# Specificity of the metabolic signatures of fish from cyanobacteria rich lakes

Benoît Sotton, Alain Paris, Séverine Le Manach, Alain Blond, Charlotte Duval, Qin Qiao, Arnaud Catherine, Audrey Combes, Valérie Pichon, Cécile Bernard, et al.

# Specificity of the metabolic signatures of fish from cyanobacteria rich lakes

*Benoît Sotton[a], Alain Paris[a], Séverine Le Manach[a], Alain Blond[a], Charlotte Duval[a], Qin Qiao[a], Arnaud Catherine[a], Audrey Combes[b], Valérie Pichon[b], Cécile Bernard[a], Benjamin Marie[a]**

[a]UMR 7245 MNHN/CNRS Molécules de communication et adaptation des microorganismes, Muséum National d'Histoire Naturelle, 12 rue Buffon, F-75231 Paris Cedex 05, France.
[b]Department of Analytical, Bioanalytical Sciences and Miniaturization (LSABM), UMR CNRS-ESPCI Paris, CBI 8231, PSL Research University, ESPCI Paris, 10 rue Vauquelin, Paris, France.

* Correspondence and requests for materials should be addressed to B.M. (bmarie@mnhn.fr)

## Abstract

With the increasing impact of the global warming, occurrences of cyanobacterial blooms in aquatic ecosystems are becoming a main worldwide ecological concern. Due to their capacity to produce potential toxic metabolites, interactions between the cyanobacteria, their cyanotoxins and the surrounding freshwater organisms have been investigated during the last past years. Non-targeted metabolomic analyses have the powerful capacity to study simultaneously a high number of metabolites and thus to investigate in depth the molecular signatures between various organisms encountering different environmental scenario, and potentially facing cyanobacterial blooms.

In this way, the liver metabolomes of two fish species (*Perca fluviatilis* and *Lepomis gibbosus*) colonizing various peri-urban lakes of the Île-de-France region displaying high biomass of cyanobacteria, or not, were investigated. The fish metabolome hydrophilic fraction was analyzed by $^{1}$H-NMR analysis coupled with Batman peak treatment for the quantification and the annotation attempt of the metabolites. The results suggest that similar metabolome profiles occur in both fish species, for individuals collected from cyanobacterial blooming lakes compared to organism from non-cyanobacterial dominant environments. Overall, such environmental metabolomic pilot study provides new research perspectives in ecology and ecotoxicology fields, and may notably provide new information concerning the cyanobacteria/fish ecotoxicological interactions.



**Keywords:** fish; cyanobacteria; metabolomics; ecotoxicology; NMR

# 1. Introduction

Direct and indirect anthropogenic disturbances affect the global functioning and stability of freshwater ecosystems. Global warming, physical alteration, habitat loss, water withdrawal, pollutions, overexploitation and the introduction of non-native species constitute the main threats affecting freshwater ecosystems and their biocenoses (Dudgeon et al., 2006; Bunn 2016; Revenga et al., 2005). Over the past century, transformation of natural landscapes for industrial and urban needs have led to a general increase of chemical nutrients concentrations in freshwater ecosystems, and notably of phosphorus and nitrogen species that have caused the massive eutrophication of freshwater bodies (Schwarzenbach et al., 2010). One of the main consequences of this non-natural eutrophication process is a shift in the composition of the primary producers due to the appearance of noxious autotrophic bacteria, such as cyanobacteria (Chorus and Bartram 1999; O'Neil et al., 2012). In worldwide freshwater ecosystems, cyanobacteria are mainly present during summers, forming large surface scums generally accompanied by the presence of toxic compounds that have been pointed out to affect all trophic food web compartments (Codd et al., 2005; Ferrao-Filho et al., 2011; Sotton et al., 2015). *Microcystis*, *Aphanizomenon*, *Anabaena*, *Cylindrospermopsis* and *Planktothrix* are the main bloom-forming freshwater cyanobacterial genera found during summers in lentic waterbodies, and all have been often reported to produce potent toxic compounds, so-called cyanotoxins (Codd et al., 2017). Among them, the microcystins (MCs), a family of hepatotoxins consisting of more than 230 variants, are the most studied due to their high biological activity and their wide occurrence during freshwater cyanobacterial blooms (Catherine et al., 2017). For exposed macro-organisms, MCs may induce inhibition of the protein phosphatases 1 (PP-1) and 2A (PP-2A) as well as occurrence of a cellular oxidative stress *via* the formation of reactive oxygen species (ROS), with different physiological consequences depending on the organism and the species studied (Amado et al., 2010; Malbrouck and Kestemont 2006). The effects of freshwater cyanobacteria and their respective cyanotoxins, notably the MCs, have been widely studied on fish maintained in microcosm or mesocosm conditions, and fish have been proposed as valuable indicators of environmental disturbances associated to cyanobacteria proliferations (Malbrouck and Kestemont 2006; Bols et al., 2001; Le Manach et al., 2018). However, the actual knowledge on the genuine cyanobacteria bloom impairs for natural populations of fishes is mainly deduced from short-time experimentation generally performed in micro- or mesocosms with high concentrations of purified toxins. Furthermore, most of these studies were focused on the mechanisms involved in the dynamic of the MC accumulation-detoxification. It appears that

there is still a lack in our understanding of the real ecotoxicological effects of cyanobacterial biomass, particularly on the natural ichtyofauna population. Indeed, cyanobacterial blooms are producing at the same time a "cocktail" of potentially bioactive compounds, and are also potentially modifying other important ecological parameters of the water bodies.

The metabolome of an organism is constituted by the set of primary metabolites synthesized at a given time. It represents its metabolic chemical picture, which can potentially be altered when the ecological and environmental stress conditions change (Fiehn 2002; Bundy et al., 2008). The observed metabolome regulations represent the chemical phenotypic response of an organism that could potentially response to counter-act the negative effects of stressors present in the ecosystems, and thus to adapt to the variable conditions (Hamilton et al., 2016). In this way, metabolomic studies have become a relevant approach to describe and analyze the integrated response of the organisms under specific environmental context and scenario (Franzosa et al., 2015; McLean 2013; Hultman et al., 2015). Then, the changes in the primary metabolite concentrations help to provide valuable and useful information concerning the physiological processes involved in the homeostatic or dys-regulated responses of the organisms encountering environmental stresses from potentially multiple origins. Nuclear magnetic resonance (NMR)-based metabolomics has been proved to be a powerful approach to address hypotheses relating to fish physiology or pollutant induced toxicity or diseases (Samuelsson and Larsson 2008; Viant 2008). However, despite its high potential for the understanding of the molecular mechanisms implicated in the ecotoxicological responses of organisms, the investigation of new qualitative and quantitative biomarkers characterizing the interactions between populations and their biotopes in various ecological context remains still rare and difficult to achieve (Sardans et al., 2013; Cappello et al., 2016; 2017; Qiao et al., 2016).

Thus, as toxic cyanobacterial blooms may represent important ecotoxicological and ecological constraints in freshwater environments, it can be supposed that organisms exposed to cyanobacterial blooms exhibit characteristic metabolome signatures, compared to others non-exposed to cyanobacterial dominant conditions. Their metabolism may respond and/or counter-act the potential negative effects of the cyanobacteria and thus adapt to the local environmental pressures. To date, no studies have been carried out using a global metabolomic approach in order to assess the specific metabolic changes that could be observed in fish exposed to cyanobacterial blooms in contrasted aquatic natural ecosystems. In this way, during the summer 2015, young fish of two representative and common species of freshwater lakes from the

European temperate regions, the perch (*Perca fluviatilis*) and the pumpkinseed sunfish (*Lepomis gibbosus*), have been sampled in eight peri-urban lakes of the Île-de-France region contrasted by their phytoplanktonic community composition ("presence" or "absence" of cyanobacterial blooms). $^{1}$H-NMR metabolomic analyses were performed on the fish liver in order to investigate the global metabolome local specificities of the two fish species collected from a gradient of distinct ecological contexts (comprising the cyanobacteria dominance) and to further identify the metabolic signatures related to these potential specific phenotypic responses.

## 2. Experimental section

### *2.1. Lakes and physico-chemical parameters measurements*

*In-situ* sampling campaigns were performed during summer 2015 in eight lakes of the Île-de-France region (France), chosen for their different dominant phytoplankton communities already described in previous studies and notably the presence or the absence of recurrent cyanobacterial blooms (Catherine et al., 2008; Maloufi et al., 2016). Cergy "les Galets" lake (C), Champs-sur-Marne "la Sablière" lake (CM), Maurepas "la Courance" lake (M), Rueil "Saint Cucufa" lake (R), Verneuil "la grosse Pierre" lake (V), Varenne-sur-Seine "Le grand Marais" lake (VS), Fontenay-sur-loing "le grand Fontenay" lake (F) and Triel lake (T) (Supplementary figure S1) were thus studied and sampled with electric fishing device for capturing fish alive. Previous investigation of the fish guild populating every one of these lakes (using both electric device and fish net sampling) had indicated that only the perch (*Perca fluviatilis*) and pumpkinseed sunfish (*Lepomis gibbosus*) were presents in all or almost all the lakes and were further selected as bioindicative species for further metabolomic analyses.

In all lakes, dissolved oxygen ($O_2$) concentrations, pH, temperatures and conductivity were measured in the water column using a multiparameter probe (YSI EXO2) (Supplementary table S1).

### *2.2. Phytoplankton sampling*

In every lake, sub-surface chlorophyll-*a* equivalent concentrations attributed to the four-main phytoplankton groups (Chlorophyta, Diatoms, Cyanobacteria, Cryptophyta) were measured with an *in-situ* fluorometer (Fluoroprobe II, Bbe-Moldenke, Germany) and samples of water were filtered through 1.2 μm GF/C filters (Nucleopore, Whatman) and stored at -80°C until total chlorophyll-*a* concentrations analyses using the ethanolic extraction as described by

Yepremian and co-workers (2017). Samples of water were fixed in Lugol iodine solution and kept at 4°C until the identification. The estimation (%) of the abundance of the different cyanobacterial genera was performed using an Utermohl's counting chamber and an inverted microscope as described by Catherine et al. (2016). In parallel, phytoplankton biomass was collected using an Apstein's type phytoplankton net (20-µm mesh size) and kept at -80°C until metabolite characterization by mass spectrometry (MS) and MCs content analyses.

*2.3. Metabolite characterization of phytoplanktonic biomass by mass spectrometry*

The collected biomass from the sampled lakes were freeze-dried and then sonicated in 75% methanol, centrifuged at 4°C (15,000 g, 10 min). This step was performed twice, and the pooled supernatants were collected with a Pasteur pipet and kept at -80°C prior to analysis. The supernatant was transferred and acidified with formic acid and 5 µL were analyzed in triplicats on an HPLC (Ultimate 3000, ThermoFisher Scientific) coupled to an electrospray ionization and quadrupole time-of-flight hybrid mass spectrometer (ESI-QqTOF, QStar® Pulsar i, Applied Biosystems®, France).

High-performance liquid chromatography (HPLC) of 5 $\mu$L of each of the metabolite extracts was performed on a $C_{18}$ column (Discovery® Bio wide pore, 15cm x 1mm, 5 µm, Sigma) at a 50 $\mu L.min^{-1}$ flow rate with a gradient of acetonitrile in 0.1% formic acid (10 to 80% for 60 min). The metabolite contents were then analyzed on positive mode using information dependent acquisition (IDA), which allows switching between mass spectrometry (MS) and tandem mass spectrometry (MS/MS) experiments, as previously described (Marie et al., 2012). The data was acquired and analyzed with the Analyst QS software (Version 1.1). Peak lists were generated from MS spectra acquired between 10 and 55 min, filtering noise threshold at 2% maximal intensity and combining various charge states and related isotopic pattern. Metabolite annotation was attempted according to the accurate mass of the molecules, then to their respective MS/MS fragmentation pattern with regard to an in-house database of more than 700 cyanobacterial metabolites.

A molecular network was created using the online workflow at GNPS (Wang et al., 2016). The data was filtered by removing all MS/MS peaks within +/- 17 Da of the precursor m/z. The data was then clustered with MS-Cluster with a parent mass tolerance of 1.0 Da and a MS/MS fragment ion tolerance of 0.5 Da to create consensus spectra. Further, consensus spectra that contained less than 2 spectra were discarded. A network was then created where edges were filtered to have a cosine score above 0.6 and more than 5 matched peaks. Further edges between two nodes were kept in the network if and only if each of the nodes appeared in each other's

respective top 10 most similar nodes. The spectra in the network were then searched against GNPS' spectral libraries. The library spectra were filtered in the same manner as the input data. All matches kept between network spectra and library spectra were required to have a score above 0.6 and at least 5 matched peaks. All GNPS results are available at http://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=a158701fa7894ab08c0df4c60994ccf1

*2.4. Cyanotoxin quantification*

The 75% methanol extracts of the phytoplankton biomass were further analyzed by ELISA (enzyme-linked immunosorbent assay), with the microcystins (Adda-specific) Kit (AD4G2 antibody, Abraxis), in order to globally measure the total MC content, and integrating the signal obtained for all MC variants in presence. Prior to analysis, samples were dissolved with the ELISA sample diluent to reach a methanol concentration below 5% to avoid any interactive effect and to stay in the detection range of the kit (0.1-5 $\mu g.L^{-1}$) for all samples. The results were expressed in microcystin-LR equivalents (µg eq. MC-LR $.mg^{-1}$ dry weight).

The free and bound β-Methylamino-L-alanine (BMAA) were also measured in the phytoplankton biomass according to the hydrophobic interaction liquid chromatography (HILIC)/MS-MS based methods described previously (Combes et al., 2014; Faassen et al., 2016). This method used solid-phase extraction based on mixed mode sorbent to concentrate and clean up the phytoplankton extract that contained free BMAA. After the acidic lysis of the phytoplankton biomass, the bound fraction of BMAA was analyzed by LC/MS-MS. This quantitative method has proved to be specific and reliable in a range of concentration level from 0.25 to 1.6 ng $mg^{-1}$ dry weight.

*2.5. Fish sampling and tissue extraction procedure for metabolomic analysis*

Immature individuals of perch (*Perca fluviatilis*) and pumpkinseed sunfish (*Lepomis gibbosus*), two representative fish species of European freshwater lakes (that were selected according to their presence in almost, if not, all investigated lakes), were targeted by electric fishing (FEG 8000, EFKO, Leutkirch, Germany) performed in the riparian area of every lake. Alive caught fish (n<6-10 per lake and per species) were directly measured, weighed (Table S2), briefly euthanized by neck dislocation and then liver of each individual was shortly sampled, deep-frozen in liquid nitrogen and kept at -80°C until metabolomics analyses, in accordance with animal ethical concerns and regulations.

Individual liver metabolome extraction was carried out using the methanol/chloroform/water (ratio 2/2/1.8) method, on the basis of existing literature (Lin et al., 2007; Wu et al., 2008).

Briefly, fresh frozen livers were weighed, homogenized in ice-cold methanol (8 mL per gram of tissue, AnalaR Normapur, min. 99.8 %, VWR, Pennsylvania, USA) and ice cold milliQ water (2.5 $mL.g^{-1}$), and then vortexed for 1 min. Subsequently, ice cold chloroform (4 $mL.g^{-1}$, Normapur, 99.3 %, VWR, Pennsylvania, USA) and milliQ water (4 $mL.g^{-1}$) were added to extract the hydrophobic metabolites. Then, the mixture was vortexed for 1 min and incubated on ice for 10 min to obtain a complete solvent partition. The resulting supernatant was then centrifuged at 4°C for 10 min at 2,000 g, resulting in a biphasic solution. The upper polar and lower non-polar layers were carefully removed. The upper polar fraction was then transferred to 2-mL Eppendorf tubes, dried under Speed-vac device (Speed-vac Plus SC110A, Savant) and then kept at -80°C until NMR analysis. Prior to $^{1}$H-NMR measurement, the polar tissue extracts were dissolved in 550 µL of 0.1 M sodium phosphate buffer (pH 7.0) prepared in $D_2O$ (10% v/v) containing 0.25 mM sodium-3-tri-methylsilylpropionate (TMSP) as an internal standard. Finally, the resulting samples were transferred to 5-mm NMR tubes (Norell, France) and immediately analyzed by $^{1}$H-NMR.

*2.6. $^{1}$H-NMR spectroscopy and metabolite quantification*.

All NMR data were recorded at 298 K on a 600 MHz Bruker AVANCE III HD spectrometer equipped with a 5-mm TCI CryoProbe ($^{1}$H-$^{13}$C-$^{15}$N) with Z-gradient. One-dimensional $^{1}$H-NMR spectra were acquired using a standard Bruker noesygppr1d pulse sequence to suppress water resonance. Each spectrum consisted of 256 scans of 32,768 data points with a spectral width of 7.2 kHz, a relaxation delay of 2.5 s and an acquisition time of 2.3 s.

The relative metabolite quantification was performed using the BATMAN (an acronym for Bayesian AuTomated Metabolite Analyser for NMR spectra) R-package (Hoa et al., 2014), which deconvolutes peaks from 1-dimensional $^{1}$H-NMR spectra to automatically assign chemical shifts to specific metabolites from a target list and then estimate their respective concentrations. This can be achieved thanks to an implementation of a Bayesian-based modelling procedure. BATMAN uses, in a two-component joint model, resonances of every assigned proton from a list of catalogued metabolites (this library is provided with the BATMAN R-package, and includes the chemical shift information of various targeted metabolites measured at pH 7.0) and noisy information to finally reconstruct the empirical NMR spectrum. But, in absence of confirmatory analytical methods, they were further mentioned as candidate metabolic biomarkers sharing the same $^{1}$H-NMR parameters with the catalogued metabolites. Therefore, 222 potential metabolites were quantified from Bruker spectra files using the following parameters: i) the chemical shift regions belonging to the two

following regions: 0.5 to 4.60 ppm and 5.40 to 10.0 ppm, ii) 400 burn-in iterations, iii) 200 post-burn-in iterations and iv) 5000 iterations for batman rerun. Calculations were performed on a HP Z820 workstation using two 3.30 GHz Intel Xeon© CPU E5 processors and 64 Go RAM by activating 12-parallel threads processing.

### *2.7. Statistical exploration of data*

The mixOmics library (Rohart et al., 2017) was used to carry out the multivariate analyses (Tenenhaus et al., 2014). Regularized canonical correlation analysis (rCCA) is a multivariate statistical method used to assess correlations between two multivariate datasets acquired on the same individuals. Here, it was used as a factorial discriminant analysis that modeled the relationships between the species (*Perca* and *Lepomis*), the quantitative chemical data measured in the different lakes (total Chl-*a* concentrations, Chl-*a* estimated concentrations related to cyanobacterial biomass (BBEcya), MCs concentrations, $O_2$ concentrations, pH and conductivity), together with the lake of sampling, constituting the Y matrix, and the semi-quantitative levels of metabolites determined by the BATMAN algorithm with no normalization of the data (X matrix) (supplementary data 1). The *rcc()* function was used to define the canonical correlations and the canonical variates, the *network()* function was used to produce the network of interactions, using default parameters. Cross-validation of MANOVA results was obtained thanks to a bootstrap-based procedure with corrected *p*-value ($P<0.05$) by applying a random assignment of any statistical individual to a given group of treatments (Krishnamoorthy and Lu, 2010). In order to evaluate the effects of the experimental factors (species, chemical data and their interaction) on the relative concentrations of metabolites highlighted by multivariate analyses and the subsequent relevant networks, simple two-way ANOVAs followed by a Student-Newman-Keuls post-hoc test were performed.

## 3. Results

### *3.1. Phytoplankton, microcystins and chemical conditions of the studied lakes*

During 2015 summer sampling campaign (Supplementary fig. S1), distinct phytoplanktonic compositions and concentrations were observed in the eight lakes targeted (Fig. 1). Among them, F, T, V and VS, exhibited remarkable high and/or dominance of cyanobacterial-specific chlorophyll-*a* concentrations reaching 106, 14.8, 135 and 48.3 µg eq. Chl *a* $L^{-1}$, respectively (Fig. 1A). In the other lakes, chlorophyll-*a* concentrations of the total phytoplanktonic community ranged between 7.2 and 241.5 µg. Chl *a* $L^{-1}$ (supplementary table S1) but were

mainly dominated by chlorophytes, diatoms and/or cryptophytes (Fig. 1A), major phytoplanktonic phyla, that are so far not known to produce any toxins nor other ecological concerns as cyanobacteria does. In these lakes, only low concentrations of cyanobacterial-specific chlorophyll-*a*, ranging from 2.9 to 10.9 µg eq. Chl *a* $L^{-1}$, were measured (Fig. 1A). Furthermore, in the cyanobacteria-dominated lakes, different cyanobacterial genera were observed during samplings (Fig. 1B). F and T were dominated by *Planktothrix* (94% of the cyanobacteria present in the sample) and *Pseudanabaena* (71%), respectively, whereas V and VS, were both dominated by *Microcystis* (Fig 1B). Furthermore, we notice that the modeste blue green signal observed in C corresponds to 100% *Spirulina* sp., a cyanobacteria genus that is so far known not to produce any known toxic compounds, as it was further confirmed.

Microcystin concentration was investigated in order to estimate the amount of the cyanobacterial bioactive compounds present in the various lake water (Fig. 2; supplementary table S2). MC-LR equivalent concentrations were measured at concentrations ranging from above the quantification limit (> 1 µg.$g^{-1}$ DW) to 3367 µg.$g^{-1}$ DW eq. MC-LR. The highest MCs concentrations were found in lakes dominated by cyanobacteria and particularly in F, V and VS where 2067 µg.$g^{-1}$, 166 µg.$g^{-1}$ and 3367 µg.$g^{-1}$ MC-LR eq. were measured, respectively. Interestingly, V that exhibited the highest cyanobacterial concentrations (dominated by *Microcystis*) measured in this study, did not exhibit the highest MCs concentrations (3367 µg.$g^{-1}$ MC-LR eq. in VS) illustrating the fact that MCs concentrations are not linearly linked to the cyanobacterial biomass measured in the lakes, as it is also dependent on the production of metabolite by the dominant cyanobacterial clones, which may or not be producing MCs. On one side, none of these samples presented detectable amount of the potential neurotoxin BMAA (neither in free or bound form), saxitoxins nor anatoxins. On the other side, various classes of cyanobacterial secondary metabolites comprising different variants of hepatotoxic microcystins, together with various other potentially bioactive metabolite families, such as anabaenopeptins, aeruginosins, microginins and cyanopeptolins, were detected in the phytoplanktonic biomass of collected from the lakes where cyanotoxin-producing cyanobacteria reached higher concentrations (Fig. 2; supplementary figure S2-S3).

*3.2. Metabolomic analysis of sampled fish*

$^{1}$H-NMR raw files (supplementary fig. S4) were preprocessed thanks to the R BATMAN library (Hao et al., 2014) to get a relative quantification of a preselected set of metabolites replacing an analytical assignment based on different and complementary analyses. This approach allows for each individual spectrum to determine a metabolite quantification, even for very low

concentrated metabolites when compared to the main ones, as the differences in their relative concentrations could have been validated by both uni- and multivariate statistical tests (Hao et al., 2014). A multivariate analysis was performed by using a rCCA analysis between the matrix of the relative concentrations of the 222 metabolites (X) and the dummy matrix (Y) corresponding to the following variables (supplementary data S1): *i.e.* fish species, cyanobacterial-specific Chl-*a* and MCs concentrations, $O_2$, pH, conductivity and the total Chlorophyll-*a* concentrations and their interaction using the mixOmics package in R and a MANOVA bootstrap procedure applied to the dataset in order to highlight the significant effects of the environmental variables on the observed metabotypes (*i.e.* the specific metabolic profiles according to the environmental factors).

MANOVA bootstraps reveal significant effects of species ($p = 0$), cyanobacterial-specific chl-*a* (BBEcya) concentrations ($p < 10^{-43}$), MC concentrations ($p < 10^{-11}$), pH ($p < 10^{-28}$), $O_2$ ($p < 10^{-28}$), conductivity ($p < 10^{-47}$), total chlorophyll-a concentrations ($p < 0.011$) from the rCCA model. However, a clear species effect is observable through the dimension 1 ($p < 10^{-16}$) with *Perca* and *Lepomis* clearly separated by the first dimension whatever the lake considered (fig. 2A). On the dimension 2, a clear effect correlated with cyanobacterial biomass is observable ($F = 1931$, $p < 10^{-16}$), as fish coming from cyanobacterial dominated lakes (in green) are clearly separated by this dimension whatever the fish species considered (Fig. 3A). In addition, pH, $O_2$ concentrations and conductivity variables ($F = 599$ and $p < 10^{-16}$, $F = 867$ and $p < 10^{-16}$, *and* $F = 419$ and $p < 10^{-16}$, respectively) are also significantly correlated to the different metabotypes observed on this dimension but to a much lesser extent compared to cyanobacterial biomass contents (Fig. 3A). In addition, no significant interactive effects between species and the environmental factors could have been observed in this analysis suggesting that the metabotypes of the two species seem to be similarly driven by the different environmental factors.

The relevance network based on the second rCCA dimension specifically highlights metabolites discriminating the fish from dominated and non-dominated cyanobacteria lakes and linking to the correlated environmental factors (mainly BBEcya concentrations but also pH) (Fig. 3B). Overall, it appears that all these metabolites shown by this network exhibit negative correlation with both cyanobacterial concentrations and pH values. In addition, two-way ANOVA (Table 1) were performed on all putative metabolites highlighted in the figure 3B. A significant effect of the BBEcya and pH is shown for putative metabolites annotated as 16a-hydroxyestrone, acetylcholine, androsterone, glycerophosphocholine, isovalerylcarnitine, L-glutamine and L-palmitoylcarnitine whereas for glycerol, homo-L-arginine and L-arginine, a significant effect

of the BBEcya concentrations, only, have been confirmed by two-way ANOVA analyses (Table 1). Furthermore, no interaction effect between the species and the BBEcya concentrations or pH has been observed for any metabolites suggesting the occurrence, for both fish species and for these two discriminating environmental factors, the occurrence of a similar metabolic responses. Negative correlations between relative concentrations of certain metabolites and the BBEcya concentrations and the pH conditions are corresponding to lowest relative concentrations of these metabolites in the liver of fish captured in cyanobacterial dominated lakes (Table 1; supplementary figure S5-S6). Globally, it appears that liver of fish sampled in T, F and V lakes, and in VS lake in a lesser extent, exhibits the most significant differences in the relative concentrations of the highlighted metabolites compared to fish coming from other investigated lakes, where cyanobacteria are notably in low or not dominant proportions.

# Discussion

*4.1. Diversity of ecosystems: environmental, chemical and biological conditions*

Cyanobacterial blooms are frequent phenoma in peri-urban lakes from the Île-de-France region due to the eutrophic and hyper-eutrophic states of these environments. Interestingly, these freshwater aquatic ecosystems are characterized by a wide range of environmental conditions and anthropogenic pressures that influence the phytoplankton biomass and compositions present in these ecosystems (Maloufi et al., 2016; Catherine et al., 2010; 2016). During summer 2015 sampling, in the lakes from Rueil (R), Cergy (C), Maurepas (M) and Champs-sur-Marne (CM), diatoms or chlorophyta were dominant and cyanobacteria were present with chl-*a* concentrations never exceeding 5 µg eq. chl-*a* $L^{-1}$ (supplementary table S1). In the other lakes, the *Microcystis* genera appeared largely dominant, such as in Varenne-sur-Seine (VS) and in Verneuil (V) lakes with cyanobacterial-specific chl-*a* (BBEcya) concentrations reaching 135 and 48 µg eq. chl-*a* $L^{-1}$, respectively. In the Fontenay (F) lake, *Planktothrix* genera was found to be dominant with BBEcya concentration reaching 106 µg eq. chl-*a* $L^{-1}$. Interestingly, in Triel (T) lake, the cyanobacteria belonging to the *Pseudanabaena* and *Anabaena* genera were present at relatively lower but even substantial BBEcya concentrations (around 15 µg eq. chl-*a* $L^{-1}$). However, even though cyanobacterial amount in T appears maybe not as high as in other cyanobacterial-dominated lakes of this study (F, V and VS), the presence in this T lake of noticeable and recurrent amounts of other contaminants, such as heavy metals, polychlorinated biphenyls (PCBs) and polycyclic aromatic hydrocarbons (PAHs) was measured in both water and fish samples (Teil et al., 2014; Gaspéri et al., 2010; Azimi and Rocher 2016), and would

indicate why the fish present in the T lake exhibit an even more divergent metabolic profile, when compared to those of the non-cyanobacteria-dominated (R, C, M and CM) and the other cyanobacteria-dominated (V, VS and C) lakes.

*4.2. Which cyanobacterial-related parameters may influence the fish metabolome?*

Cyanobacteria are well known to produce a wide variety of potentially toxic secondary metabolites including MCs, that have been detected in remarkable amount especially in the VS and F lakes, together with various other cyanobacterial bioactive compounds, such as those that have been highlighted by the global network analysis, including cyanopeptolins, aeruginosins, microginins and anabaenopeptins. Previous experimental studies have been able to show that cyanobacterial exposure could induce significant physiological impairs in fish exhibiting different metabolomic dysregulation, in relation with the various secondary metabolite content of the investigated strains (Sotton et al., 2017a; 2017b). Nevertheless, the genuine ecological consequences of the production of this chemical bioactive (including potent toxic) metabolites by cyanobacteria remain largely uncovered.

Indeed, in addition to the production of such potentially toxic secondary metabolites, changes in physico-chemical parameters of the water-bodies have also been reported during cyanobacterial bloom episodes. Massive decrease of dissolved $O_2$ concentrations may occur during important bloom senescence events, due to the bacterial degradation of cyanobacterial cell biomass, inducing sometimes spectacular massive death of various fish species, also described as fish-kill phenomenon (Pearl and Paul 2011). During the present study, no noticeable depletion of the dissolved $O_2$ concentrations of neither surface nor bottom water or any fish-kill have been observed, indicating that the high cyanobacterial biomass (observed in V, VS, T and F lakes) were not in senescence. Furthermore, $O_2$ concentrations does not appear to significantly drive the liver metabolic differences observed in the fish species studied, suggesting that no metabolic disturbances linked to $O_2$ concentration variation were observed here. Among the other environmental factors measured on the studied lakes that are noticeably correlated with metabolic differences in fish livers based on rCCA analyses, the pH exhibits large variations between sampled aquatic ecosystems. Indeed, it appears that lakes dominated by cyanobacteria show higher pH values (above 9), excepted in Fontenay (F) lake where the value is more similar (between 7.1 and 8.3) to those found in lakes where cyanobacteria are in less abundance (supplementary table S1). Such high pH value phenomenon has been already reported to be directly related with cyanobacteria photosynthesis process that removes carbon dioxide from the water and increases hydroxide ion concentration (Lopez-Archilla et al., 2004).

However, in the present study, it is not possible to directly conclude whether the elevated pH observed in cyanobacterial dominated lakes is a consequence of the photosynthetic activity of cyanobacteria or is due to local geochemical conditions of the respective lake. To better characterize the causality of these elevated pH in further investigation, it should be suitable to monitor pH value before, during and at the end of the bloom. Further studies need now to be performed in order to disentangle the specific and/or synergic effects of the various bioactive metabolite production together with those of other physicochemical parameters associated with cyanobacterial blooms.

*4.3. Metabolic changes correlated to perturbed lakes: cyanobacterial concentrations and pH as environmental drivers of fish liver metabolome?*

Thanks to NMR and multivariate analyses, the present results show that comparable metabolic changes are observable in both fish species exposed to higher cyanobacterial biomass (F, VS, T and V lakes). in these analyses the most constraining factors correlated to the metabolic changes in fish were the cyanobacterial concentrations and the pH values that exhibit the strongest correlations with the changes of relative concentrations of various metabolites in the two fish species sampled. In this study, it is also very likely that the combination of both high pH values and high cyanobacterial biomass could lead to the more contrasted metabolome changes in fish. This would be in agreement with the fact that fish from F lake, presenting high cyanobacterial and MCs concentrations but moderate pH value around 7.5, exhibit intermediate metabolomes compared to fish captured in T, V and VS (high cyanobacterial and MCs concentrations and high pH) and fish captured in other lakes (C, CM, M, R), not dominated by cyanobacteria.

The present results indicate that MCs may not be in this case the most constraining factor for fish from those natural ecosystems. However, in addition to MC cyanobacteria produce a large set of bioactive compounds that could also influence the fish health and the complexity of such secondary metabolite mixtures aims at being more characterized in an ecotoxicological point of view. One should take also into account that fish may not necessarily be exposed by direct ingestion of cyanobacterial cells, and that trophic transfer may be another source of cyanobacterial bioactive (including toxic) metabolites (Sotton et al., 2014). It would be now interesting to consider other fish species feeding on phytoplankton in order to test whether they are indeed exposed to higher cyanotoxin contents, by direct ingestion of cyanobacterial cells, and whether they exhibit similar or even more drastic metabolome alterations than such carnivorous ones.

Interestingly, the fish collected from the T lake exhibits one of the most divergent global metabotypes, suggesting that it could potentially be influenced by the occurrence of multi-stress conditions, comprising cyanobacteria proliferations together with the other heavy metal, PCB and PAH contaminants, that have been observed in this specific environments (Catherine et al., 2016). Indeed, T lake was not characterized by the highest concentration of cyanobacteria observed during this study. However, due to the presence of other pollutants already monitored in past studies in this pound and not in the other one (Teil et al., 2014; Gaspéri et al., 2010; Azimi and Rocher 2016), this observation suggests that additive and/or synergistic effects of multi-pollutants together with cyanobacterial bloom seem to be involved in similar metabolic variations than those of fish from pounds which are the most stressed by cyanobacterial blooms. But such assumption now need further investigations before been fully validated. On order to fully validate potential metabolite biomarkers of cyanobacterial exposure pressure on fish liver metabolome, further investigation would now aim at investigating a larger set of lake from a wider geographic area, and also to firmly confirm the metabolite identification by orthogonal approaches such as 2-D NMR, spiking of standard, and or to develop in parallel mass spectrometry based metabolomic investigations.

## 5. Conclusion

The present study demonstrates that cyanobacterial blooms occurrence can locally be correlated with metabolic divergences in relation to the fish response to environmental stress. Such metabolomic analyses support experimentally observed processes but also give us innovative perspectives for the characterization of the environmental and genuine interactions between fish and cyanobacteria proliferations.

In this way, further studies should be performed in order to investigate longitudinal variations along seasonal variations, fish life history, populational genetic distance and the specific changes of fish metabolomes during bloom episodes, which could help us to better understand and characterize the main environmental factors potentially disturbing fish physiology. Also, the understanding of the ecosystem functioning under ecotoxicological pressure induced by cyanobacterial blooms would gain at investigating other fish species feeding directly on phytoplankton organisms. It should highlight whether higher metabolome and toxicological changes would be driven by directly feeding on cyanobacterial cells and higher contact with with the various potentially toxic secondary metabolites, or by other ecological parameters impacted by cyanobacterial proliferations. Furthermore, complementary omic analyses, such as

proteomics and transcriptomics, would provide additional support to generate system biology understanding of the global organism response to environmental stress, and help at characterizing potential biomarkers useable at the fish population levels in various aquatic environments.

**Conflicts of interest**

None

**Declaration of interest**

None

## Acknowledgments

The NMR and the MS spectra were respectively acquired at the Plateau technique de Résonance Magnétique Nucléaire and the Plateau technique de spectrométrie de masse bio-organique, UMR 7245 CNRS/MNHN Molécules de Communication et d'Adaptation des Micro-organismes, Muséum National d'Histoire Naturelle, Paris, France. This work was supported by grants from the Sorbonne Universités "DANCE", "Procytox" projects and from CNRS (Défi ENVIROMICS "Toxcyfish" project) and from the "region Île-de-France" (R2DS N°2015-11) awarded to Dr. Benjamin Marie. The authors would like to thank the French minister for the research for the financial supports to Séverine Le Manach. Qin Qiao PhD was founded by the China Scolarship Council.

## References

Amado, L., Monserrat, J., 2010. Oxidative stress generation by microcystins in aquatic animals: Why and how. *Environment International* 36, 226–235.

Azimi, S., Rocher, V., 2016 Influence of the water quality improvement on fish population in the Seine River (Paris , France) over the 1990 – 2013 period. *Science of the Total Environment* 542, 955–964.

Bols, N. C., Brubacher, J. L., Ganassin, R. C., Lee, L. E. J., 2001. Ecotoxicology and innate immunity in fish. *Developmental and Comparative Immunology* 25(8–9), 853–873.

Bundy, J. G., Davey, M. P., Viant, M. R., 2008. Environmental metabolomics: a critical review and future perspectives. *Metabolomics* 5(1), 3–21.

Bunn, S. E., 2016. Grand challenge for the future of freshwater ecosystems. *Frontiers in*

509 *Environmental Science 4*, 1–4.
510 Cappello, T., Fernandes, D., Maisano, M., Casano, A., Bonastre, M., Bebianno, M. J., ... &
511 Porte, C., 2017. Sex steroids and metabolic responses in mussels *Mytilus*
512 *galloprovincialis* exposed to drospirenone. *Ecotoxicology and environmental safety*
513 143, 166-172
514 Cappello, T., Brandão, F., Guilherme, S., Santos, M. A., Maisano, M., Mauceri, A., Canário,
515 J., Pacheco, M., Pereira, P., 2016. Insights into the mechanisms underlying mercury-
516 induced oxidative stress in gills of wild fish (*Liza aurata*) combining $^{1}$H NMR
517 metabolomics and conventional biochemical assays. *Science of the Total Environment*
518 *548–549*, 13–24.
519 Catherine, A., Troussellier, M., Bernard, C., 2008. Design and application of a stratified
520 sampling strategy to study the regional distribution of cyanobacteria (Ile-de-France,
521 France). *Water research* 42(20), 4989–5001.
522 Catherine, A., Mouillot, D., Escoffier, N., Bernard, C., Troussellier, M., 2010. Cost effective
523 prediction of the eutrophication status of lakes and reservoirs. *Freshwater Biology*
524 55(11), 2425–2435.
525 Catherine, A., Selma, M., Mouillot, D., Troussellier, M., Bernard, C., 2016. Patterns and
526 multi-scale drivers of phytoplankton species richness in temperate peri-urban lakes.
527 *Science of The Total Environment* 559, 74–83.
528 Catherine, A., Bernard, C., Spoof, L., Bruno, M., 2017. Microcystins and Noduarins. In
529 *Handbook on Cyanobacterial Monitoring and Cyanotoxin Analysis*, Meriluoto, J.,
530 Spoof, L., Codd, G. A., Eds., John Wiley and Sons, p 576.
531 Chorus, I., Bartram, J. *Toxic cyanobacteria in water: a guide to their public health*
532 *consequences, monitoring and management.*, E & FN Spon, 1999.
533 Codd, G. A., Morrison, L. F., Metcalf, J. S. 2005. Cyanobacterial toxins: risk management for
534 health protection. *Toxicol. Appl. Pharm.* 203(3), 264–272.
535 Codd, G. A., Meriluoto, J., Metcalf, J. S., 2017. Introduction: Cyanobacteria, Cyanotoxins,
536 Their Human Impact, and Risk Management. In *Handbook of Cyanobacterial*
537 *Monitoring and Cyanotoxin Analysis*, Meriluoto, J., Spoof, L., Codd, G. A., Eds., John
538 Wiley and Sons, p 548.
539 Combes, A., El Abdellaoui, S., Vial, J., Lagrange, E., Pichon, V., 2014. Development of an
540 analytical procedure for quantifying the underivatized neurotoxin β-N-methylamino-l-
541 alanine in brain tissues. *Analytical and Bioanalytical Chemistry* 406(19), 4627–4636.
542 Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z.-I., Knowler, D. J., Lévêque,

C., Naiman, R. J., Prieur-Richard, A.-H., Soto, D., Stiassny, M. L. J., et al., 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews of the Cambridge Philosophical Society* 81 (2), 163–182.

Faassen, E. J., Antoniou, M. G., Beekman-Lukassen, W., Blahova, L., Chernova, E., Christophoridis, C., Combes, A., Edwards, C., Fastner, J., Harmsen, J., et al., 2016. A collaborative evaluation of LC-MS/MS based methods for BMAA analysis: Soluble bound BMAA found to be an important fraction. *Marine Drugs* 14(3).

Ferrão-Filho, A. da S., Kozlowsky-Suzuki, B., 2011. Cyanotoxins: bioaccumulation and effects on aquatic animals. *Marine drugs* 9(12), 2729–2772.

Fiehn, O., 2002. Metabolomics - the link between genotyopes and phenotypes. *Plant Molecular Biology* 48, 155–171.

Franzosa, E. A., Hsu, T., Sirota-Madi, A., Shafquat, A., Abu-Ali, G., Morgan, X. C., Huttenhower, C., 2015. Sequencing and beyond: integrating molecular "omics" for microbial community profiling. *Nature reviews. Microbiology* 13(6), 360–372.

Gaspéri, J., Moreau-Guigon, E., Labadie, P., Blanchard, M., Teil, M.-J., Tlili, K., Dinh, T. ., Tran, C., Alliot, F., Desportes, A., et al., 2010 Contamination de la Seine par les micropolluants organiques : Effet des conditions hydriques et de l'urbanisation. https://hal-enpc.archives-ouvertes.fr/hal-00766537

Hamilton, P. B., Cowx, I. G., Oleksiak, M. F., Griffiths, A. M., Grahn, M., Stevens, J. R., Carvalho, G. R., Nicol, E., Tyler, C. R., 2016. Population-level consequences for wild fish exposed to sublethal concentrations of chemicals ??? a critical review. *Fish and Fisheries* 17(3), 545–566.

Hao, J., Liebeke, M., Astle, W., De Iorio, M., Bundy, J. G., Ebbels, T. M. D., 2014. Bayesian deconvolution and quantification of metabolites in complex 1D NMR spectra using BATMAN. *Nature protocols* 9(6), 1416–1427.

Hultman, J., Waldrop, M. P., Mackelprang, R., David, M. M., McFarland, J., Blazewicz, S. J., Harden, J., Turetsky, M. R., McGuire, A. D., Shah, M. B., et al., 2015. Multi-omics of permafrost, active layer and thermokarst bog soil microbiomes. *Nature 521*, 208–212.

Krishnamoorthy, K., & Lu, F., 2010. A parametric bootstrap solution to the MANOVA under heteroscedasticity. *Journal of Statistical Computation and Simulation* 80(8), 873-887.

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.

577 *Environmental pollution, 234*, 523-537.
578 Lin, C. Y., Wu, H., Tjeerdema, R. S., Viant, M. R., 2007. Evaluation of metabolite extraction
579 strategies from tissue samples using NMR metabolomics. *Metabolomics* 3(1),
580 55–67.
581 Lopez-Archilla, A., Moreira, D., Lopez-Garcia, P., Guerrero, C., 2004. Phytoplankton
582 diversity and cyanobacterial dominance in a hypereutrophic shallow lake with
583 biologically produced alkaline pH. *Extremophiles* 8, 109–115.
584 Malbrouck, C., Kestemont, P., 2006. Effects of microcystins on fish. *Environ. Toxicol. Chem.*
585 25(1), 72–86.
586 Maloufi, S., Catherine, A., Mouillot, D., Louvard, C., Couté, A., Bernard, C., Troussellier,
587 M., 2016. Environmental heterogeneity among lakes promotes hyper ß-diversity across
588 phytoplankton communities. *Freshwater Biology* 61(5), 633–645.
589 Marie, B., Huet, H., Marie, A., Djediat, C., Puiseux-Dao, S., Catherine, A., Trinchet, I.,
590 Edery, M., 2012. Effects of a toxic cyanobacterial bloom (*Planktothrix agardhii*) on
591 fish: Insights from histopathological and quantitative proteomic assessments following
592 the oral exposure of medaka fish (*Oryzias latipes*). *Aquatic toxicology (Amsterdam,*
593 *Netherlands)* 114–115C, 39–48.
594 McLean, T. I., 2013. "Eco-omics": A Review of the Application of Genomics,
595 Transcriptomics, and Proteomics for the Study of the Ecology of Harmful Algae.
596 *Microbial Ecology* 65(4), 901–915.
597 Paerl, H. W., Paul, V. J., 2011. Climate Change: Links to Global Expansion of Harmful
598 Cyanobacteria. *Water Research* 46(5), 1363–1349.
599 O'Neil, J. M., Davis, T. W., Burford, M. A., Gobler, C. J., 2012. The rise of harmful
600 cyanobacteria blooms: The potential roles of eutrophication and climate change.
601 *Harmful Algae* 14, 313–334.
602 Qiao, Q., Le Manach, S., Huet, H., Duvernois-Berthet, E., Chaouch, S., Duval, C., Sotton, B.,
603 Ponger, L., Marie, A., Mathéron, L., et al., 2016. An integrated omic analysis of
604 hepatic alteration in medaka fish chronically exposed to cyanotoxins with possible
605 mechanisms of reproductive toxicity. *Environ. pollut.* 219, 119–131.
606 Revenga, C., Campbell, I., Abell, R., de Villiers, P., Bryer, M., 2005. Prospects for
607 monitoring freshwater ecosystems towards the 2010 targets. *Philosophical*
608 *transactions of the Royal Society of London. Series B, Biological sciences* 360(1454),
609 397–413.
610 Rohart, F., Gautier, B., Singh, A., & Le Cao, K. A., 2017. mixOmics: An R package for

611 ‘omics feature selection and multiple data integration. *PLoS computational biology*
612 13(11), e1005752.
613 Samuelsson, L. M., Larsson, D. G. J., 2008. Contributions from metabolomics to fish
614 research. *Molecular bioSystems* 4(10), 974–979.
615 Sardans, J., Peguelas, J., Rivas-Ubach, A., 2011 Ecological metabolomics: Overview of
616 current developments and future challenges. *Chemoecology* 21(4), 191–225.
617 Schwarzenbach, R. P., Egli, T., Hofstetter, T. B., von Gunten, U., Wehrli, B., 2010. Global
618 Water Pollution and Human Health. *Annual Review of Environment and Resources*
619 35(1), 109–136.
620 Sotton, B., Guillard, J., Anneville, O., Maréchal, M., Savichtcheva, O., Domaizon, I., 2014.
621 Trophic transfer of microcystins through the lake pelagic food web: Evidence for the
622 role of zooplankton as a vector in fish contamination. *Science of The Total*
623 *Environment* 466–467, 152–163.
624 Sotton, B., Domaizon, I., Anneville, O., Cattanéo, F., Guillard, J., 2015. Nodularin and
625 cylindrospermopsin: a review of their effects on fish. *Reviews in Fish Biology and*
626 *Fisheries* 25, 1–19.
627 Sotton, B., Paris, A., Le Manach, S., Blond, A., Lacroix, G., Millot, A., ... & Marie, B.,
628 2017a. Global metabolome changes induced by cyanobacterial blooms in three
629 representative fish species. *Science of the Total Environment* 590, 333-342.
630 Sotton, B., Paris, A., Manach, S. Le, Blond, A., Lacroix, G., Duval, C., Huet, H., Qiao, Q.,
631 Labrut, S., Chiapetta, G., et al., 2017b. Metabolic changes in Medaka fish induced by
632 cyanobacterial exposures in mesocosms : an integrative approach combining
633 proteomic and metabolomic analyses. *Scientific reports* 7, 1–13.
634 Teil, M.-J., Tlili, K., Blanchard, M., Labadie, P., Alliot, F., Chevreuil, M., 2014.
635 Polychlorinated Biphenyls, Polybrominated Diphenyl Ethers, and Phthalates in Roach
636 from the Seine River Basin (France): Impact of Densely Urbanized Areas. *Archives of*
637 *environmental contamination and toxicology* 66, 41–57.
638 Tenenhaus, A., Philippe, C., Guillemot, V., Le Cao, K. A., Grill, J., & Frouin, V., 2014.
639 Variable selection for generalized canonical correlation analysis. *Biostatistics* 15(3),
640 569-583.
641 Wang, Mingxun, et al., 2016. Sharing and community curation of mass spectrometry data with
642 Global Natural Products Social Molecular Networking. *Nature Biotechnology* 34/8,
643 828-837.
644 Wu, H., Southam, A. D., Hines, A., Viant, M. R., 2008. High-throughput tissue extraction

protocol for NMR- and MS-based metabolomics. *Analytical biochemistry* 372 (2), 204–212.

Yepremian, C., Catherine, A., Bernard, C., 2017. SOP202: Chlorophyll a extraction. In *Handbook on Cyanobacterial Monitoring and Cyanotoxin Analysis*, Meriluoto, J., Spoof, L., Codd, G. ., Eds., p 458.

# Table

| Putative metabolites | BBECya | | pH | |
|---|---|---|---|---|
| | *F-value* | *p-value* | *F-value* | *p-value* |
| **16a-Hydroxyestrone** | 26.604 | **$1.02e^{-06}$** | 8.295 | **0.00472** |
| **Acetylcholine** | 43.903 | **$1.08e^{-09}$** | 5.892 | **0.01672** |
| **Androsterone** | 19.932 | **$1.85e^{-05}$** | 10.776 | **0.00135** |
| **Glycerophosphocholine** | 38.040 | **$1e^{-08}$** | 8.892 | **0.00348** |
| **Glycerol** | 29.455 | **$3.1e^{-07}$** | 3.908 | 0.0504 |
| **Homo-L-arginine** | 32.501 | **$8.95e^{-08}$** | 1.795 | 0.18287 |
| **Isovalerylcarnitine** | 32.028 | **$1.08e^{-07}$** | 9.392 | **0.0027** |
| **L-Arginine** | 30.269 | **$2.22e^{-07}$** | 2.228 | 0.13820 |
| **L-Glutamine** | 30.004 | **$2.47e^{-07}$** | 14.025 | **0.000280** |
| **L-Palmitoylcarnitine** | 24.965 | **$2.05e^{-06}$** | 12.888 | **0.000483** |

**Table 1**: Effects of significantly-related environmental factors on the putative metabolite expressions revealed by two-ways ANOVA analyses performed on the environmental parameter and metabolite matrix for all individuals (supplementary data 1). Bold p-values correspond to p-values < 0.05.

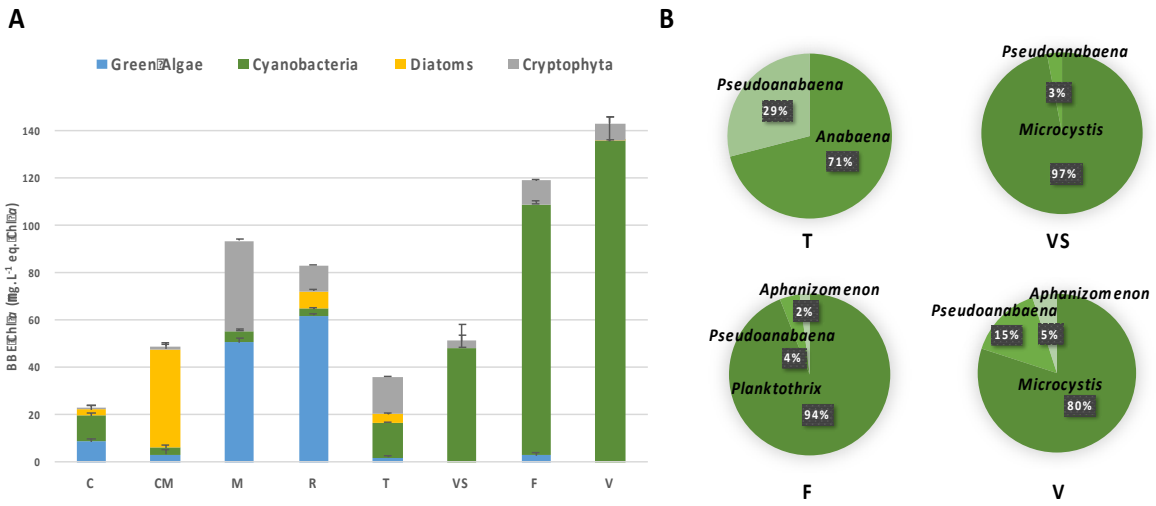
## Legends of figures

**Figure 1. Phytoplanktonic chlorophyll–a concentrations (µg.L$^{-1}$ eq. Chl-a) in the sampled lakes (A) and the cyanobacterial genus composition observed in the cyanobacteria-dominated lakes (B).** ND = non-detected. Cergy (C), Champs-sur-Marne (CM), Maurepas lake (M), Rueil lake (R), Verneuil lake (V), Varenne-sur-Seine lake (VS), Fontenay lake (F) and Triel lake (T).

**Figure 2. Diversity of metabolites observed in the lakes, dominated or not by cyanobacterial proliferations revealed by molecular networking based on MS/MS fragmentation data.** When various phospholipids are observed in all lakes, potentially bio-active metabolites known to be produced by certain cyanobacteria, such as microcystins (hepatotoxin), aeruginosins, cyanopeptolins, anabaenopeptins, microginins, are only observed in cyanobacterial dominated lakes.

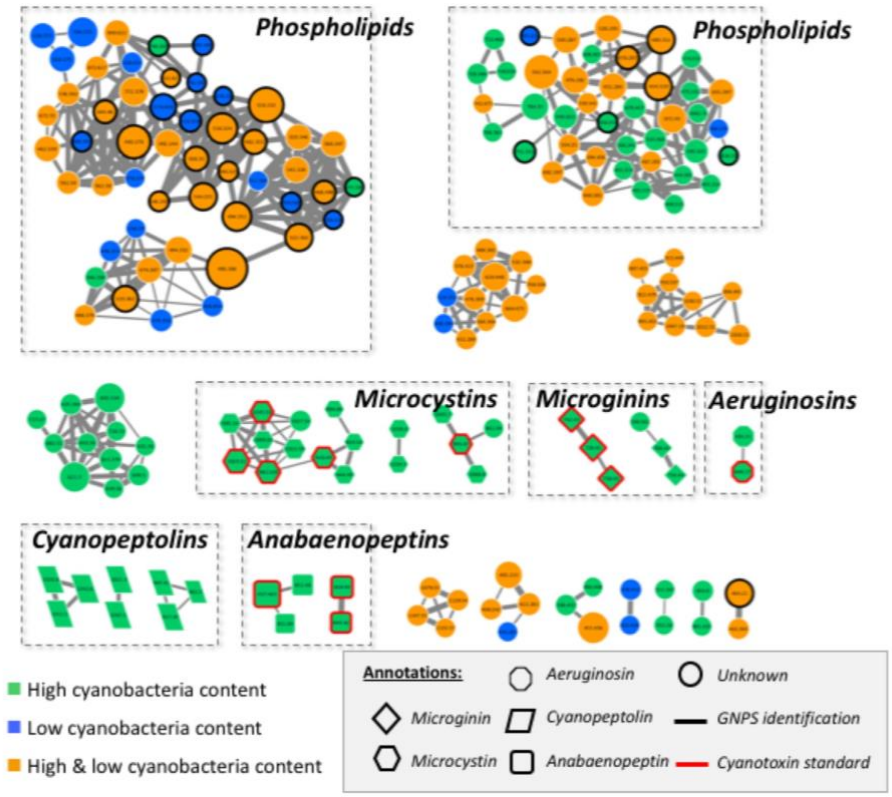
**Figure 3.** $^{1}$H-NMR liver metabolomes and relevance network of fish sampled in the different lakes. The individual plots of regularized canonical correlations analysis (rCCA) for dimensions 1–2 (A). Perch individuals are on the left side and pumpkinseed individuals on the right side of the graphic. Lakes are represented by their respective letters that in blue correspond to control lakes and in green to perturbated lakes. Cergy (C), Champs-sur-Marne (CM), Maurepas lake (M), Rueil lake (R), Verneuil lake (V), Varenne-sur-Seine lake (VS), Fontenay lake (F) and Triel lake (T). Relevance network providing from rCCA analysis on the dimension 2 (B). Putative metabolites above a correlation threshold of |0.3| were kept, revealing that the 2 most correlated variables are pH and BBEcya (cyanobacterial specific content of chlorophyll a measured by BBE flurorescence probe). Green edges correspond to negative correlations with the discriminant ecological factors.

686 **Fig. 1.**


A
Green Algae
Cyanobacteria
Diatoms
Cryptophyta
140
120
100
80
60
40
20
0
C
CM
M
R
T
VS
F
V
B
Pseudoanabaena
29%
Anabaena
71%
T
Pseudoanabaena
3%
Microcystis
97%
VS
Aphanizomenon
2%
Pseudoanabaena
4%
Planktothrix
94%
F
Aphanizomenon
Pseudoanabaena
15%
5%
Microcystis
80%
V

687

688 **Fig. 2.**

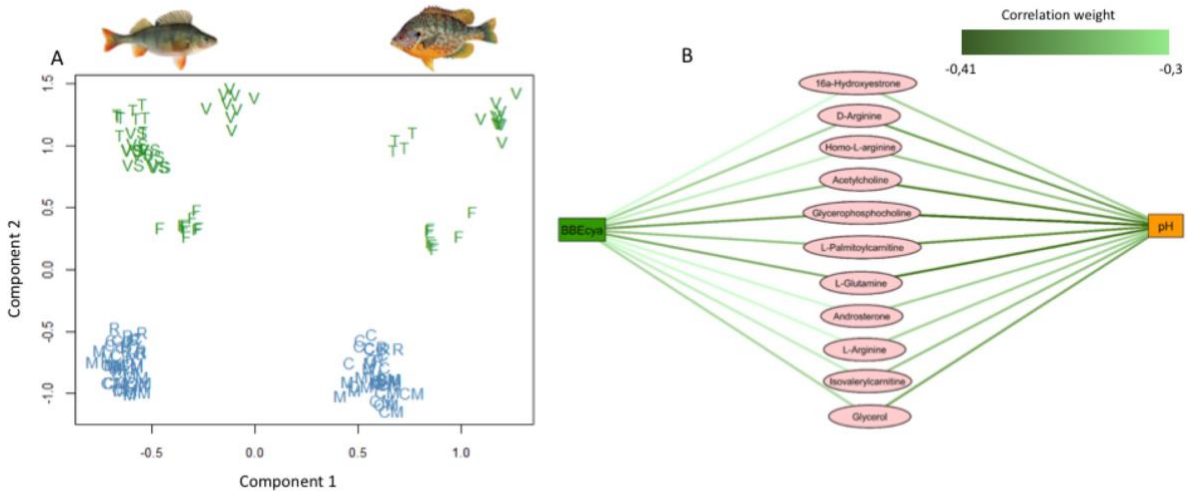



689

690

691

692 **Fig. 3.**


A
Component 2
Component 1
-0.5
0.0
0.5
1.0
1.5
-1.0
B
Correlation weight
-0,41
-0,3
16a-Hydroxyestrone
D-Arginine
Homo-L-arginine
Acetylcholine
Glycerophosphocholine
L-Palmitoylcarnitine
L-Glutamine
Androsterone
L-Arginine
Isovalerylcarnitine
Glycerol
BBEcya
pH

693