated MS spectra (< 0.5 ppm for each sample, as an internal calibrant of Na formate was injected at the beginning of each sample analysis) within a 1-15 min window of the LC gradient, with a filtering of 5,000 count of minimal intensity, a minimal occurrence in at least 50% of all samples, and combining all charge states and related isotopic forms using MetaboScape 4.0 software (Bruker).

Additionally, five pools of six different individuals randomly selected for *Perca* and *Lepomis* (quality check samples) and the phytoplanktonic biomass extracts from the 8 lakes were similarly eluted then analysed on positive autoMS/MS mode at 2-4 Hz on the 50-1500 *m/z* range for further metabolite annotation. Molecular networks were performed with Global

Natural Product Social network (http://gnps.ucsd.edu) and/or MetGem (http://metgem.github.io) softwares, as previously described (Kim Tiam et al., 2019; Le Manach et al., 2019) for cyanobacterial and fish metabolite annotation was assayed using in-house (Le Manach et al., 2019) or CyanoMetDB (Jones et al., 2021), and GNPS, Mona, human metabolome database (HMDB) and Massbank MS/MS spectral libraries. Water samples of each lake, concentrated on a 20-μm mesh size, were also extracted with 75% methanol (2 min sonication, 5 min centrifugation at 15,000 g - 4°C) and then similarly analysed in triplicates on LC-HRMS on autoMS/MS positive mode, as earlier described, and metabolite list annotated with same pipelines.

2.4. Data matrix treatment

The resulting the intensity data tables of the 222 metabolites (^{1}H NMR) and 1252 analytes (LC-MS) were further treated for normalization with Pareto's scaling, inspected data representation by PCA and heatmap with hierarchical clustering (Euclidean distance) and then analysed to investigate the influences of "Species" and "Lakes" parameters on the datasets by PERMANOVA, PLS-DA or multi-block PLS-DA using MixOmics and Mixomics DIABLO R Package (Rohart et al., 2017; Singh et al., 2019), MetaboAnalyst 5 (Chong et al., 2019) and MicrobiomeAnalyst tools (Chong et al., 2020).

3. RESULTS AND DISCUSSION

3.1. Distinction of *Perca* and *Lepomis* liver LC-MS metabolome

Figure 1 illustrates the global metabolite profiles analyzed by LC-MS for each fish (comprising *Perca* and *Lepomis*) on a heatmap with hierarchical classifications and individual plot representation of principal component analyses (PCA). A very strong and significant discrimination between the "species" was achieved (PERMANOVA F-value = 94.92; R^2 = 0.43; P-value <0.001), when a slighter but still significant discrimination can be also retrieved for "lakes" (PERMANOVA F-value = 4.96; R^2 = 0.22; P-value <0.001).

In parallel, a molecular network performed with fragmentation data was obtained by LC-MS/MS analyses of *Perca* and *Lepomis* metabolomes. This approach supports the global annotation process of the fish liver metabolites considering both structural identities and similarities (Nothias et al., 2020). The GNPS and the *t*-SNE representation of the molecular network highlights that most, if not all, of the known metabolites (presenting structural identity or analogy hits) are largely shared between the two species (*e.g.* nucleic acids, carnitines, glutathiones, lipids, saccharides, …). Moreover, numerous species-specific cluster metabolites remain uncharacterized, with relative molecules and clusters presenting no match within the public chemical databases (supplementary figure S2) (da Silva et al., 2015). Taken together, these analyses illustrate for these two species the various molecular specificities of their respective metabolomes. It also shows the substantial portion of the specific liver molecular metabolism that still remains uncharacterized and that could be

related to species-specific nutritional, physiological or toxicological functions (Brusle et al., 2017).

3.2. Comparison of *Perca* and *Lepomis* LC-MS metabolome between lakes

In addition, the LC-MS datasets of the two fish species considered separately show even clearer discriminations of the metabolomes of fish originating from the different lakes. The un-supervised PCA and the heatmap with hierarchical classification (Figure 2) present distinct environmental signatures that support reliable discrimination of the lake of sampling for both *Perca* and *Leptomis* (Fig. 2A-B and 2C-D, respectively). These observations strongly suggest that the locality seems to globally influence the local LC-MS metabolome composition for both species. However, the liver metabolomes of individuals from Fontenay-sur-Loing lake (Fon) exhibit quite species-specific relative positions on PCA (Figure 2A and 2C). Such differences of the relative position of metabolomes from Fon between the two species suggests that: *i/* for most localities, the specific "lake" signature of the metabolome appears in good agreement between the two species, *ii/* for some specific environments (*e.g.* Fon), the metabolome signatures are variable from one species to the other one. Thus, the species-specific signature of the *Perca* and *Lepomis* metabolomes in Fon lake may rather traduce the specific responsiveness and/or sensitivity to a local factor occurring at this particular location.

Interestingly, the two hierarchical classifications performed on LC-MS metabolome datasets of *Perca* and *Lepomis,* respectively, overall show a global similarity of the metabolome of fish originating from the same lake, together with an overall distinction of 3 principal lake clusters. Indeed, these analyses show very similar lake relationships for both species, grouping together all fish from the lakes of 'Cer and Mau', from 'Cha, Fon and Rue', and from Tri, Var and Ver' for *Perca,* or 'Tri and Ver' for *Lepomis*, as this latter was not retrieved in Var (Figure 2B and 2D). This observation suggests that these localities could somehow present comparable, if not analogous, environmental conditions. This hypothesis was firstly considered regarding NMR metabolomics and potential explicative environmental factors (physico-chemical and biological factors – supplementary table S2) and had suggested a potential influence of cyanobacterial proliferation (BBEcya and pH) (Sotton et al., 2019). Indeed, as initially shown, the Fon, Tri, Var and Ver lakes encounter remarkable high amounts of cyanobacteria belonging to the *Planktothix*, *Pseudo-Anabaena*, *Aphanizomenon*, *Anabaena* or *Microcystis* genera. Differently, the other lakes (Cer, Mau, Cha and Rue) exhibit phytoplanktonic biomasses principally dominated by green algae, diatoms or cryptophytes (Fig. 3A-B). A molecular network performed with LC-MS/MS data obtained from the phytoplanktonic biomass of these 8 lakes shows that noticeable amounts of cyanopeptides are observed in the water column of the cyanobacterial-rich lakes (Fon, Tri, Var and Ver), with the specific presence of most noxious cyanopeptides (microcystins, cyanopeptolins and anabaenopeptins) being detected in Tri, Var and Ver lakes. In a different manner, Fon water exhibits rather less toxic cyanopeptides (Janssen 2019), such as microginins and aeruginosins

(Fig. 3C). Taken together, these observations suggest that the proliferation and the subsequent production of such potentially noxious cyanobacterial metabolites could be also considered as an explicative factor of the local metabolome singularity for the liver of *Perca* and *Lepomis* from Tri, Var and Ver lakes, as previously postulated (Sotton et al., 2019).

3.3. Discriminating metabolites between the different lake groups

Thus, we further investigate the fish metabolome distinctions considering together fishes from lakes presenting noxious cyanobacterial metabolites ("Tri/Var/Ver" group for *Perca* and "Tri/Ver" group for *Lepomis*), or not ("Cer/Mau/Cha/Fon/Rue" group). For this purpose, we used PLS-DA to identify metabolites that discriminate the individuals belonging to these two lake groups (Figure 4). Indeed, while unsupervised PCA were first used to evaluate the global dispersion between species and sampling groups, a supervised model such as partial least square differential analysis (PLS-DA) allows to maximize the separation between the different sample classes (*e.g.* "lake groups") and to further provide information on discriminating features (variable importance on projection - VIP). For both fish species, the metabolomes present a consistent discriminant analysis scoring (*Perca*: accuracy = 1, performance R^2 = 0.78, predictability Q^2 = 0.72; *Lepomis*: accuracy = 0.93, R^2 = 0.59, Q^2 = 0.29) and large sets of discriminating variables (*Perca* $VIP_{>1.5}$=109 and *Lepomis* $VIP_{>1.5}$=93) (Figure 4A-B). Among these metabolites, 33 components have been successfully annotated considering their respective molecular mass, the isotopic pattern, together with the fragmentation pattern of all components with GNPS and MetGem algorithms and a manual curation. Overall, the metabolite semi-quantification of the annotated best VIP reveals largely similar variations between *Perca* and *Lepomis* for these two groups of lakes (Figure 4C; Supp. fig. S3). Indeed, except for the reduced-glutathione, which present higher concentrations in the livers of *Perca* from "Tri/Var/Ver" lakes, when livers of *Lepomis* exhibits lower concentrations than in the those of fish of the "Cer/Mau/Cha/Fon/Rue" group, all other metabolites present similar quantification patterns between the two species. This discrepancy suggests that glutathione-dependent redox mechanisms could be more extensive in *Perca* than in *Lepomis* livers.

Globally, the main differences in the different metabolite concentration between these two lake groups show that fishes collected from the "Cer/Mau/Cha/Fon/Rue" present higher energetic, anti-oxidant/detoxification and lipids reserves (*e.g.* various nucleic acids, glutathiones, flavonoids and lipo-phosphocholines) than in "Tri/Var/Ver" lakes. In addition, higher amounts of amino acids, di-peptides and acylcarnitines in the liver of fish originating from the "Tri/Var/Ver" lakes could be related to higher protein catabolism and beta-oxidation, that could denote of higher energetic requirements for those specific organisms. Interestingly, the "Tri/Var/Ver" lakes present much more cyanobacteria and related noxious metabolites, that may represent a patent stress source for fishes colonizing those environments. Taken together, these observations are in agreement to the hypothesis of an occurrence of higher stress conditions locally occurring in the Tri, Var and Ver lakes, that could be associated with the local production of potentially noxious metabolites by

cyanobacterial proliferation (Sotton et al., 2017a, 2017b; Le Manach et al., 2018).

In general, various genetic or phenotypic factors (such as development stages, contamination levels, predator/parasite pressures, food availability, …) could influence and explain the local discrepancy of the metabolome of fishes collected from different environments (reviewed in Marie 2020). However, in the present case, the fact that *Perca* and *Lepomis* metabolomes analyzed in parallel present similar metabolite variations clearly indicates that local environmental constraints may drive such phenotypic co-variations. Then, these present results claim for the high informativeness of LC-MS-based metabolomic imprints for molecular phenotyping on sentinel species and the exploration of environmental stresses (Lohr et al., 2019).

3.4. Comparison of ^{1}H NMR and LC-MS metabolomics for fish livers

Considering the previous ^{1}H NMR metabolomics datasets retrieved from the same samples on these two fish species (Sotton et al., 2019), the liver metabolome analyzed by NMR shows more limited specificity according to “species” or “lakes” factor than those analyzed by LC-MS. Although the 222 potential metabolites quantified by NMR on 132 fish (78 *Perca* and 54 *Lepomis*) present no obvious structuration of the individual dataset according to “species” or “lakes” variables, when observed on PCA or hierarchical classification, slight but significant relations are however supported by PERMANOVA (F-values = 4.15 and 3.63; R^2 = 0.03 and 0.17; P-values < 0.01 and 0.01, for “species” and “lakes” respectively, Supp. fig. S4). Indeed, according to the relatively lower performances of the hierarchical classifications and PERMANOVA and the PLS-DA (Supp. fig. S4 and S5), the NMR liver metabolomes exhibit a limited discrimination on “lakes” for both species, when compared to LC-MS metabolomics performances.

Overall, regarding both hierarchical classification, individual plots of PCA or PERMANOVA performances, the dataset obtained from the LC-MS metabolome analysis presents a clearer discriminant pattern between the two species or the 8 lakes, than the NMR dataset generated on the same samples (with the same extraction procedure) (Fig. 1 and 2; Supp. fig. S4 and S5). This observation is further confirmed by direct comparison of the NMR and LC-MS metabolome dataset using MixOmics Diablo tool for multi-block PLS-DA (Supp. fig. S6 and S7). Indeed, although the two datasets appear mostly congruent, the individual plot distribution carries much more variance on the PLS-DA projection for the LC-MS analysis than for the NMR, considering both fish species.

On one side, the ^{1}H NMR metabolomics allows the reliable quantification of main liver metabolites (belonging principally to amino acids, sugars, TCA metabolites, …), by detecting characteristic chemical groups, with regards to their respective position within the molecules (Martin et al., 2015). These primary metabolites are key contributors to the cellular metabolism and are supposed to be responsive to variable physiological conditions of fish encountering peculiar biological outcomes and/or environmental constraints (Roques

et al., 2020). However, this approach remains focused on few principal metabolites involved in main cellular pathways which measured cellular contents (expressed in *ppm*) can be globally involved in multiple regulation processes, and not being specific of subtle molecular regulation processes. On the other hand, ^{1}H NMR presents only low sensitivity and an overall limited discrimination capability for molecules presenting similar chemical structures, such as lipids, that exhibits an aggregated and unspecific signal on NMR spectra (Emwas et al., 2019). For this reason, it is supposed to present lower discriminating power than more selective and less qualitative, but more sensitive, high-resolution mass spectrometry-based metabolomics approaches.

On the other side, although LC-MS-based metabolomic remains a largely selective analytical approach (according to both LC separation and ionisation respective properties of each molecule), it presents un-precedent analytical capabilities in terms of sensitivity, dynamic range and number of characterized components and offer now unprecedented perspectives for the investigation for molecular phenotyping applied to environmental sciences (Beyoglu and Idle, 2020). We assume, that a single LC-MS fingerprint analyse does not embed the whole metabolite picture of the biological compartment (the fish liver, in the present case), especially because of its selectivity performances (Gika et al., 2019). Indeed, this noticeable selectivity can have multifactorial origin along the analytical procedures and be related to both the extraction solvent, the selected chromatographic gradient and column, together with the ionisation principle and mode. However, it provided specific and precise measurements of a large number of components that can endorse high-throughput and in-depth phenotyping. For comparison, ^{1}H NMR metabolomics allow the global quantification of the CH_2-CH_2 or CH_2-CH=CH fatty acid bounds (Marchand et al., 2018), when LC-MS metabolomics can discriminate and characterize hundreds of different lipids belonging to various sub-classes (Dreier et al., 2020), providing undoubtedly more informative and discriminant features. For these reasons, recent efforts are still been performed in order to compared the inputs of both NMR- and LC-MS-based metabolomics in order to explore the respective advantages and caveats of both approaches for environmental investigations (Ameline et al., 2019; Hayden et al., 2019; Labine et al., 2020). In addition, multi-omics integration tools, such as MixOmics DIABLO (Singh et al., 2019), also provides innovative solutions for the comparison of heterogeneous dataset joint matrices obtained from a single set of samples or individuals (supplementary figures S6-S7).

However, the robustness and the repeatability of the operational analytical pipeline constitute key requirements of first order for environmental metabolomics, as recently experienced by clinic development of environmental toxicology (Silva et al., 2020). Nowadays, it also appears even more crucial to carefully take into account the socio-economic and the environmental costs when considering the development of different pipelines for operational and sustainable metabolomic applications (Jones et al., 2013).

3.5. Use of LC-MS metabolomics fingerprint for environmental assessment?

Although the value of LC-MS metabolomics for investigating the impacts of environmental stressors or contaminants, and their respective mode of action has been well-explored in medical sciences (Beyoğlu et al., 2020; Zhang et al., 2020) or in ecotoxicology laboratory-based studies on aquatic models (Le Manach et al., 2018; Huang et al., 2017; Ekman et al., 2015; Gil-Solsona et al., 2017), such methods have been used only faintly in field research so far (Bundy et al., 2009; Lankadurai et al., 2013). Apart from a limited number of evidence of the utility of NMR-based metabolomics in environmental fish studies (Capello et al., 2016; Sotton et al., 2019; Wei et al., 2018), few other examples indicate that field-based LC-MS metabolomics constitutes an emerging and powerful, but still underused, approach for increasing our understanding of *in situ* biological, physiological, ecological or ecotoxicological processes (Meador et al., 2020; Reverter et al., 2017; Goode et al., 2020).

So far, aquatic ecotoxicology has been widely using fishes as reference models to investigate both physical or chemical contaminants (Li et al., 2021; Mallik et al., 2021), and their deleterious effects on diverse biological mechanisms including their behaviour (Almroth et al., 2021) or specific enzymatic activities (Villeneuve et al., 2021), on short or longer terms (Thoré et al., 2021). Among the large list of the characters of interest on which ecotoxicological studies aim at being dedicated, the metabolome appears as one of the main promising functional traits (Martynuik et al., 2016). Indeed, metabolomics seem to present many key advantages for environmental studies, such as its immediate connexion with the global physiology of the organisms (Rosato et al., 2018), its pertinence for all model or non-model organisms (Hillyer et al., 2022; Reverter et al., 2017), the multi-variable character of the dataset and its statistical strength (Chu et al., 2019; Mendez et al., 2019), and its appliance for non-lethal and minimal-stress approaches (Colette et al., 2016). However, applying metabolomics to ecotoxicological questions is still currently limited by a lack of conceptual, technical and data frameworks, which require to be tailored to reach its full potential for applied environmental purposes (Walker et al., 2022).

Our analysis constitutes one of the first attempts to push forward the potential of high-throughput molecular phenotyping, and especially through LC-MS-based metabolomics, for environmental assessment. Indeed, this organism molecular phenotyping supported by multi-variable chemometric investigation offers rich biological information that serves at describing specific phenotypic plasticity. Moreover, this molecular variability/responsiveness can further be confronted to local environmental factors (such as noxious cyanobacterial metabolites in the present case) in order to demonstrate correlation or causality relationships.

As described by Pompfret and co-workers (2019), environmental metabolomics exhibit very promising perspectives for operational bio-monitoring applications, according to its reliability, its reproducibility, and its high predictive potential. However, these authors also point out that the responsiveness and robustness of the bio-indicative object, which is characterized through the analytical prism of the metabolomics, remains crucial and have to

be carefully evaluated and tested with an appropriate experimental design (Pompfret et al., 2019). However, for ethical concerns, fish bio-monitoring would also gain at been less invasive and deleterious for the organisms. To this end, non-lethal mucus sampling has been investigated by LC-MS and have demonstrated the remarkable informativeness of this approach for environmental studies (Reverter et al. 2017), but these efforts still remain explorative and require additional investigations.

4. Conclusions

The main challenge facing us remains to be able to develop more integrative approaches to connect chemical, biological and ecological assessment, in the context of anthropized natural environments confronting multi-stressor pressures. A major caveat of the use of fish environmental bio-indication by metabolomics remains maybe the lack of standardized methodology and, overall, of dedicated databases (Viant et al., 2019) fulfilled by studies considering together different species, populations, development stages, seasons and environments. These data could be used to provide baseline reference values that would further support machine learning or artificial neural network tools for the training of decision-making models (Mendez et al., 2019).

Author contributions

B.M. designed experiment, and performed sample analysis. B.M. and A.G. analysed data. B.M. wrote the manuscript. All authors have given approval to the final version of the manuscript.

Notes

The authors declare not conflict of interest

ACKNOWLEDGMENTS

We would like to thanks B. Sotton for his collaboration on NMR data acquisition and treatment. We are also greatful to A. Paris for its support and prolific discussions. This work was supported by grants from CNRS (Défi ENVIROMICS “Toxcyfish” project). The NMR and the MS spectra were respectively acquired at the Plateau technique de Résonance Magnétique Nucleaire and the Plateau technique de spectrométrie de masse bio-organique, Muséum National d’Histoire Naturelle, Paris, France. This work benefitted from the French GDR "Aquatic Ecotoxicology" framework which aims at fostering stimulating scientific discussions and collaborations for more integrative approaches.

REFERENCES

444 Almroth, B. C., Cartine, J., Jönander, C., Karlsson, M., Langlois, J., Lindström, M., ... & Sturve,
445 J. (2021). Assessing the effects of textile leachates in fish using multiple testing
446 methods: From gene expression to behavior. *Ecotoxicology and Environmental*
447 *Safety*, *207*, 111523.
448 Amado, L., Monserrat, J., 2010. Oxidative stress generation by microcystins in aquatic
449 animals: Why and how. *Environment International* 36, 226–235.
450 Ameline, A., Garnier, D., Gheddar, L., Richeval, C., Gaulier, J. M., Raul, J. S., & Kintz, P. (2019).
451 Identification and analytical characterization of seven NPS, by combination of 1H
452 NMR spectroscopy, GC–MS and UPLC–MS/MS®, to resolve a complex toxicological
453 fatal case. *Forensic science international*, *298*, 140-148.
454 Bahamonde, P. A.; Feswick, A.; Isaacs, M. A.; Munkittrick, K. R.; Martyniuk, C. J. (2016)
455 Defining the role of omics in assessing ecosystem health: Perspectives from the
456 Canadian environmental monitoring program. *Environmental toxicology and*
457 *chemistry 35*(1), 20-35.
458 Beyoğlu, D.; Idle, J. R. (2020). Metabolomic and lipidomic biomarkers for premalignant liver
459 disease diagnosis and therapy. *Metabolites 10*(2), 50.
460 Brusle, J., & i Anadon, G. G. (2017). The structure and function of fish liver. In *Fish*
461 *morphology* (pp. 77-93). Routledge.
462 Bundy, J. G.; Davey, M. P.; Viant, M. R. (2009) Environmental metabolomics: a critical review
463 and future perspectives. *Metabolomics 5*(1), 3.
464 Burford, M. A., Carey, C. C., Hamilton, D. P., Huisman, J., Paerl, H. W., Wood, S. A., & Wulff,
465 A. (2020) Perspective: Advancing the research agenda for improving understanding of
466 cyanobacteria in a future of global change. *Harmful Algae*, *91*, 101601.
467 Cappello, T.; Brandão, F.; Guilherme, S.; Santos, M. A.; Maisano, M.; Mauceri, A.; … Pereira,
468 P. (2016) Insights into the mechanisms underlying mercury-induced oxidative stress
469 in gills of wild fish (*Liza aurata*) combining ^{1}H NMR metabolomics and conventional
470 biochemical assays. *Science of the Total Environment 548*, 13-24.
471 Chong, J.; Wishart, D. S.; Xia, J. (2019) Using metaboanalyst 4.0 for comprehensive and
472 integrative metabolomics data analysis. *Current protocols in bioinformatics 68*(1):e86.
473 Chong, J.; Liu, P.; Zhou, G.; Xia, J. (2020) Using MicrobiomeAnalyst for comprehensive
474 statistical, functional, and meta-analysis of microbiome data. *Nature Protocols* 15,
475 799-821.
476 Chu, S. H., Huang, M., Kelly, R. S., Benedetti, E., Siddiqui, J. K., Zeleznik, O. A., ... &
477 Consortium of Metabolomics Studies Statistics Working Group. (2019). Integration of
478 metabolomic and other omics data in population-based study designs: An
479 epidemiological perspective. *Metabolites*, *9*(6), 117.
480 Collette, T. W., Skelton, D. M., Davis, J. M., Cavallin, J. E., Jensen, K. M., Kahl, M. D., ... &
481 Ekman, D. R. (2016). Metabolite profiles of repeatedly sampled urine from male
482 fathead minnows (Pimephales promelas) contain unique lipid signatures following
483 exposure to anti-androgens. *Comparative Biochemistry and Physiology Part D:*
484 *Genomics and Proteomics*, *19*, 190-198.

485 Cordier T., Alonso-Sáez L., Apothéloz-Perret-Gentil L., Aylagas E.; Bohan D. A., Bouchez A.;,...
486 Keeley, N. (2021) Ecosystems monitoring powered by environmental genomics: a
487 review of current strategies with an implementation roadmap. *Molecular Ecology* 30,
488 2937-58.
489 da Silva, R. R.; Dorrestein, P. C.; Quinn, R. A. (2015) Illuminating the dark matter in
490 metabolomics. *Proceedings of the National Academy of Sciences 112*(41), 12549-
491 12550.
492 Davis, J. M., Ekman, D. R., Teng, Q., Ankley, G. T., Berninger, J. P., Cavallin, J. E., ... & Collette,
493 T. W. (2016). Linking field-based metabolomics and chemical analyses to prioritize
494 contaminants of emerging concern in the Great Lakes basin. *Environmental*
495 *toxicology and chemistry*, *35*(10), 2493-2502.
496 De Marco, G., Brandão, F., Pereira, P., Pacheco, M., & Cappello, T. (2022). Organ-Specific
497 Metabolome Deciphering Cell Pathways to Cope with Mercury in Wild Fish (Golden
498 Grey Mullet *Chelon auratus*). *Animals*, *12*(1), 79.
499 Dreier, D. A.; Bowden, J. A.; Aristizabal-Henao, J. J.; Denslow, N. D.; Martyniuk, C. J. (2020)
500 Ecotoxico-lipidomics: An emerging concept to understand chemical-metabolic
501 relationships in comparative fish models. *Comparative Biochemistry and Physiology*
502 *Part D: Genomics and Proteomics 36,* 100742.
503 Ekman, D. R.; Skelton, D. M.; Davis, J. M.; Villeneuve, D. L.; Cavallin, J. E.; Schroeder, A.; ...
504 Collette, T. W. (2015) Metabolite profiling of fish skin mucus: a novel approach for
505 minimally-invasive environmental exposure monitoring and surveillance.
506 *Environmental Science & Technology 49*(5), 3091-3100.
507 Emwas, A. H.; Roy, R.; McKay, R. T.; Tenori, L.; Saccenti, E.; Gowda, G. A.; ... Wishart, D. S.
508 (2019) NMR spectroscopy for metabolomics research. *Metabolites 9*(7), 123.
509 Ferrão-Filho, A. da S., Kozlowsky-Suzuki, B., 2011. Cyanotoxins: bioaccumulation and effects
510 on aquatic animals. *Marine drugs* 9(12), 2729–2772.
511 Gika, H.; Virgiliou, C.; Theodoridis, G.; Plumb, R. S.; Wilson, I. D. (2019) Untargeted LC/MS-
512 based metabolic phenotyping (metabonomics/metabolomics): The state of the art.
513 *Journal of Chromatography B 1117*, 136-147.
514 Gil-Solsona, R., Nácher-Mestre, J., Lacalle-Bergeron, L., Sancho, J. V., Calduch-Giner, J. A.,
515 Hernández, F., & Pérez-Sánchez, J. (2017) Untargeted metabolomics approach for
516 unraveling robust biomarkers of nutritional status in fasted gilthead sea bream
517 (*Sparus aurata*). *PeerJ 5*, e2920.
518 Goode, K. L.; Dunphy, B. J.; Parsons, D. M. (2020) Environmental metabolomics as an
519 ecological indicator: Metabolite profiles in juvenile fish discriminate sites with
520 different nursery habitat qualities. *Ecological Indicators 115*, 106361.
521 Hamilton, P.; Cowx, I.; et al.; Tyler, C. R. (2016) Population-level consequences for wild fish
522 exposed to sublethal concentrations of chemicals–a critical review. *Fish and Fisheries*
523 *17*(3), 545-566.
524 Hao, J.; Liebeke, M.; Astle, W.; De Iorio, M.; Bundy, J. G.; Ebbels, T. M. D. (2014) Bayesian
525 deconvolution and quantification of metabolites in complex ^{1}D NMR spectra using

526 BATMAN. *Nature protocols* 9(6), 1416–1427.
527 Harke, M. J., Steffen, M. M., Gobler, C. J., Otten, T. G., Wilhelm, S. W., Wood, S. A., & Paerl,
528 H. W. (2016). A review of the global ecology, genomics, and biogeography of the toxic
529 cyanobacterium, Microcystis spp. *Harmful algae*, *54*, 4-20.
530 Harris, J. H. (1995) The use of fish in ecological assessments. *Australian Journal of Ecology*
531 *20*(1), 65-80.
532 Hayden, H. L., Rochfort, S. J., Ezernieks, V., Savin, K. W., & Mele, P. M. (2019). Metabolomics
533 approaches for the discrimination of disease suppressive soils for Rhizoctonia solani
534 AG8 in cereal crops using 1H NMR and LC-MS. *Science of the Total Environment*, *651*,
535 1627-1638.
536 Hillyer, K. E., Raes, E., Karsh, K., Holmes, B., Bissett, A., & Beale, D. J. (2022). Metabolomics as
537 a tool for in situ study of chronic metal exposure in estuarine invertebrates.
538 *Environmental Pollution*, *292*, 118408.
539 Huang, S. S.; Benskin, J. P.; Veldhoen, N.; Chandramouli, B.; Butler, H.; Helbing, C. C.;
540 Cosgrove, J. R. (2017) A multi-omic approach to elucidate low-dose effects of
541 xenobiotics in zebrafish (*Danio rerio*) larvae. *Aquatic toxicology 182*, 102-112.
542 Janssen, E. M. L. (2019) Cyanobacterial peptides beyond microcystins–A review on co-
543 occurrence, toxicity, and challenges for risk assessment. *Water Research*, *151*, 488-
544 499.
545 Jeffries, K. M., Jackson, L. J., Ikonomou, M. G., & Habibi, H. R. (2010). Presence of natural and
546 anthropogenic organic contaminants and potential fish health impacts along two
547 river gradients in Alberta, Canada. *Environmental Toxicology and Chemistry*, *29*(10),
548 2379-2387.
549 Jones, O. A., Maguire, M. L., Griffin, J. L., Dias, D. A., Spurgeon, D. J., & Svendsen, C. (2013).
550 Metabolomics and its use in ecology. *Austral Ecology*, *38*(6), 713-720.
551 Jones, M. R., Pinto, E., Torres, M. A., Dörr, F., Mazur-Marzec, H., Szubert, K., ... & Janssen, E.
552 M. L. (2021). CyanoMetDB, a comprehensive public database of secondary
553 metabolites from cyanobacteria. *Water Research*, *196*, 117017.
554 Jousse, C., Dalle, C., Abila, A., Traïkia, M., Diogon, M., Lyan, B., ... & Delbac, F. (2020). A
555 combined LC-MS and NMR approach to reveal metabolic changes in the hemolymph
556 of honeybees infected by the gut parasite Nosema ceranae. *Journal of Invertebrate*
557 *Pathology*, *176*, 107478.
558 Kim Tiam, S.; Gugger, M.; Demay, J.; Le Manach, S.; Duval, C.; Bernard, C.; Marie, B. (2019)
559 Insights into the Diversity of Secondary Metabolites of *Planktothrix* Using a Biphasic
560 Approach Combining Global Genomics and Metabolomics. *Toxins 11*(9), 498.
561 Labine, L. M., & Simpson, M. J. (2020). The use of nuclear magnetic resonance (NMR) and
562 mass spectrometry (MS)–based metabolomics in environmental exposure
563 assessment. *Current Opinion in Environmental Science & Health*, *15*, 7-15.
564 Lankadurai, B. P.; Nagato, E. G.; Simpson, M. J. (2013) Environmental metabolomics: an
565 emerging approach to study organism responses to environmental stressors.
566 *Environmental Reviews 21*(3), 180-205.

Le Manach, S.; Sotton, B.; Huet, H.; Duval, C.; Paris, A.; Marie, A.; ... ; Marie, B. (2018) Physiological effects caused by microcystin-producing and non-microcystin producing *Microcystis aeruginosa* on medaka fish: A proteomic and metabolomic study on liver. *Environmental pollution 234*, 523-537.

Le Manach, S., Duval, C., Marie, A., Djediat, C., Catherine, A., Edery, M., ... & Marie, B. (2019) Global metabolomic characterizations of *Microcystis* spp. highlights clonal diversity in natural bloom-forming populations and expands metabolite structural diversity. *Frontiers in microbiology*, *10*, 791.

Li, H., Jing, T., Li, T., Huang, X., Gao, Y., Zhu, J., ... & Mu, W. (2021). Ecotoxicological effects of pyraclostrobin on tilapia (Oreochromis niloticus) via various exposure routes. *Environmental Pollution*, *285*, 117188.

Lohr, K. E.; Khattri, R. B.; Guingab-Cagmat, J.; Camp, E. F.; Merritt, M. E.; Garrett, T. J., Patterson, J. T. (2019) Metabolomic profiles differ among unique genotypes of a threatened Caribbean coral. *Scientific reports 9*(1), 1-11.

Lorenz, C. M. (2003) Bioindicators for ecosystem management, with special reference to freshwater systems. In *Trace Metals and other Contaminants in the Environment* (Vol. 6, pp. 123-152). Elsevier.

Mallik, A., Xavier, K. M., Naidu, B. C., & Nayak, B. B. (2021). Ecotoxicological and physiological risks of microplastics on fish and their possible mitigation measures. *Science of The Total Environment*, *779*, 146433.

Maloufi, S.; Catherine, A.; Mouillot, D.; Louvard, C.; Couté, A.; Bernard, C.; Troussellier, M. (2017) Environmental heterogeneity among lakes promotes hyper β-diversity across phytoplankton communities. *Freshwater Biology*, *61*(5), 633-645.

Marchand, J.; Martineau, E.; Guitton, Y.; Le Bizec, B.; Dervilly-Pinel, G.; Giraudeau, P. A (2018) multidimensional ^{1}H NMR lipidomics workflow to address chemical food safety issues. *Metabolomics 14*(5), 60.

Marie, B. (2020) Disentangling of the ecotoxicological signal using “omics” analyses, a lesson from the survey of the impact of cyanobacterial proliferations on fishes. *Science of The Total Environment 736,* 139701.

Marjan, P., Bragg, L. M., MacLatchy, D. L., Servos, M. R., & Martyniuk, C. J. (2017). How does reference site selection influence interpretation of omics data?: Evaluating liver transcriptome responses in male rainbow darter (*Etheostoma caeruleum*) across an urban environment. *Environmental science & technology*, *51*(11), 6470-6479.

Martin, J. C.; Maillot, M.; Mazerolles, G.; Verdu, A.; Lyan, B.; Migne, C.; ... ; Manach, C. (2015) Can we trust untargeted metabolomics? Results of the metabo-ring initiative, a large-scale, multi-instrument inter-laboratory study. *Metabolomics 11*(4), 807-821.

Martyniuk, C. J., & Simmons, D. B. (2016). Spotlight on environmental omics and toxicology: a long way in a short time. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, *19*, 97-101.

McClanahan, T. R.; Schroeder, R. E.; Friedlander, A. M.; Vigliola, L.; Wantiez, L.; Caselle, J. E.; ...; Oddenyo, R. M. (2019) Global baselines and benchmarks for fish biomass:

comparing remote reefs and fisheries closures. *Marine Ecology Progress Series*, *612*, 167-192.

Meador, J. P.; Bettcher, L. F.; Ellenberger, M. C.; Senn, T. D. (2020) Metabolomic profiling for juvenile Chinook salmon exposed to contaminants of emerging concern. *Science of The Total Environment 747*, 141097.

Melvin, S. D., Lanctôt, C. M., Doriean, N. J., Bennett, W. W., & Carroll, A. R. (2019). NMR-based lipidomics of fish from a metal (loid) contaminated wetland show differences consistent with effects on cellular membranes and energy storage. *Science of the Total Environment*, *654*, 284-291.

Mendez, K. M., Reinke, S. N., & Broadhurst, D. I. (2019). A comparative evaluation of the generalised predictive ability of eight machine learning algorithms across ten clinical metabolomics data sets for binary classification. *Metabolomics*, *15*(12), 1-15.

Nothias, LF., Petras, D., Schmid, R. *et al.* (2020) Feature-based molecular networking in the GNPS analysis environment. *Nat Methods* 17, 905–908.

Pomfret, S. M.; Brua, R. B.; Izral, N. M.; Yates, A. G. (2020) Metabolomics for biomonitoring: an evaluation of the metabolome as an indicator of aquatic ecosystem health. *Environmental Reviews*, *28*(1), 89-98.

Reverter, M.; Sasal, P.; Banaigs, B.; Lecchini, D.; Lecellier, G.; Tapissier-Bontemps, N. (2017) Fish mucus metabolome reveals fish life-history traits. *Coral Reefs 36*(2), 463-475.

Rohart, F.; Gautier, B.; Singh, A.; Lê Cao, K. A. (2017) mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS computational biology 13*(11), e1005752.

Roques, S.; Deborde, C.; Richard, N., Marchand; Y., Larroquet, L.; Prigent, S.; ...; Fauconneau, B. (2020) Proton-NMR Metabolomics of Rainbow Trout Fed a Plant-Based Diet Supplemented with Graded Levels of a Protein-Rich Yeast Fraction Reveal Several Metabolic Processes Involved in Growth. *J. Nutrition 150*(9), 2268-2277.

Rosato, A., Tenori, L., Cascante, M., Carulla, P. R. D. A., Dos Santos, V. A. M., & Saccenti, E. (2018). From correlation to causation: analysis of metabolomics data using systems biology approaches. *Metabolomics*, *14*(4), 1-20.

Sardans, J., Peguelas, J., Rivas-Ubach, A., 2011 Ecological metabolomics: Overview of current developments and future challenges. *Chemoecology* 21(4), 191–225.

Silva, R. A., Pereira, T. C., Souza, A. R., & Ribeiro, P. R. (2020). 1H NMR-based metabolite profiling for biomarker identification. *Clinica Chimica Acta*, *502*, 269-279.

Simmons, D. B., Benskin, J. P., Cosgrove, J. R., Duncker, B. P., Ekman, D. R., Martyniuk, C. J., & Sherry, J. P. (2015). Omics for aquatic ecotoxicology: Control of extraneous variability to enhance the analysis of environmental effects. *Environmental toxicology and chemistry*, *34*(8), 1693-1704.

Singh, A.; Shannon, C. P.; Gautier, B.; Rohart, F.; Vacher, M.; Tebbutt, S. J.; Lê Cao, K. A. (2019) DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics 35*(17), 3055-3062.

Sotton, B., Paris, A., Le Manach, S., Blond, A., Lacroix, G., Millot, A., ... & Marie, B. (2017a) Metabolic changes in Medaka fish induced by cyanobacterial exposures in mesocosms: an integrative approach combining proteomic and metabolomic analyses. *Scientific reports*, *7*(1), 1-13.

Sotton, B., Paris, A., Le Manach, S., Blond, A., Lacroix, G., Millot, A., ... & Marie, B. (2017b) Global metabolome changes induced by cyanobacterial blooms in three representative fish species. *Science of the Total Environment*, *590*, 333-342.

Sotton, B.; Paris, A.; Le Manach, S.; Blond, A.; Duval, C.; Qiao, Q.; ...; Marie, B. (2019) Specificity of the metabolic signatures of fish from cyanobacteria rich lakes. *Chemosphere 226*, 183-191.

Thoré, E. S., Philippe, C., Brendonck, L., & Pinceel, T. (2021). Towards improved fish tests in ecotoxicology-efficient chronic and multi-generational testing with the killifish Nothobranchius furzeri. *Chemosphere*, 129697.

Viant, M. R.; Ebbels, T. M.; Beger, R. D.; Ekman, D. R.; Epps, D. J.; Kamp, H.; ...; Rocca-Serra, P. (2019) Use cases, best practice and reporting standards for metabolomics in regulatory toxicology. *Nature communications 10*(1), 1-10.

Villeneuve, D. L., Blackwell, B. R., Cavallin, J. E., Cheng, W. Y., Feifarek, D. J., Jensen, K. M., ... & Ankley, G. T. (2021). Case Study in 21st Century Ecotoxicology: Using In Vitro Aromatase Inhibition Data to Predict Short-Term In Vivo Responses in Adult Female Fish. *Environmental Toxicology and Chemistry*, *40*(4), 1155-1170.

Walker, T. W., Alexander, J. M., Allard, P. M., Baines, O., Baldy, V., Bardgett, R. D., ... & Salguero-Gómez, R. (2022). Functional Traits 2.0: The power of the metabolome for ecology. *Journal of Ecology*, *110*(1), 4-20.

Wei, F.; Sakata, K.; Asakura, T.; Date, Y.; Kikuchi, J. (2018) Systemic homeostasis in metabolome, ionome, and microbiome of wild yellowfin goby in estuarine ecosystem. *Scientific reports 8*(1), 1-12.

Zhang, X. W.; Li, Q. H.; Dou, J. J. (2020) Mass spectrometry-based metabolomics in health and medical science: a systematic review. *RSC Advances 10*(6), 3092-3104.

Legends of figures

Figure 1. Visualization of the dataset structuration for LC-MS metabolomics (comprising 132 fish and 1252 analytes) on PCA (A) and heatmap with hierarchical clustering (B) for *Perca* (n=78) and *Lepomis* (n=54) fish collected during the 7-10th of September 2015 within 8 pounds of the peri-urban Paris' area.

Figure 2. Individual plot of PCA (A and C) and heatmap with hierarchical classification (B and D) of LC-MS liver metabolomes of *Perca* (n=78; A-B) and *Lepomis* (n=54; C-D) collected during the 7-10th of September 2015 within 8 pounds of the peri-urban Paris' area.

Figure 3. Phytoplankton composition estimated by BBE measurment of the 8 lake sub-surface water (A), corresponding cyanobacteria relative composition for Fon, Tri, Var and Ver (B), and molecular networking of metabolites extracted from the filtered biomass of the respective water of the 8 lakes generated with *t*-SNE algorithm, with cyanobacteria peptide clusters indicated in bold and potentially more noxious family clusters (microcystin, cyanopeptolmin and anabaenopeptin) indicated in red (C).

Figure 4. Individual plots of PLS-DA, testing the difference between the different lake groups presenting, or not, high content of potentially noxious cyanobactérial metabolites (Cha/Cer/Fon/Mau/Rue and Tri/Var/Ver, respectively) according to the LC-MS metabolomes of *Perca* (A) and *Lepomis* (B), and corresponding list of annotated best VIP list (metabolite score > 1.5) of *Perca* and *Lepomis* metabolomes (C). Accuracy, predictability and quality performances and number of VIP metabolite score > 1.5) for sampling lake discrimination are indicated for each species. Dots between brackets indicate when the metabolite concentration is more important in the Cha/Cer/Fon/Mau/Rue (green) or Tri/Var/Ver (red) lake group. * indicates that this metabolite was also observed as being negatively correlated with cyanobacteria occurrence within the different lake according to NMR analysis (Sotton et al., 2019).

Figure 1.

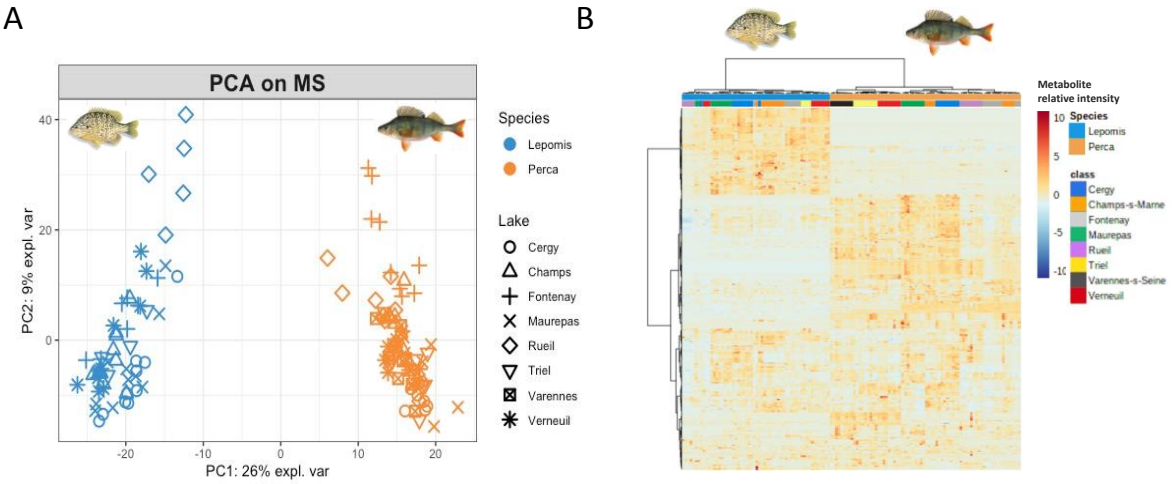

A
PCA on MS
PC2: 9% expl. var
PC1: 26% expl. var
Species
Lepomis
Perca
Lake
Cergy
Champs
Fontenay
Maurepas
Rueil
Triel
Varennes
Verneuil
B
Metabolite relative intensity
Species
Lepomis
Perca
class
Cergy
Champs-s-Marne
Fontenay
Maurepas
Rueil
Triel
Varennes-s-Seine
Verneuil

Figure 2.

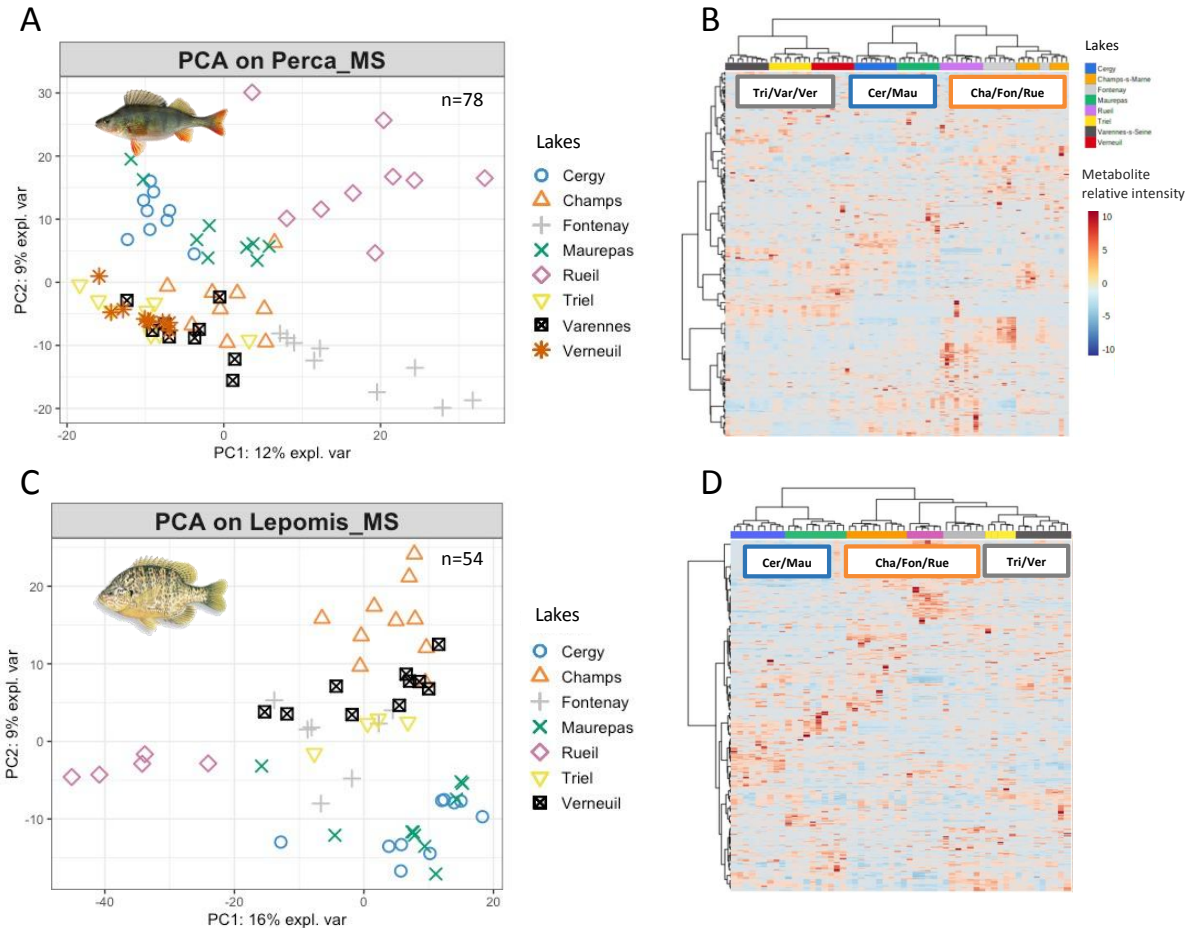

A
PCA on Perca_MS
n=78
PC2: 9% expl. var
PC1: 12% expl. var
Lakes
Cergy
Champs
Fontenay
Maurepas
Rueil
Triel
Varennes
Verneuil
B
Tri/Var/Ver
Cer/Mau
Cha/Fon/Rue
Lakes
Metabolite relative intensity
C
PCA on Lepomis_MS
n=54
PC2: 9% expl. var
PC1: 16% expl. var
Lakes
Cergy
Champs
Fontenay
Maurepas
Rueil
Triel
Verneuil
D
Cer/Mau
Cha/Fon/Rue
Tri/Ver

Figure 3.

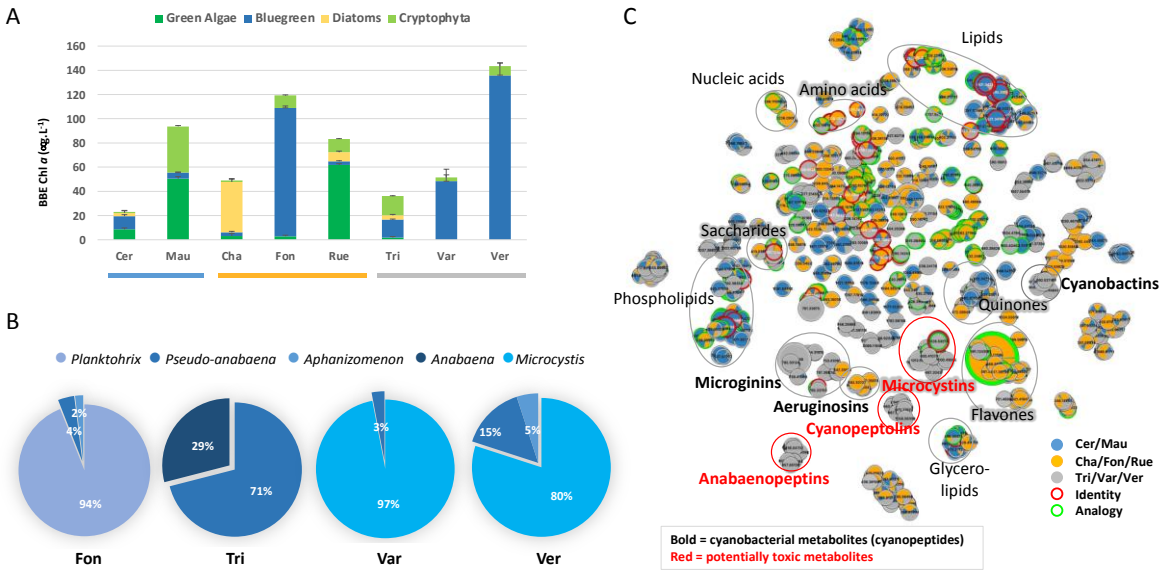
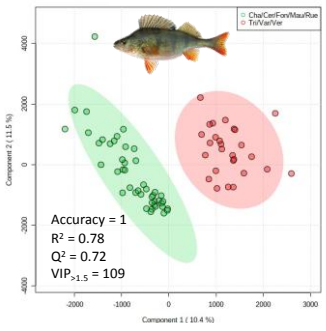
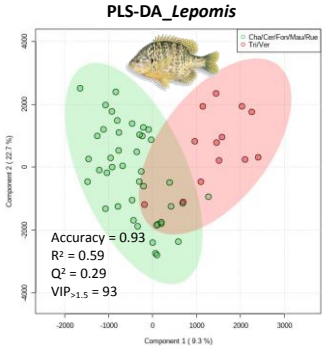

A
Green Algae
Bluegreen
Diatoms
Cryptophyta
BBE Chl a (ng.L-1)
160
140
120
100
80
60
40
20
0
Cer
Mau
Cha
Fon
Rue
Tri
Var
Ver
B
Planktothrix
Pseudo-anabaena
Aphanizomenon
Anabaena
Microcystis
2%
4%
94%
29%
71%
3%
97%
15%
5%
80%
Fon
Tri
Var
Ver
C
Lipids
Nucleic acids
Amino acids
Saccharides
Phospholipids
Cyanobactins
Quinones
Microginins
Microcystins
Aeruginosins
Flavones
Cyanopeptolins
Anabaenopeptins
Glycero-
lipids
Cer/Mau
Cha/Fon/Rue
Tri/Var/Ver
Identity
Analogy
Bold = cyanobacterial metabolites (cyanopeptides)
Red = potentially toxic metabolites

Figure 4.

A



B



C

MW	Annotation	Comp. 1_*Perca*	Comp. 1_*Lepomis*
267.09601	Adenosine	9.33 (●<●)	14.3 (●<●)
347.06212	Adenosine 5′-mono-phosphate	2.02 (●<●)	7.64 (●<●)
307.08355	Glutathione (reduced)	7.49 (●<●)	5.28 (●<●)
335.11488	Trp-Met	5.64 (●<●)	0.17
203.11734	Acetyl-carnitine	5.27 (●<●)	2.25 (●<●)
284.13698	Analog: Isoflavonoid	4.84 (●<●)	1.53 (●<●)
463.07428	Adenylosuccinic acid	4.06 (●<●)	0.07
612.15076	Glutathione (oxidized)	3.82 (●<●)	0.05
245.16244	Valeryl-carnitine*	3.76 (●<●)	1.64 (●<●)
463.07419	Adenylosuccinic acid	3.46 (●<●)	1.99 (●<●)
383.10736	Analog: Cyanidin arabonosine	3.41 (●<●)	0.41
567.33206	Lysophosphatidylcholine 22:6	3.34 (●<●)	1.60 (●<●)
626.16765	Analog: N-acetylglutamic acid	1.51 (●<●)	3.08 (●<●)
181.07481	Tyr	1.55 (●<●)	2.67 (●<●)
250.02948	Analog: N-acetylglutamic acid	2.57 (●<●)	0.03
300.13247	Analog: Hydroxyflavanone	2.50 (●<●)	1.03
559.07138	Adenosine 5'-diphosphoribose	2.49 (●<●)	0.59
404.00264	Uridine 5'-diphosphate	2.40 (●<●)	0.55
135.05505	Adenine	1.77 (●<●)	2.38 (●<●)
607.08199	Uridine-5-diphosphoacetylgalactosamine	2.37 (●<●)	0.14
363.05757	Guanosine 5-monophosphate	2.33 (●<●)	1.59 (●<●)
324.10545	Analog: 6-O-α-D-glucopyranosyl-D-fructose	1.89 (●<●)	2.30 (●<●)
392.16974	Trp-Trp	0.02	2.29 (●<●)
321.06967	Cytidine triphosphate	2.28 (●<●)	0.19
278.15223	Tyr-Pro	2.21 (●<●)	0.02
521.19585	*sn*-glycero-3-phosphocholine	1.69 (●<●)	2.21 (●<●)
665.12496	1,4-Dihydronicotinamide adenine dinucleotide	1.54 (●<●)	2.16 (●<●)
204.07959	Trp	2.13 (●<●)	0.10
242.06799	Thymidine	2.12 (●<●)	0.36
283.09143	Guanosine	0.44	2.08 (●<●)
429.14971	Analog: Adenosine 5'-monophosphate	1.54 (●<●)	2.06 (●<●)
257.10277	*sn*-glycero-3-phosphocholine	2.04 (●<●)	0.60
325.37028	N-Oleoylethanolamine	2.03 (●<●)	0.77