Limnospira (Cyanobacteria) chemical fingerprint reveals local molecular adaptation

Théotime Roussel[1,3], Cédric Hubas[2], Sébastien Halary[1], Mathias Chynel[2], Charlotte Duval[1], Jean-Paul Cadoret[3], Tarik Meziane[2], Léa Vernès[3], Claude Yéprémian[1], Cécile Bernard[1*], Benjamin Marie[1*]

[1] UMR7245 MCAM MNHN-CNRS, Muséum National d'Histoire Naturelle, CP 39, 12 rue Buffon, F-75231 Paris Cedex 05, France

[2] UMR 7208 BOREA MNHN-CNRS-SU-IRD, Muséum d'Histoire Naturelle, CP 53, 61 rue Buffon, F-75231 Paris Cedex 05, France

[3] Algama, 81 rue Réaumur, 75002 PARIS France

* Correspondance: benjamin.marie@mnhn.fr; cecile.bernard@mnhn.fr (+33140793212, +33140793183, BM; CB)

Marie B., https://orcid.org/0000-0001-9880-5541

Bernard C., http://orcid.org/0000-0001-7032-3989

Abstract

Limnospira can colonize a wide variety of environments (*e.g.*, freshwater, brackish, alkaline or alkaline-saline water) and develop dominant and even permanent blooms that limit overshadowed adjacent phototrophs diversity, especially in alkaline and saline environments. Previous phylogenomic analysis of *Limnospira* allowed us to distinguish two major phylogenetic clades (I and II) but failed to clearly segregate strains according to their respective habitats in terms of salinity or biogeography. In the present work, we attempt to determine whether *Limnospira* displays metabolic signatures specific to its different habitats, particularly brackish or alkaline-saline ecosystems, and question the impact of accessory gene repertoires on respective chemical adaptations.

The study of the metabolomic diversity of 93 strains of *Limnospira* from the Paris Museum Collection, grown under standardized lab culture conditions, showed clearly distinct chemical fingerprints that were correlated with the respective biogeographic origins of the strains. The molecules that most distinguish the different *Limnospira* geographic groups are sugars, lipids, peptides, photosynthetic pigments, and antioxidant molecules. Interestingly, these molecule enrichments might represent adaptation traits to the local conditions encountered in their respective sampling environments concerning salinity, light and oxidative stress. We hypothesize that within extreme environments, such as those colonized by *Limnospira*, a large set of flexible genes can provide remarkable adaptation to specific local environmental conditions (*e.g.*, salinity, light, and oxidative pressure). Thus, the occurrence within *Limnospira* population genomes of a specific set of flexible genes potentially involved in the production of certain metabolites may provide valuable adaptative traits that may support the bloom persistence beyond environmental condition variations.

Importance

Limnospira are ubiquitous cyanobacteria able to colonize and dominate a wide range of alkaline-saline environments around the world according to remarkable adaptative strategies. Phylogenomic analysis of *Limnospira platensis* allowed to distinguish two major phylogenetic clades (I and II) but failed to clearly segregate strains following their habitats in terms of salinity or biogeography. One can presume that the genes found within this variable portion of the genome of these clades could be involved in *L. platensis* adaptation to local environmental conditions. In the present paper, we attempt to determine whether *Limnospira platensis* displays metabolic signatures specific to its different habitats, particularly brackish or alkaline-saline ecosystems, and question the impact of accessory gene repertoire on respective chemical adaptation.

Keywords: Cyanobacteria, metabolomics, genomics, adaptation, micro-diversity.

Introduction

Certain phytoplanktonic organisms can durably colonize their environment. This capability appears to be balanced by their specific ability to perform efficient primary production, to avoid being overshadowed by competitors, and to deal with physico-chemical environmental constraints. This adaptive power is strongly dependent on the energetic cost they must pay for it, in terms of genomic or physiological expenses[1].

Cyanobacteria are a group of photoautotrophic microorganisms that have colonized a wide range of environments around the world[2-3]. The large diversity of cyanobacteria species is accompanied by a great diversity of primary and/or secondary/accessory metabolites, many of which presenting noticeable biological activities of central interest in key cellular or ecological functions[4-5]. The genus *Limnospira* (anc. *Arthrospira*), better known under its commercial/vernacular name 'spirulina', is one of the most extensively cyanobacteria taxa studied so far (*e.g.* more than 300 publications per year since 2020 on PubMed), especially due to its industrial potential (*e.g.* agri-food, pharmaceuticals, cosmetics)[6-7]. Indeed, cultivated *Limnospira* present a remarkably high growth rate under appropriate culture conditions (more than one division per day)[8-9], and are very rich in proteins (up to 70% of the dry weight)[10], unsaturated fatty acids[11], vitamins (at least 9)[12-13], and various molecules with antioxidant properties (*e.g.* polyphenols, carotenoids, ergothioneine, and biopterins)[12,14]. To this aim, the physiological aptitudes of certain *Limnospira* strains have been examined (*i.e.*, biomass or molecule production affected by modifying culture parameters)[8,15-18], whereas only few studies have focused on the global ecological trait of this taxon[19].

In a recent study based on a polyphasic approach combined with comparative genomics, we showed that the *Limnospira* genus is monospecific, represented by one single species, *Limnospira platensis*, which is cosmopolitan[20]. Yet, *Limnospira platensis* seems to develop a

singular ecological strategy, as this species is able to colonize a wide variety of environments (*e.g.*, freshwater, brackish, alkaline or alkaline-saline water)[19-20] in which it dominates the phytoplanktonic community by developing permanent blooms that limit overshadowed adjacent phototroph diversity[19,21].

Phylogenomic analysis of *L. platensis* allowed us to distinguish two major phylogenetic clades (I and II), but failed to clearly segregate strains following their habitats in terms of salinity or biogeography[20]. Such discrepancies have already been observed in other taxa, for instance for the widespread extremophilic bacteria *Salinibacter ruber*[22]. However, in this case, non-targeted metabolites analysis succeeded to segregate the strains, highlighting biogeographical metabolic signatures related to local adaptations. Recently, the same approach has suggested the existence of distinct chemical signatures between different ecotypes of the oceanic picocyanobacterium *Prochlorococcus marinus*, each being adapted to a specific niche characterized in terms of light intensity and temperature[23].

Concerning the species *L. platensis*, the flexible part of the genome remarkably represents up to 55% of the genome of each strain, and differs from one strain to another, carries a large variety of genes (certain of them presenting regulatory and defensive functions)[20], many of which present so-far unknown functions and could potentially be related to the specific capabilities of the strains in terms of physiology and ecology. One can presume that the genes found within this variable portion of the genome could be involved in *L. platensis* adaptation to local environmental conditions. These genes could indeed be responsible of carrying traits that allow the species to adjust to specific local conditions and support the maintenance of constant blooms. Ultimately, this local adaptation might give *L. platensis* an advantage to overpass other phytoplankton competitors within these environments.

In the present paper, we attempt to determine whether *Limnospira platensis* displays metabolic signatures specific to its different habitats, particularly brackish or alkaline-saline ecosystems, and question the impact of accessory gene repertoires on respective chemical adaptations. In this aim, we have investigated a wide panel of strains from different environments and maintained them in collections under identical culture conditions. We used two types of extraction and chemical analysis approaches, pursuing a large diversity of metabolites potentially involved in various adaptative traits. First, we applied a targeted approach, focusing on photosynthetic pigments and fatty acids known to be involved in adaptation to different environmental conditions (*e.g.*, light, salinity, and pH). Then, we performed an untargeted approach in parallel to obtain an overall chemical fingerprint of each strain and look for other potential differences.

Material and Methods

Strains and culture conditions

93 strains of *L. platensis* were studied. 89 strains were obtained from the Paris Museum Collection (PMC)[24]: 72 from two Camargue area, a humid region located in the Rhône delta (South of France) mostly consisting of marshland and brine lagoons, 13 from the lake Dziani Dzaha, an alkaline and thalassohaline lake located on the island Petite-Terre of Mayotte, three from lake Natron, an alkaline-saline lake located between Tanzania and Kenya, and one strain from historical collection with no environmental data available. Three additional strains were provided by other collections: *L. platensis* PCC 8005 and *L. platensis* PCC 7345 (Pasteur Culture Collection of Cyanobacteria, France) from unknown locations, and *L. platensis* SAG 85.79 (Culture Collection of Algae at Göttingen University, Germany) from the lake Natron. one strain provided by Spirulina Solutions (https://spirulinasolutions.fr/): Paracas R14 from a freshwater lake in Peru. All information about strains, sampling site properties, year of

isolation, and genomic information is given in Supplementary Table S1. These 93 strains were maintained as monospecific but non-axenic cultures, growing in 30 mL of sensibly improved Spirulina medium[24] at 24°C with a photon flux density of 30 µmol photon.m^{-2}.s^{-1} under a 16:8-h light:dark photoperiod. When investigating the metabolite production of the various strains, careful attention was paid to maintaining cultures in the growing phase to keep the level of heterotrophic bacteria as low as possible. We assume it remains negligible comparing to cyanobacteria development, especially in terms of biovolume, biomass, and metabolite contents, and might not critically influence the present observation on *Limnospira* strain distinctions. Each month, 1/10 of each culture were used to inoculated new cultures. The remaining 9/10 were centrifuged (10 min, 3,220 g, 3 times). The supernatants were discarded, and the pellets were stored at -80°C for at least 3h, and then lyophilization was performed (-47°C, 0.03 mBar for 16h) (Labconco, Kansas City, MO, USA). Six successive cycles of culture-biomass sampling have been made and pooled. The six freeze-dried biomass samples of each strain have been mixed together and homogenized prior to the different chemical analyses for both pigments, fatty acids and untargeted metabolites.

Pigments

Hydrophilic pigments of the 93 strains were analyzed by spectrophotometry according to Yéprémian et al.[25-26]. Ten mg of freeze-dried biomass were incubated with 10 mL of a saline buffer solution (NaCl 0.15 mol.L^{-1}, KCl 27 mmol.L^{-1}, Na_2HPO_4 80 mmol.L^{-1}, KH_2PO_4 20 mmol.L^{-1}, and Na_2-EDTA 10 mmol.L^{-1}; pH = 7.5) for 16h, at 4°C in the dark. Samples were then centrifuged (10 min, 3,220 g), and 1 mL of the supernatant was analyzed with a spectrophotometer (CARY-60, Agilent) at 565, 620, and 650 nm, corrected by subtracting the value recorded at 750 nm (< 0.01 absorbance unit). The values reported as concentration

(mg.g^{-1}) of C-phycocyanin and allophycocyanin (absence of phycoerythrin) were obtained with calculus according to de Marsac and Houmard[27].

Lipophilic pigments of the 93 strains were analyzed by high performance liquid chromatography (HPLC). The complete procedure is detailed in supplementary material and methods. The relative abundance of each pigment (%) was calculated from its respective concentration in the sample (mg.g^{-1} dw for chlorophyll *a* derivative, and μg.g^{-1} dw for other pigments).

Fatty acids

Ten mg of freeze-dried biomass were used for the extraction of fatty acids from 91 strains (all 93 except *L. platensis* PMC 1041.18 and PMC 1044.18). Prior to extraction, an internal standard (tricosanoic acid: 23:0) was added in each sample. Lipids were extracted according to the protocol of Bligh and Dyer[29] as modified by Meziane et al.[30]. The complete procedure is detailed in supplementary material and methods. Fatty acids were identified with a mass spectrometer (Agilent 5977B GC/MSD) according to the comparison of retention times of commercial fatty acid standards (Supelco 37). We then reported the values as % of total FA and concentration (mg.g^{-1} dw).

Mass spectrometry

The intracellular metabolite contents of 88 strains (all 93 except *L. platensis* PMC 1287.21, PMC 1288.21, PMC 1298.21, PMC 1300.21, and PMC 1302.21) were analyzed as previously described[31]. The complete procedure is detailed in supplementary material and methods. Intracellular metabolite contents were analyzed using an electrospray ionization hybrid quadrupole time-of-flight (ESI-QqTOF) high-resolution mass spectrometer (Compact,

Brucker, Bremen, Germany) in the range 50–1500 *m/z*. Metabolite annotations were made from MS2 data by generating a molecular network for the comparison of fragmentation profiles using the MetGem software (version 1.3.6) with GNPS algorithm (supplementary figure S1).

Statistical analyses

The MetaboAnalyst 5.0 platform (www.metaboanalyst.ca) was used to perform data matrix normalization according to *Pareto* and mean-centered and divided by the standard deviation for untargeted metabolite, and for fatty acid and pigment matrices, respectively.

All three matrices (annotated ions, fatty acids, and pigments) were combined in a single matrix further called "total metabolite matrix". The weight of each group of variables was adjusted to avoid the analysis being influenced by the largest group. The weights are identical for the variables of the same group.

Clade based on presence/absence of coding gene sequences was obtained with dendrogram (Jaccard's dissimilarity index, Ward reconstruction, Silouhette analysis to find the optimal number of clusters) using software R (version 1.4.1103).

Hierarchical clustering on principal components (HCPC) using Euclidian distance and a principal component analysis (PCA) were performed to find different groups of strain using software R (version 1.4.1103). An analysis of similarities (ANOSIM) was performed on the total metabolite matrix to test differences between factors. Multiple correspondence analysis (MCA) was performed to identify correlations between tested factors using software R (version 1.4.1103). For each tested factor, multiple factor analyses (MFA) and between class analyses (BCA) were performed simultaneously (BC-MFA, code available at https://github.com/Hubas-prog/BC-MFA) on the combined matrix, as describe in Michelet et al.[32]. The percentage of total inertia explained (TIE) by the instrumental variable was

systematically calculated. A threshold based on the square cosine of the coordinates of the variables was applied for highlighting the most discriminating variables with a square cosine superior to 0.4 ($cos^2 > 0.4$). A Monte-Carlo test was systematically performed to test the significance of the BCA ordination, using software R (version 1.4.1103). For each factor, principal component analyses (PCA) and heatmap analyses were performed on each of the three matrices using the MetaboAnalyst 5.0 platform.

Differences between each class of factor have been searched for each of the 215 molecules with Kruskal-Wallis tests followed by pairwise Wilcoxon tests. For total fatty acid means comparison, normality was present (Shapiro-Wilk test, $p > 0.05$) and homogeneity of the variances was not present (Bartlett test, $p < 0.05$), thus a Welch test was performed. All tests were run using software R (version 1.4.1103).

Results

Pigment composition

Fourty photosynthetic pigments were detected and quantified by HPLC-DAD in the 93 analyzed *L. platensis* strains: two phycobiliproteins (C-phycocyanin and allophycocyanin), 12 chlorophylls or derivatives (chlorophylls *a*, *d*, and pheophytins), and 26 carotenoids (five carotenes, four carotenoid glycosides, six xanthophylls, one unknown carotene, and ten unknown carotenoids) (Figure 1A; Supplementary Table S2). A mean of 27.9 ± 1.8 pigments was detected per strain, ranging from 20 to 32 for strain *L. platensis* PMC 1223.20, PMC 1278.20 and PMC 1279.20, and *L. platensis* PMC 1041.18, respectively (Supplementary Table 2). 12 pigments were detected in all strains: C-phycocyanin, allophycocyanin, chlorophyll *a*, chlorophyll *a* allomer, chlorophyll *a* epimer, two pheophytins *a*, β,φ-carotene, β,ε-carotene-like, zeaxanthin, echinenone, echinenone/cantaxanthin-like, and two unknown carotenoids (UC10, and UC9) (Supplementary Table S3). The most abundant pigments were

C-phycocyanin (64.5 ± 3.1% of total pigment composition) and allophycocyanin (25 ± 1.6% of total pigment composition). The chlorophyll *a* allomer was the most abundant form of chlorophyll *a* in 90 out of 93 strains, representing 6.6 ± 2.3% of the total pigment composition. The three other strains (*L. platensis* PMC 1223.20, PMC 1278.20, and PMC 1279.20) had the usual form of chlorophyll *a* as the main chlorophyll *a*, representing 13.6 ± 2.3% of the total pigment composition (Supplementary Table 2). Carotenoids represented 1.19 ± 0.6% of the total pigment composition. The most abundant carotenoids were zeaxanthin (34.2 ± 5.5% of the total carotenoid composition), myxoxanthophyll (14.9 ± 5.6% of total carotenoid composition), β,β-carotene (9.9 ± 1.3% of total carotenoid composition), and UC9 (425, 448, 470nm) (8.7 ± 3.6% of total carotenoid composition) (Supplementary Table S2).

Fatty acid composition

Fourteen fatty acids were detected and quantified by GC-FID and GC-MS in the 91 analyzed *L. platensis* strains. A mean of 52.32 ± 23.77 mg fatty acids.g^{-1} dry weight (dw) was quantified, ranging from 13.78 to 110.59 mg.g^{-1} for the strains *L. platensis* PMC 1259.20 and *L. platensis* PMC 894.15, respectively (Supplementary Table S3). Among the 14 fatty acids detected, three were saturated fatty acids (14:0, 16:0, and 18:0), two were branched fatty acids (15:0iso, and 15:0anteiso), four were mono-unsaturated fatty acids (16:1ω7, 16:1ω9, 18:1ω7, and 18:1ω9), and five were poly-unsaturated fatty acids (16:2ω6, 18:2ω6, 18:3ω6, 20:2ω6, and 20:3ω6). Five fatty acids represented > 90% of the total fatty acids: 16:0 (38.6 ± 1.8%), 18:3ω6 (24.6 ± 2.1%), 18:2ω6 (17.1 ± 1.5%), 16:1ω9 (6.2 ± 1.1%), and 16:1ω7 (5.6 ± 0.8%) (Supplementary Table S3). Interestingly, the relative quantities of these molecules within the different strains allow to most distinguish a group of “lower” and “higher” fatty acid-

containing strains according to heatmap with hierarchical clustering with no obvious relation with the respective metadata of the strains (Figure 1B).

Metabolite composition by LC-MS

A total of 2,851 ions were detected by HPLC-MS in the 88 analyzed *L. platensis* strains. A mean of 1,311.7 ± 175.1 ions was detected, ranging from 1,023 to 1,686 for strain *L. platensis* PMC 1044.18 and *L. platensis* SAG 85.79, respectively (Supplementary Table S4). 1013 ions (35.5%) exhibit an intensity superior to a threshold of 5,000 counts in single MS and were analyzed by HPLC-MS/MS. Among them, 221 ions (15.9% of fragmented ion, and 5.6% of total ions) have been clustered and annotated with MetGem and GNPS (*e.g.* phospholipids, saccharides, peptides, …) and 25 have been annotated at the molecular levels (*e.g.* ergothioneine, glutathione, biopterin, nicotinamide adenine dinucleotide) (Figure 1C; Supplementary Table S4), whereas other were annotated according to MS/MS fragmentation pattern similarity revealed by molecular networking (Supplementary Figure S1). Analyses of the 2,851 total ions and the 221 annotated ions showed similar results (Data not shown). We further focused only of the 221 annotated ions (Figure 1C).

Groups of strains

Based on all dataset from both fatty acids, pigments and annotated ions, a complete matrix further called “total metabolite matrix” was assembled for performing global metabolite analyses (see Materials and methods). Considering this total metabolite matrix, a multiple factor analysis (MFA) combined to a hierarchical clustering on principal components (HCPC) was performed considering the effect of the sampling site factor and highlighted three distinct groups of strains (Figure 2A and 2B, respectively). Remarkably, those three groups that were statistically different (Figure 2A, ANOSIM $p < 0.001$) and also similarly discriminate

regarding both fatty acids, pigments and annotated ions, when considered separately (supplementary Figure S2A-C), were overall matching with the sampling location factor (Figure 2C). Indeed, the strains isolated from different sampling sites were clearly clustered into different groups: Camargue region in group 1, Lake Dziani Dzaha in group 3, and another sampling site in group 2 (Figure 2B). Indeed, these analyses showed that strains coming from the lake Natron, Peru, and from the three unknown locations were overall grouped together in a single group 3 (Figure 2A-B). Thus, those strains of group 3 were further considered together in a site group named “Others”. According to the multiple correspondence analysis (Figure 2D), the correlation between these three groups (1=“Camargue”, 3=“Dziani” and 2=“Others”) and the different environmental factors was investigated. Group 1 mostly correlated to the range of pH of 7-8, salinity of 10-35‰, brackish habitat, and years of isolation 2019 and 2020, while Group 3 correlated to the range of pH of 9-11, salinity of 35-70‰, alkaline-saline habitat, and years of isolation 2010 and 2014, and Group 2 with other pH, salinity, habitat, and year of isolation prior to 2000 (Figure 2 C-D). The three “sampling site” groups highlighted by the HCPC were used for further analysis regarding their respective metabolite signatures.

Sampling sites

MFA and BCA analyses were performed considering the three groups that were previously separated by HCPC and clearly correspond to the different sampling sites (Figure 3A-B). The primary axis of the BCA analysis accounted for 63.1% of the variability. It primarily distinguished the *L. platensis* samples taken at Dziani (group 3) from the others, based on their chemical composition (Figure 3B). The second axis explained 37% of total inertia and mainly opposed HCPC group 2 from all other strains (Figure 3B). Among fatty acids, pigments, and annotated ions, 31 most discriminant variants were observed and present a cos^2

> 0.4 (Figure 3C): two fatty acids, nine pigments, and 20 annotated ions (26 related to axis 1, and five to axis 2). In total, 126 molecules showed significative difference between at least two groups (Kruskal-Wallis test, $p < 0.001$). The two fatty acids were the PUFAs with the longest carbon chains, 20:2ω6 and 20:3ω6 that were significantly more present in the strains from lake Dziani Dzaha in concentration and proportion of total fatty acids (Figure 4, Kruskal-Wallis test, $p < 0.001$). Based on fatty acids only, two groups of 42 and 49 strains have been discriminated by hierarchal clustering, not related to any of the environmental factor (Supplementary Figure 2B). They significantly differed by the total fatty acid biomass (30 ± 8 $mg.g^{-}$and 71 ± 15 $mg.g^{-1}$, respectively).

Six of the pigments were only present in the strains from lake Dziani Dzaha (Figures 1A and 4, Kruskal-Wallis test, $p < 0.001$). Among them, three were recognized as resembling caloxanthin and nostoxanthin and an analog of β,β-carotene , and three were unknown carotenoids. The strains from lake Dziani Dzaha did not possessed the usual form of β,β-carotene, that was only present in the strains from Camargue and other locations. The allophycocyanin were significantly more present in the strains from other locations (Figure 4, Kruskal-Wallis test, $p < 0.001$). Among the 20 most discriminant annotated metabolites (Figure 5), 13 were significantly more abundant in the strains from lake from lake Dziani Dzaha, including four saccharides (MW = 254.1022, 596.2184, 684.2349 Da, and sucrose), three little peptides of 2 or 3 amino acids (MW = 373.2214, 359.2057, and 260.1376 Da), two phospholipids (MW = 467.304, and 479.3041 Da), one lipid (MW = 455.3042 Da), and two unsaturated fatty acid glycerol annotated as monoolein and alpha-linolenoyl-glycerol (Kruskal-Wallis test, $p < 0.001$). Three molecules were significantly less abundant in the strains from lake Dziani Dzaha: two saccharides (MW = 285.1065, and 268.0803 Da), and S-methylglutathione (Kruskal-Wallis test, $p < 0.001$). Four molecules were significantly more present in the strains from other locations: two peptides (MW = 359.1869, and 373.1986 Da),

one fatty acid ester (MW = 588.3765 Da), and oxidized glutathione (Kruskal-Wallis test, $p < 0.001$).

Gene content

Among the 93 *Limnospira* strains presently investigated, 33 of them present already available genomic information (supplementary table S1) that could explored for the detecting distinctive genetic traits. that could be related to the clustered strains according to their chemical signatures. Interestingly, the analysis of the presence/absence of coding genes perfectly match with distance pattern between the strains determined according to the core-genome sequence comparison, by distinguishing two main groups of strains (I and II) (supplementary figure S3). The first group is divided into two subgroups (I.1 and I.2); the clade I.1 contains 23 strains from the Camargue region, and the clade I.2 four strains from Lake Dziani Dzaha, two strains from Lake Natron (*L. platensis* PMC 1042.18 and SAG 85.79, and the strain *L. platensis* PCC 8005. The clade II contains the last three strains (*L. platensis* PCC 7345, PMC 289.06, and Paracas R14).

Interestingly, these three coding gene clades (I.1, I.2 and II) can also be clearly discriminated according to both MFA and BC-MFA analyses (Supplementary Figure S4A-B). The chemical composition of all three clades was mostly separated by the first axis of the MFA and the BC-MFA, while all clade II were separated from clades I.1 and I.2 by the second axis of the BC-MFA. Among fatty acids, pigments, and annotated ions, 18 most discriminant molecules were observed on BC-MFA variable plot to present a $\cos^2 > 0.4$ (Supplementary Figure 4C): seven fatty acids, seven pigments, and four annotated ions (17 related to axis 1, and one to axis 2). In total, 63 molecules showed significant difference between at least two groups (Kruskal-Wallis test, $p < 0.01$). The seven fatty acids, as well as the seven pigments, were less abundant in clade I.1 (Supplementary Figure S5, Kruskal-Wallis test, $p < 0.01$). Among the four most

discriminant annotated ions, two saccharides (MW = 536.1602 and 430.133 Da) were significantly more abundant in the strains of clade I.1. The two other annotated ions were significantly more abundant in the strains of clade II: oxidized glutathione and a fatty acid ester (MW = 588.3765 Da) (Supplementary Figure S5, Kruskal-Wallis test, $p < 0.01$).

Discussion

In the present work, we have developed an original approach based on the investigation specific adaptation of the different chemotypes to their original environmental conditions, through the observation of their regular chemical traits. Thus, we hypothesized that the original chemical traits of these different genotypes might offer them remarkable adaptative features that have been shaped by the specific conditions of their respective originating environments.

Indeed, analyses of the pigment, lipid, and small metabolite contents of various *Limnospira platensis* strains cultivated under identical laboratory conditions provide unprecedented data that offers a plain answer to the foremost question: is the diversity of these molecules driven by selective pressures on *Limnospira platensis* (clade I or II)[20] and/or by the environment from which the strains are isolated? Our results clearly show that the sampling location that offers diverse environmental conditions (*e.g.* salinity, pH, light, or temperature) constitutes the variables explaining the distribution of *Limnospira platensis* in three distinct chemotype groups. The molecules that most discriminate these three groups of the *Limnospira platensis* strains studied, under the standard conditions of lab cultures, are saccharides, fatty acids, peptides, photosynthetic pigments, and antioxidant molecules. Interestingly, these molecule enrichments might constitute physiological adaptation traits to the local conditions encountered in their respective sampling environments, potentially in relation to local salinity, light and oxidative stress. These points are then critically debated in the following discussion.

Osmoregulation and salinity resistance

The ponds of the Camargue in France and Lake Dziani Dzaha in Mayotte are environments with different salinities. In the brackish ponds of Camargue, close to the Mediterranean Sea, salinity varies between 12-35‰. Lake Dziani Dzaha is a thalassohaline lake. Here, salinity is much higher, reaching up to 70‰[33]. Strains from Lake Dziani Dzaha have greater sugar diversity, have more long-chain polyunsaturated fatty acids compared with strains isolated from other habitats under our experimental conditions. They are also the only ones to produce nostoxanthin and caloxanthin. This specific chemical signature of strains from Lake Dziani Dzaha testifies to the adaptation of the *Limnospira platensis* species to resist to higher salinity in this specific environment.

In cyanobacteria in general, several families of molecules are involved in osmoregulation and detoxification of stress due to high salinity[34-35]. Cyanobacteria produce various sugars (*e.g.* glucose, sucrose, trehalose or glycine betaine) to prevent the osmotic stress caused by high salinity[36]. In *Synechococcus*, long-chain polyunsaturated fatty acids help combat the oxidative stress caused by high salinity as they appear to be produced to a greater extent with increasing salinity[37]. Interestingly, it has also been shown that certain carotenoids, nostoxanthin and caloxanthin, are produced by a bacterium of the genus *Sphingomonas* that appears resistant to high salinity[38]. In addition, the nostoxanthin and caloxanthin-producing bacterium *Sphingomonas nostxanthinifaciens* may have a role in reducing the amount of reactive oxygen species (ROS) in *Arabidopsis thaliana* when the two organisms are co-cultured in high salt concentrations[39]. Taken together, these finding support the idea that sugar and carotenoid molecules might support high salinity adaptation process in *Limnospira platensis*.

Light intensity

Facing high light, cyanobacteria modify chlorophyll, phycobiliprotein, and carotenoid composition to improve electron transfer and manage excess energy[40-42]. In *Synechococcus*, various carotenoids are crucial for high light survival. Strains lacking carotenoid biosynthesis genes (*cru* genes) have lower growth rates and higher ROS levels compared to carotenoid-producing strains[40].

The different *Limnospira platensis* strains relied on a wide diversity of chlorophylls (n = 5-12) and carotenoids (n = 13-18) that could accordingly be involved in photo-pigment adaptation. One of the most striking results is the presence of six exclusive carotenoids that appear exclusively in strains from Lake Dziani Dzaha. This lake has a mean surface irradiance of 2,000 µmol photon.s^{-1}.m^{-2}, which decreases by 99% in the first 20 cm of surface water to around 20 µmol photon.s^{-1}.m^{-2}.[33] *Limnospira* living in the photic zone of the lake are therefore subjected to both high and strong variations in light intensity over shallow depths. We found 18 carotenoid pigments that have an unknown chemical structure (putative analogues of other pigments, n = 7, or totally unknown pigments, n = 11). This observation illustrates the ability of *Limnospira* to diversify its pigment content in response to a given environment. The great diversity of pigments produced by *Limnospira platensis* was shown by the characterization of up to 48 hydrophobic pigments (22 carotenoids and 26 chlorophylls) by LC-APCI-FT-ICR-MS[43], some of which could very likely correspond to those unknown pigments observed in our study.

Oxidative stress

In a general manner, both high salinity, light and UV exposures are interdependently involved in ROS production and subsequent cellular oxidative stress[44]. In addition to primary mechanisms listed above, aiming at reducing the adverse effect of oxidative stress, cyanobacteria also produce a large variety of other molecules (other than carotenoids and

long-chain polyunsaturated fatty acids). These molecules are specifically involved in buffering ROS production and combating oxidative stress deleterious effects, such as ergothioneine, glutathione, biopterins and their derivatives[14,45-46]. The cellular extracts of all *Limnospira platensis* strains presently investigated contained significant amounts of each of these molecules, that might constitute a set of molecular resource for this organism to deal with remarkable local cellular stressors.

***Limnospira platensis* presents a peculiar ecological strategy**

Cyanobacteria have colonized many types of habitats (*e.g.* oceanic waters, fresh waters, alkaline waters, benthic or pelagic, terrestrial, symbiotic)[2]. Aquatic species have for example developed different ecological strategies to adapt to their environment. Some cyanobacteria live in diversified communities of various species being concomitant or consecutive along the proliferative events. These include benthic cyanobacteria, which form biofilms (*e.g. Phormidium*, *Lyngbya*, *Oscillatoria*)[47], and pelagic cyanobacteria, which can form blooms (*e.g. Microcystis*, *Aphanizomenon, Raphidiopsis, …*)[48-50]. The genomes of these cyanobacteria are remarkably large (> 5Mb) and contain many BGCs[51-53], producing various bioactive secondary metabolites (including cyclic peptides, lipopeptides, alkaloids, macrolides, polyketides, etc.) that enable cyanobacteria to interact with their microbial community (*e.g.* anti-grazing, anti-algal, communication blocking, competition for resources)[5,54]. However, the production of these specific molecules has a high energy cost for the producing cells, as they are involving large gene sets and the synthesis of many enzymes[55]. In contrary, some marine picocyanobacteria (*e.g. Prochlorococcus*, *Synechococcus*) are exclusively dominating one type of habitat (*i.e.* the worldwide oligothrophic seas) and exhibit a unique diversity of photosynthetic pigments, allowing them to exploit a wide range of light micro-niches[56-57]. The genomes of these cyanobacteria are remarkably small (1.5-2.5Mb) and contain no BGCs[51,58].

In these cyanobacteria, adaptation to an ecological niche is encoded in both core and flexible parts of the genomes[59-60].

For *Limnospira platensis*, we denote a potential third strategy. *Limnospira platensis* genomes are relatively large, ranging between 5.5 and 6.5Mb[20,61], being comparable to the genome size of community-dwelling cyanobacteria proliferating in eutrophic waters. Yet, they remarkably contain no or only one BGC of secondary metabolites (potentially involved in the biosynthesis of cyanobactins called Arthrospiramides), as do the specialized marine picocyanobacteria *Synechococcus* and *Prochlorococcus*[51,62-63]. Unlike marine *Synechococcus* and *Prochlorococcus*, *Limnospira platensis* proliferates in many different habitat types: freshwater, brackish water, alkaline water, and saline alkaline water[20,64], and *Limnospira platensis* strains belonging to the same phylogenetic clade do not share similar habitats[20]. Moreover, in lakes where phytoplanktonic communities are dominated by *Limnospira platensis*, this species frequently generates persistent proliferations. These blooms are often accompanied by minimal surrounding diversity (except de picoeukaryote *Picocystis salinarum*)[21,63], serving as a distinctive indicator of extreme environments (*e.g.*, hypersaline conditions, high light coupled with high turbidity, elevated temperatures, high pH levels, and high sulfur contents). In such environments, only a limited number of microorganism clades can develop. This absence of surrounding microbial diversity, comprising potential competitors, could explain the absence of BGCs in the *Limnospira platensis* genomes, as it has fewer phototrophic organisms with which to interact through allelopathic compound production, as various cyanobacteria from high diversity freshwater temperate lakes do[51]. Interestingly, such an ecological strategy could support the high biomass production observed both *in situ* and *in vitro*, as lab culture conditions present many similarities (in terms of both biotic and abiotic factors) with those of the extreme environment in which *Limnospira* proliferates. Indeed, it has been shown by transcriptomics that the essential transcribed

functions expressed *in vitro* are related to primary metabolism (metabolism of sugar, fatty acids, nitrogen, CO_2, photosynthesis, and antioxidant processes)[64].

Interestingly, the genomic investigation of *Limnospira platensis* strains shows that both well-conserved (core) and flexible (pan) genome parts discriminate strains according to their respective origin locations (supplementary Figure S3). Thus, one can suppose that either common point or InDel mutations on key metabolic genes of the core-genome or variation in the flexible gene content of *Limnospira platensis* genomes may support peculiar adaptations to local environmental conditions. Interestingly, this latter flexible part of the genome is remarkably rich in transposable and regulatory elements and contains various genes, most of which have as yet unknown molecular functions, and may dynamically support evolutionary exaptation processes such as ecological niche extension due to higher stress resistance aquisition[20]. So far, the relationship between small metabolite content (*i.e.*, redox agents, saccharides and pigments/carotenoids) - chemotype - and niche partitioning (*i.e.*, ROS scavenging, osmoregulatory, and photoprotective molecular traits) - ecotype - still remains largely undetermined due to the lack of distinctive genetic markers associated with the various anti-oxidant, osmolyte and pigment production. Such assumptions now require further experimental validation.

Conclusion

Taken together, these ecological and evolutionary concerns offer new comprehension of the mechanisms that support the remarkable success encountered by *Limnospira* in natural as well as lab or industrial environments. One step further would be to try to experimentally identify the genetic and phenotypic traits that are defining the different *Limnospira* ecotypes and to determine whether *Limnospira* clade I.1, I.2 and II genotypes can colonize similar ecological

niches, as this taxon exhibits ubiquitous geographic distribution, or represents already well-distinct ecologically and functionally distinguishable taxa.

Also, key questions remain on the understanding of the mechanisms that allow *Limnospira in situ* and *in vitro* blooms to avoid density-dependent regulations of its proliferation by cyanophages, as those encountered by other blooming cyanobacteria[65]. Such adaptative mechanisms would help explain how *Limnospira platensis* can form permanent blooms in many environments. To this end, two distinct hypotheses can be further tested: either the *Limnospira platensis* genome presents a mechanism to protect itself from phage lyses, or the fraction of the population potentially lysed by phages would be replaced by rapidly dividing *Limnospira platensis*.

Author Contributions: Conceptualization, T.R., J.-P.C., L.V., C.B. and B.M.; methodology, T.R., C.H., M.C., C.Y., B.M.; investigation, T.R., C.H., T.M.; resources, C.D.; writing-original draft preparation, T.R., C.B. and B.M.; writing—review and editing, all authors. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by ANRT, through a convention Cifre n° 2020/058 awarded to the Company Algama employing T. Roussel for a PhD thesis.

Acknowledgments:

We are grateful to the anonymous reviewer for their valuable comments and suggestions, which have helped improve the quality of our manuscript. We would like to thank the UMR 7245 MCAM, Muséum National d'Histoire Naturelle, Paris, France for laboratories facilities and Algama for funds. The MS spectra were acquired at the Plateau technique de

spectrométrie de masse bio-organique, Muséum National d'Histoire Naturelle, Paris, France. The authors thank the Pasteur Collection of Cyanobacteria, the Culture Collection of Algae at Göttigen University, and Gilles Planchon for sending strains.

Conflicts of Interest: The authors declare no conflict of interest.

References

1- Reynolds, C. S. (2006). *The ecology of phytoplankton*. Cambridge University Press.

2- Whitton, B.A. (ed) (2012) Ecology of Cyanobacteria II: Their Diversity in Space and Time. Dordrecht: Springer Netherlands. Available online: https://doi.org/10.1007/978-94-007-3855-3.

3- Whitton, B.A. & Potts, M. (eds) (2007) The Ecology of Cyanobacteria: Their Diversity in Time and Space. Dordrecht: The Netherlands, Springer Science & Business Media.

4- Leflaive, J. & Ten-Hage, L. (2007) Algal and cyanobacterial secondary metabolites in freshwaters: a comparison of allelopathic compounds and toxins. Freshwater Biology, 52(2), 199–214. Available online: https://doi.org/10.1111/j.1365-2427.2006.01689.x.

5- Demay, J., Bernard, C., Reinhardt, A. & Marie, B. (2019) Natural Products from Cyanobacteria: Focus on Beneficial Activities. Marine Drugs, 17(6), 320. Available online: https://doi.org/10.3390/md17060320.

6- Nuhu, A. A. (2013). Spirulina (Arthrospira): An important source of nutritional and medicinal compounds. *Journal of Marine biology*, *2013*, 1-8.

7- Shao, W., Ebaid, R., El-Sheekh, M., Abomohra, A., & Eladel, H. (2019). Pharmaceutical applications and consequent environmental impacts of Spirulina (Arthrospira): An overview. *Grasas y Aceites*, *70*(1), e292-e292.

8- Vonshak, A. & Tomsaelli, L. (2000) Arthrospira (Spirulina): Systematics and Ecolophysiology. In Whittons, B.A. & Potts, M. The ecology of Cyanobacteria: Their Diversity in Time and Space.

9- Vonshak, A. (ed) (1997) Spirulina platensis (Arthrospira): Physiology, Cell-biology, and Biotechnology. London, Bristol, PA: Taylor & Francis.

10- Avila-Leon, I., Chuei Matsudo, M., Sato, S. & De Carvalho, J.C.M. (2012) Arthrospira platensis biomass with high protein content cultivated in continuous process using urea as nitrogen source: High protein biomass by continuous process with urea. Journal of Applied Microbiology, 112(6), 1086–1094. Available online: https://doi.org/10.1111/j.1365-2672.2012.05303.x.

11- Ötleş, S., & Pire, R. (2001). Fatty acid composition of Chlorella and Spirulina microalgae species. *Journal of AOAC international*, *84*(6), 1708-1714.

12- Martinez Andrade, K., Lauritano, C., Romano, G. & Ianora, A. (2018) Marine Microalgae with Anti-Cancer Properties. Marine Drugs, 16(5), 165. Available online: https://doi.org/10.3390/md16050165.

13- Edelmann, M., Aalto, S., Chamlagain, B., Kariluoto, S. & Piironen, V. (2019) Riboflavin, niacin, folate and vitamin B12 in commercial microalgae powders. Journal of Food Composition and Analysis, 82, 103226. Available online: https://doi.org/10.1016/j.jfca.2019.05.009.

14- Pfeiffer, C., Bauer, T., Surek, B., Schmig, E. & Gründemann, D. (2011) Cyanobacteria produce high levels of ergothioneine. Food Chemistry, 129(4), 1766–1769. Available online: https://doi.org/10.1016/j.foodchem.2011.06.047.

15- Markou, G. (2014) Effect of Various Colors of Light-Emitting Diodes (LEDs) on the Biomass Composition of Arthrospira platensis Cultivated in Semi-continuous Mode. Applied Biochemistry and Biotechnology, 172(5), 2758–2768. Available online: https://doi.org/10.1007/s12010-014-0727-3.

16- Bezerra, R.P., Matsudo, M.C., Sato, S., Perego, P., Converti, A. & Carvalho, J.C.M.D. (2012) Effects of photobioreactor configuration, nitrogen source and light intensity on the fed-batch cultivation of Arthrospira (Spirulina) platensis. Bioenergetic aspects. Biomass and Bioenergy, 37, 309–317. Available online: https://doi.org/10.1016/j.biombioe.2011.11.007.

17- Manirafasha, E., Murwanashyaka, T., Ndikubwimana, T., Rashid Ahmed, N., Liu, J., Lu, Y., Zeng, X., Ling, X. & Jing, K. (2018) Enhancement of cell growth and phycocyanin production in Arthrospira (Spirulina) platensis by metabolic stress and nitrate fed-batch. Bioresource Technology, 255, 293–301. Available online: https://doi.org/10.1016/j.biortech.2017.12.068.

18- Tanaka, K., Kishi, M., Assaye, H. & Toda, T. (2020) Low temperatures in dark period affect biomass productivity of a cyanobacterium Arthrospira platensis. Algal Research, 52, 102132. Available online: https://doi.org/10.1016/j.algal.2020.102132.

19- Sili, C., Torzillo, G. & Vonshak, A. (2012) Arthrospira (Spirulina). In Whitton, B.A. (ed) Ecology of Cyanobacteria II: Their Diversity in Space and Time. Dordrecht: Springer Netherlands. Available online: https://doi.org/10.1007/978-94-007-3855-3.

20- Roussel, T., Halary, S., Duval, C., Piquet, B., Cadoret, J. P., Vernès, L., ... & Marie, B. (2023). Monospecific renaming within the cyanobacterial genus Limnospira (Spirulina) and consequences for food authorization. *Journal of Applied Microbiology*, *134*(8), lxad159.

21- Cellamare, M., Duval, C., Drelin, Y., Djediat, C., Touibi, N., Agogué, H., ... & Bernard, C. (2018). Characterization of phototrophic microorganisms and description of new cyanobacteria isolated from the saline-alkaline crater-lake Dziani Dzaha (Mayotte, Indian Ocean). *FEMS microbiology ecology*, *94*(8), fiy108.

22- Rossello-Mora, R., Lucio, M., Pena, A., Brito-Echeverría, J., Lopez-Lopez, A., Valens-Vadell, M., ... & Schmitt-Kopplin, P. (2008). Metabolic evidence for biogeographic isolation of the extremophilic bacterium Salinibacter ruber. *The ISME journal*, *2*(3), 242-253.

23- Kujawinski EB, Braakman R, Longnecker K, Becker JW, Chisholm SW, Dooley K, Kido Soule MC, Swarr GJ, Halloran K. Metabolite diversity among representatives of divergent *Prochlorococcus* ecotypes. mSystems. 2023 Oct 26;8(5):e0126122. doi: 10.1128/msystems.01261-22. Epub 2023 Oct 10. Erratum in: mSystems. 2024 Mar 19;9(3):e0012024.

24- Hamlaoui, S., Yéprémian, C., Duval, C., Marie, B., Djédiat, C., Piquet, B., Bernard, C. & Duperron, S. (2022) The Culture Collection of Cyanobacteria and Microalgae at the French National Museum of Natural History: A Century Old But Still Alive and Kicking! Including in Memoriam: Professor Alain Cout.. Cryptogamie, Algologie, 43(3). Available online: https://doi.org/10.5252/cryptogamie-algologie2022v43a3.

25- Yéprémian, C., Catherine, A., Bernard, C., Congestri, R., Elersek, T. & Pilkaityte, R. (2017a) Phycocyanin Extraction and Determination. In Meriluoto, J., Spoof, L., & Codd, G.A. (eds) Handbook of Cyanobacterial Monitoring and Cyanotoxin Analysis. 1st edition. Wiley, 335–338. Available online: https://doi.org/10.1002/9781119068761.ch35.

26- Yéprémian, C., Catherine, A., Bernard, C., Congestri, R., Elersek, T. & Pilkaityte, R. (2017b) Chlorophyll a Extraction and Determination. In Meriluoto, J., Spoof, L., &

Codd, G.A. (eds) Handbook of Cyanobacterial Monitoring and Cyanotoxin Analysis. 1st edition. Wiley, 331–334. Available online: https://doi.org/10.1002/9781119068761.ch34.

27- De Marsac, N. T., & Houmard, J. (1988). [34] Complementary chromatic adaptation: physiological conditions and action spectra. In *Methods in enzymology* (Vol. 167, pp. 318-328). Academic Press.

28- Brotas, V. & Plante-Cuny, M.R. (2003) The use of HPLC pigment analysis to study microphytobenthos communities. Acta Oecologica, 24(1), S109-S115. Available online: https://doi.org/10.1016/S1146-609X(03)00013-4.

29- Bligh EG, Dyer WJ. A rapid method of total lipid extraction and purification. Can J Biochem Physiol. 1959 Aug;37(8):911-7. doi: 10.1139/o59-099. PMID: 13671378.

30- Meziane, T., Lee, S.Y., Mfilinge, P.L., Shin, P.K.S., Lam, M.H.W. & Tsuchiya, M. (2007) Inter-specific and geographical variations in the fatty acid composition of mangrove leaves: implications for using fatty acids as a taxonomic tool and tracers of organic matter. Marine Biology, 150(6), 1103–1113. Available online: https://doi.org/10.1007/s00227-006-0424-z.

31- Le Moigne, D., Demay, J., Reinhardt, A., Bernard, C., Kim Tiam, S., & Marie, B. (2021). Dynamics of the Metabolome of Aliinostoc sp. PMC 882.14 in Response to Light and Temperature Variations. *Metabolites*, *11*(11), 745.

32- Michelet, C., Zeppilli, D., Hubas, C., Baldrighi, E., Cuny, P., Dirberg, G., Militon, C., Walcker, R., Lamy, D., Jézéquel, R., Receveur, J., Gilbert, F., Houssainy, A.E., Dufour, A., Heimbürger-Boavida, L.-E., Bihannic, I., Sylvi, L., Vivier, B. & Michaud, E. (2021) First Assessment of the Benthic Meiofauna Sensitivity to Low Human-Impacted Mangroves in French Guiana. Forests, 12(3), 338. Available online: https://doi.org/10.3390/f12030338.

33- Leboulanger, C., Agogué, H., Bernard, C., Bouvy, M., Carré, C., Cellamare, M., Duval, C., Fouilland, E., Got, P., Intertaglia, L., Lavergne, C., Le Floc'h, E., Roques, C. & Sarazin, G. (2017) Microbial Diversity and Cyanobacterial Production in Dziani Dzaha Crater Lake, a Unique Tropical Thalassohaline Environment. PLOS ONE. Edited by J.-F. Humbert, 12(1), e0168879. Available online: https://doi.org/10.1371/journal.pone.0168879.

34- Hagemann M. Molecular biology of cyanobacterial salt acclimation. FEMS Microbiol Rev. 2011 Jan;35(1):87-123. doi: 10.1111/j.1574-6976.2010.00234.x.

35- Duval, C., Thomazeau, S., Drelin, Y., Yepremian, C., Bouvy, M., Couloux, A., ... & Bernard, C. (2018). Phylogeny and salt-tolerance of freshwater Nostocales strains: contribution to their systematics and evolution. *Harmful algae*, *73*, 58-71.

36- Mackay, M.A., Norton, R.S. & Borowitzka, L.J. (1984) Organic Osmoregulatory Solutes in Cyanobacteria. Microbiology, 130(9), 2177–2191. Available online: https://doi.org/10.1099/00221287-130-9-2177.

37- Verma, E., Singh, S., Niveshika & Mishra, A.K. (2019) Salinity-induced oxidative stress-mediated change in fatty acids composition of cyanobacterium Synechococcus sp. PCC7942. International Journal of Environmental Science and Technology, 16(2), 875–886. Available online: https://doi.org/10.1007/s13762-018-1720-0.

38- Kikukawa, H., Okaya, T., Maoka, T., Miyazaki, M., Murofushi, K., Kato, T., Hirono-Hara, Y., Katsumata, M., Miyahara, S. & Hara, K.Y. (2021) Carotenoid Nostoxanthin Production by Sphingomonas sp. SG73 Isolated from Deep Sea Sediment. Marine Drugs, 19(5), 274. Available online: https://doi.org/10.3390/md19050274.

39- Jiang, L., Seo, J., Peng, Y., Jeon, D., Lee, J.H., Kim, C.Y. & Lee, J. (2023) A nostoxanthin-producing bacterium, Sphingomonas nostoxanthinifaciens sp. nov., alleviates the salt stress of Arabidopsis seedlings by scavenging of reactive oxygen species. Frontiers

in Microbiology, 14, 1101150. Available online: https://doi.org/10.3389/fmicb.2023.1101150.

40- Zhu, Y., Graham, J.E., Ludwig, M., Xiong, W., Alvey, R.M., Shen, G. & Bryant, D.A. (2010) Roles of xanthophyll carotenoids in protection against photoinhibition and oxidative stress in the cyanobacterium Synechococcus sp. strain PCC 7002. Archives of Biochemistry and Biophysics, 504(1), 86–99. Available online: https://doi.org/10.1016/j.abb.2010.07.007.

41- Kumar, M., Kulshreshtha, J. & Singh, G.P. (2011) Growth and biopigment accumulation of cyanobacterium Spirulina platensis at different light intensities and temperature. Brazilian Journal of Microbiology, 42(3), 1128–1135. Available online: https://doi.org/10.1590/S1517-83822011000300034.

42- Masamoto, K. & Furukawa, K. (1997) Accumulation of zeaxanthin in cells of the cyanobacterium, Synechococcus sp. Strain PCC 7942 grown under high irradiance. Journal of Plant Physiology, 151(3), 257–261. Available online: https://doi.org/10.1016/S0176-1617(97)80250-7.

43- Sommella, E., Conte, G. M., Salviati, E., Pepe, G., Bertamino, A., Ostacolo, C., ... & Campiglia, P. (2018). Fast profiling of natural pigments in different spirulina (arthrospira platensis) dietary supplements by DI-FT-ICR and evaluation of their antioxidant potential by pre-column DPPH-UHPLC assay. *Molecules*, *23*(5), 1152.

44- Khorobrykh, S., Havurinne, V., Mattila, H. & Tyystjérvi, E. (2020) Oxygen and ROS in Photosynthesis. Plants, 9(1), 91. Available online: https://doi.org/10.3390/plants9010091.

45- Noguchi, Y., Ishii, A., Matsushima, A., Haishi, D., Yasumuro, K., Moriguchi, T., Wada, T., Kodera, Y., Hiroto, M., Nishimura, H., Sekine, M. & Inada, Y. (1999) Isolation of Biopterin-α-glucoside from Spirulina (Arthrospira) platensis and Its Physiologic Func-

tion. Marine Biotechnology, 1, 207–210. Available online: https://doi.org/10.1007/PL00011769.

46- Latifi, A., Ruiz, M. & Zhang, C.-C. (2009) Oxidative stress in cyanobacteria. FEMS Microbiology Reviews, 33(2), 258–278. Available online: https://doi.org/10.1111/j.1574-6976.2008.00134.x.

47- Quiblier, C., Wood, S., Echenique-Subiabre, I., Heath, M., Villeneuve, A. & Humbert, J.-F. (2013) A review of current knowledge on toxic benthic freshwater cyanobacteria – Ecology, toxin production and risk management. Water Research, 47(15), 5464–5479. Available online: https://doi.org/10.1016/j.watres.2013.06.042.

48- Cirés, S. & Ballot, A. (2016) A review of the phylogeny, ecology and toxin production of bloom-forming Aphanizomenon spp. and related species within the Nostocales (cyanobacteria). Harmful Algae, 54, 21–43. Available online: https://doi.org/10.1016/j.hal.2015.09.007.

49- Jones, A.C., Monroe, E.A., Podell, S., Hess, W.R., Klages, S., Esquenazi, E., Niessen, S., Hoover, H., Rothmann, M., Lasken, R.S., Yates, J.R., Reinhardt, R., Kube, M., Burkart, M.D., Allen, E.E., Dorrestein, P.C., Gerwick, W.H. & Gerwick, L. (2011) Genomic insights into the physiology and ecology of the marine filamentous cyanobacterium Lyngbya majuscula. Proceedings of the National Academy of Sciences, 108(21), 8815–8820. Available online: https://doi.org/10.1073/pnas.1101137108.

50- Harke, M.J., Steffen, M.M., Gobler, C.J., Otten, T.G., Wilhelm, S.W., Wood, S.A. & Paerl, H.W. (2016) A review of the global ecology, genomics, and biogeography of the toxic cyanobacterium, Microcystis spp. Harmful Algae, 54, 4–20. Available online: https://doi.org/10.1016/j.hal.2015.12.007.

51- Shih, P.M., Wu, D., Latifi, A., Axen, S.D., Fewer, D.P., Talla, E., Calteau, A., Cai, F., Tandeau De Marsac, N., Rippka, R., Herdman, M., Sivonen, K., Coursin, T., Laurent,

T., Goodwin, L., Nolan, M., Davenport, K.W., Han, C.S., Rubin, E.M., Eisen, J.A., Woyke, T., Gugger, M. & Kerfeld, C.A. (2013) Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. Proceedings of the National Academy of Sciences, 110(3), 1053–1058. Available online: https://doi.org/10.1073/pnas.1217107110.

52- Leikoski, N., Liu, L., Jokela, J., Wahlsten, M., Gugger, M., Calteau, A., Permi, P., Kerfeld, C.A., Sivonen, K. & Fewer, D.P. (2013) Genome Mining Expands the Chemical Diversity of the Cyanobactin Family to Include Highly Modified Linear Peptides. Chemistry & Biology, 20(8), 1033–1043. Available online: https://doi.org/10.1016/j.chembiol.2013.06.015.

53- Halary, S., Duperron, S., Kim Tiam, S., Duval, C., Bernard, C. & Marie, B. (2023) Intra-population genomic diversity of the bloom-forming cyanobacterium, Aphanizomenon gracile, at low spatial scale. ISME Communications, 3(1), 57. Available online: https://doi.org/10.1038/s43705-023-00263-3.

54- Leao, P.N., Engene, N., Antunes, A., Gerwick, W.H. & Vasconcelos, V. (2012) The chemical ecology of cyanobacteria. Natural Product Reports, 29(3), 372. Available online: https://doi.org/10.1039/c2np00075j.

55- Fischbach, M.A. & Walsh, C.T. (2006) Assembly-Line Enzymology for Polyketide and Nonribosomal Peptide Antibiotics: Logic, Machinery, and Mechanisms. Chemical Reviews, 106(8), 3468–3496. Available online: https://doi.org/10.1021/cr0503097.

56- Grébert, T., Dor., H., Partensky, F., Farrant, G.K., Boss, E.S., Picheral, M., Guidi, L., Pesant, S., Scanlan, D.J., Wincker, P., Acinas, S.G., Kehoe, D.M. & Garczarek, L. (2018) Light color acclimation is a key process in the global ocean distribution of Synechococcus cyanobacteria. Proceedings of the National Academy of Sciences, 115(9). Available online: https://doi.org/10.1073/pnas.1717069115.

57- Sanfilippo, J.E., Garczarek, L., Partensky, F. & Kehoe, D.M. (2019) Chromatic Acclimation in Cyanobacteria: A Diverse and Widespread Process for Optimizing Photosynthesis. Annual Review of Microbiology, 73(1), 407–433. Available online: https://doi.org/10.1146/annurev-micro-020518-115738.

58- Biller, S.J., Berube, P.M., Berta-Thompson, J.W., Kelly, L., Roggensack, S.E., Awad, L., Roache-Johnson, K.H., Ding, H., Giovannoni, S.J., Rocap, G., Moore, L.R. & Chisholm, S.W. (2014) Genomes of diverse isolates of the marine cyanobacterium Prochlorococcus. Scientific Data, 1(1), 140034. Available online: https://doi.org/10.1038/sdata.2014.34.

59- Biller, S.J., Berube, P.M., Lindell, D. & Chisholm, S.W. (2015) Prochlorococcus: the structure and function of collective diversity. Nature Reviews Microbiology, 13(1), 13–27. Available online: https://doi.org/10.1038/nrmicro3378.

60- Doré, H., Guyet, U., Leconte, J., Farrant, G.K., Alric, B., Ratin, M., Ostrowski, M., Ferrieux, M., Brillet-Guéguen, L., Hoebeke, M., Siltanen, J., Le Corguillé, G., Corre, E., Wincker, P., Scanlan, D.J., Eveillard, D., Partensky, F. & Garczarek, L. (2023) Differential global distribution of marine picocyanobacteria gene clusters reveals distinct niche-related adaptive strategies. The ISME Journal, 17(5), 720–732. Available online: https://doi.org/10.1038/s41396-023-01386-0.

61- Hicks, M., Tran-Dao, T.-K., Mulroney, L. & Bernick, D.L. (2021) De-novo Assembly of Limnospira fusiformis Using Ultra-Long Reads. Frontiers in Microbiology, 12, 657995. Available online: https://doi.org/10.3389/fmicb.2021.657995.

62- Donia, M.S. & Schmidt, E.W. (2011) Linking Chemistry and Genetics in the Growing Cyanobactin Natural Products Family. Chemistry & Biology, 18(4), 508–519. Available online: https://doi.org/10.1016/j.chembiol.2011.01.019.

63- Krienitz, L., Bock, C., Kotut, K. & Préschold, T. (2012) Genotypic diversity of Dictyosphaerium-morphospecies (Chlorellaceae, Trebouxiophyceae) in African inland waters, including the description of four new genera. Fottea, 12(2), 231–253. Available online: https://doi.org/10.5507/fot.2012.017.

64- Kumaresan, V., Sannasimuthu, A., Arasu, M.V., Al-Dhabi, N.A. & Arockiaraj, J. (2018) Molecular insight into the metabolic activities of a protein-rich micro alga, Arthrospira platensis by de novo transcriptome analysis. Molecular Biology Reports, 45(5), 829–838. Available online: https://doi.org/10.1007/s11033-018-4229-1.

65- Grasso, C.R., Pokrzywinski, K.L., Waechter, C., Rycroft, T., Zhang, Y., Aligata, A., Kramer, M. & Lamsal, A. (2022) A Review of Cyanophage–Host Relationships: Highlighting Cyanophages as a Potential Cyanobacteria Control Strategy. Toxins, 14(6), 385. Available online: https://doi.org/10.3390/toxins14060385.

Figures

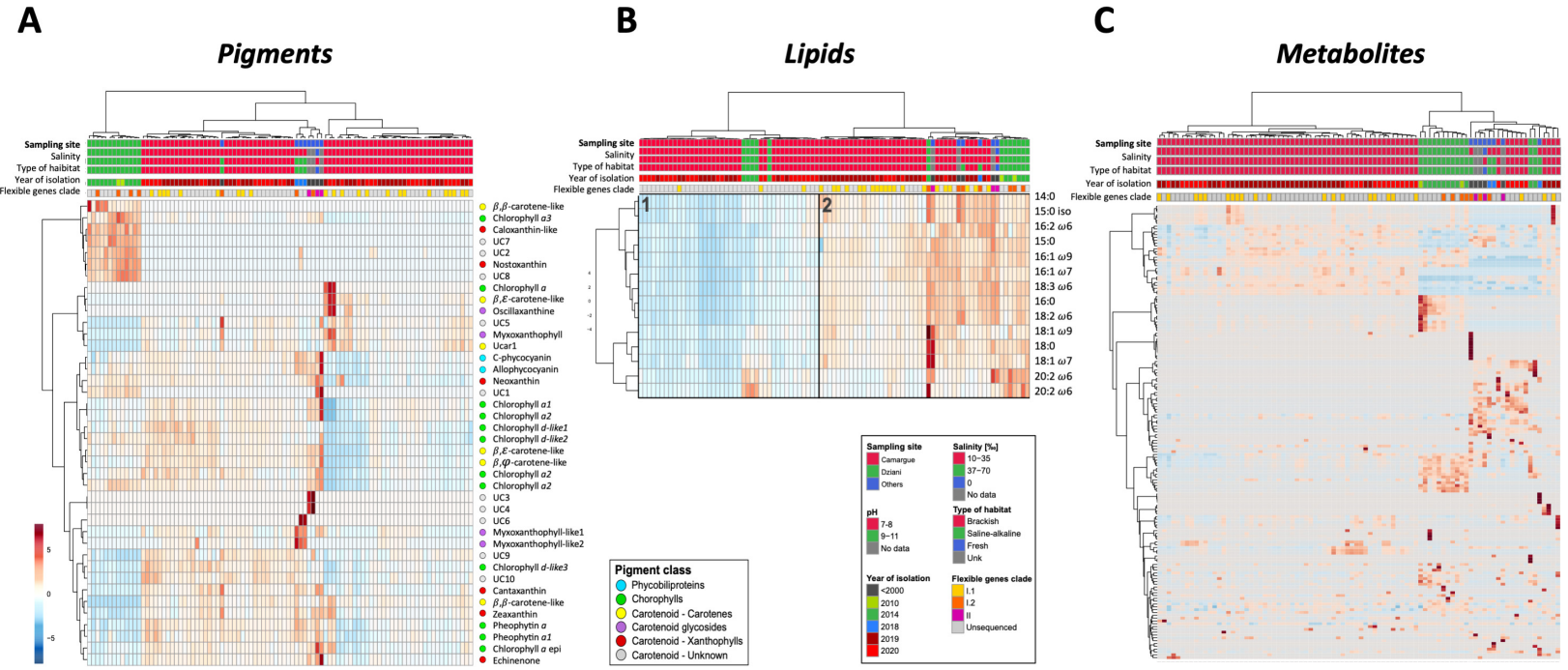
Figure 1. Heatmap analyses of pigment, lipid, and metabolite matrices. Heatmap analysis of the 40 pigment concentrations of the 93 analyzed *L. platensis* strains (normalization according to each feature) (A). Heatmap analysis of the normalized 14 fatty acid concentrations of the 91 analyzed *L. platensis* strains (B). Total fatty acid concentration ($mg.g^{-1}$ dw) discriminates between 2 groups of lower (1) and higher (2) lipid contents (Welch test, $p < 0.001$). (C) Heatmap analysis of the 161 HPLC-MS/MS annotated ions of the 88 analyzed *L. platensis* strains, expressed in normalized relative abundances.

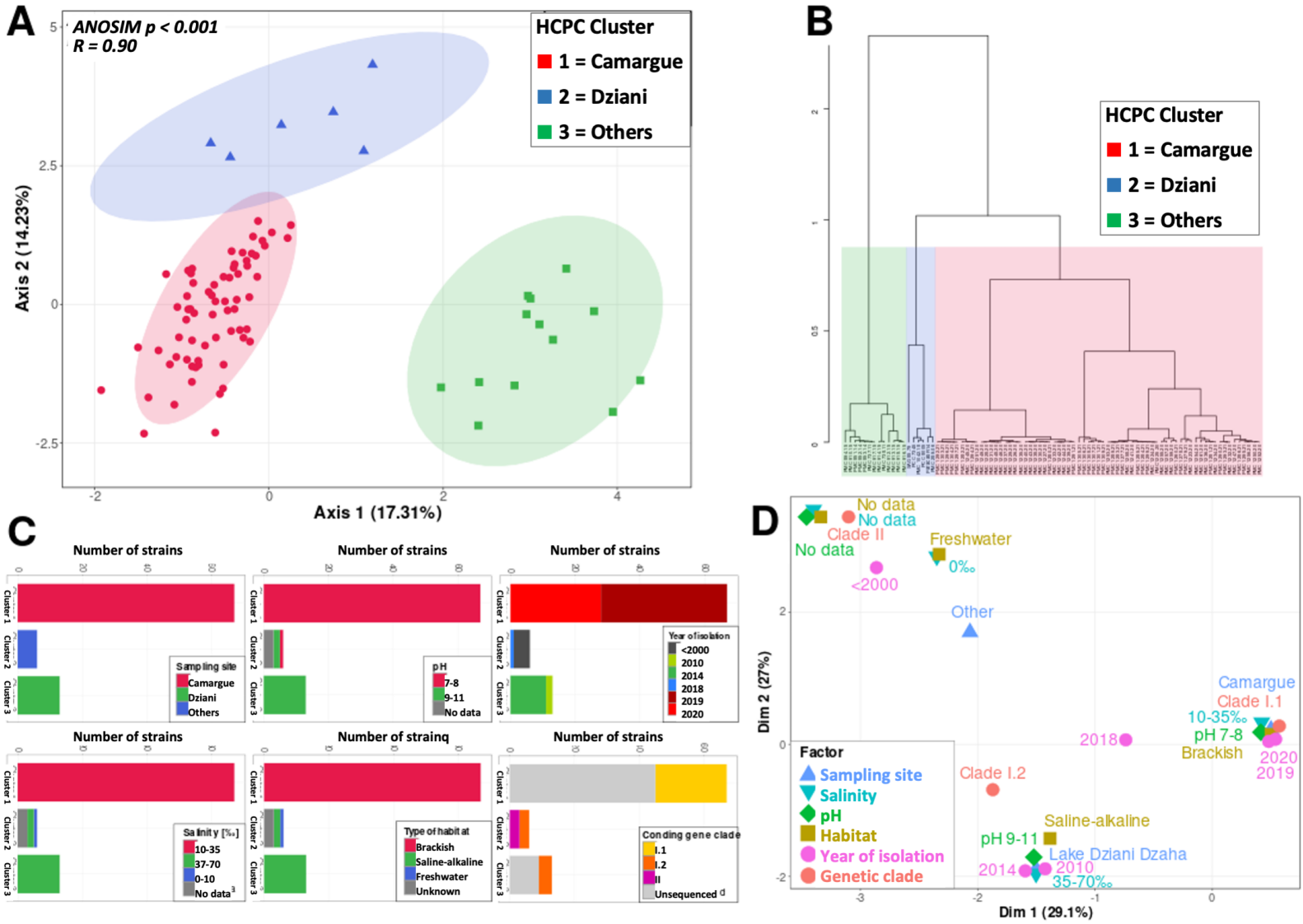
Figure 2. Principal component analysis (PCA) (A) represents the total matrix metabolite composition of the strains regarding the groups given by the hierarchical clustering on principal components (HCPC) (B). Characterization of each group regarding the sic tested factors (C) and the multiple correspondence analysis (D).

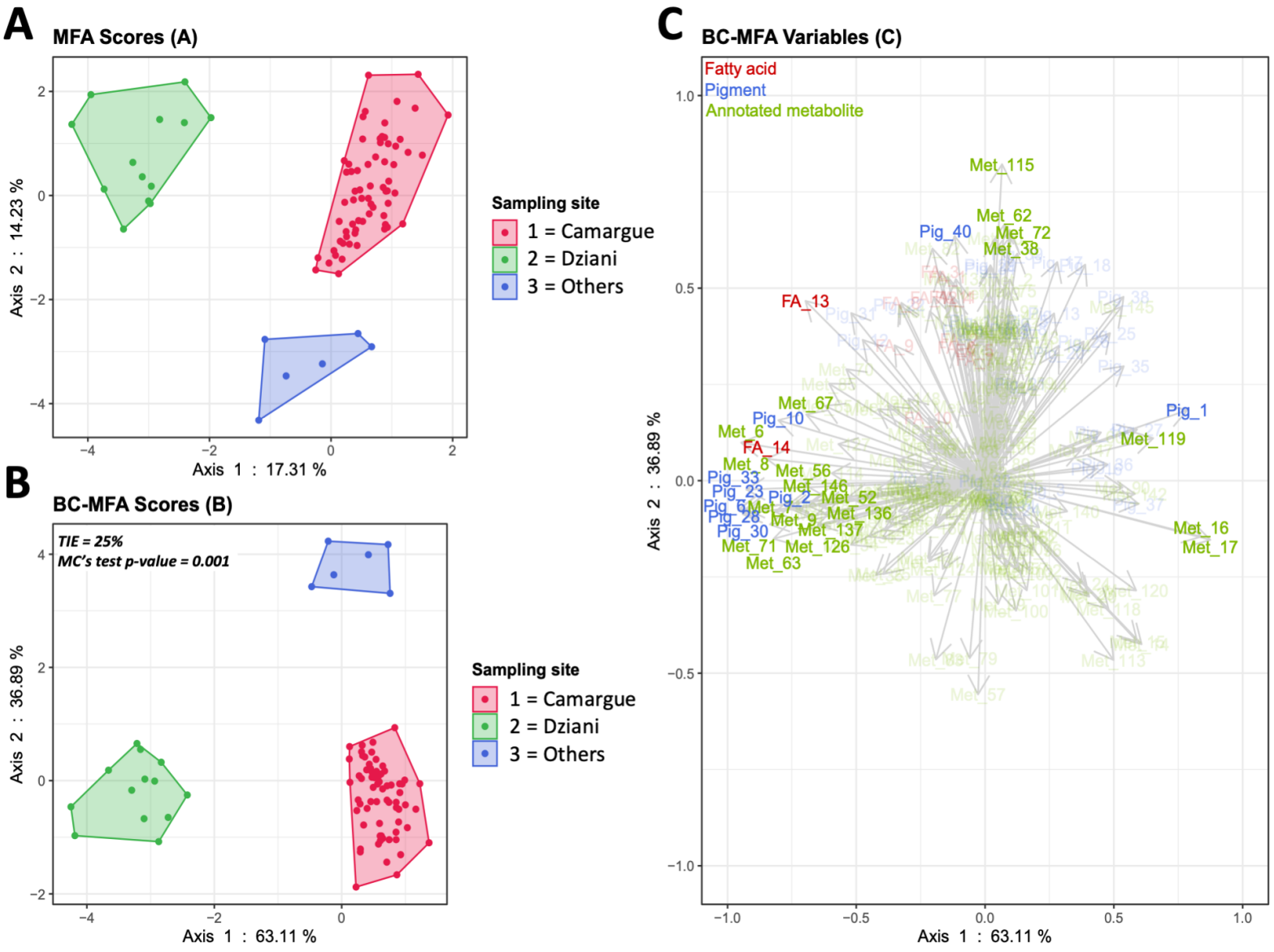
Figure 3. Multiple factor analysis (MFA: A), and between class analysis with a cos^2 threshold of 0.4 (BCA: B, cos^2 plot: C) based on the chemical composition of the strains. The tested factor is the sampling site.

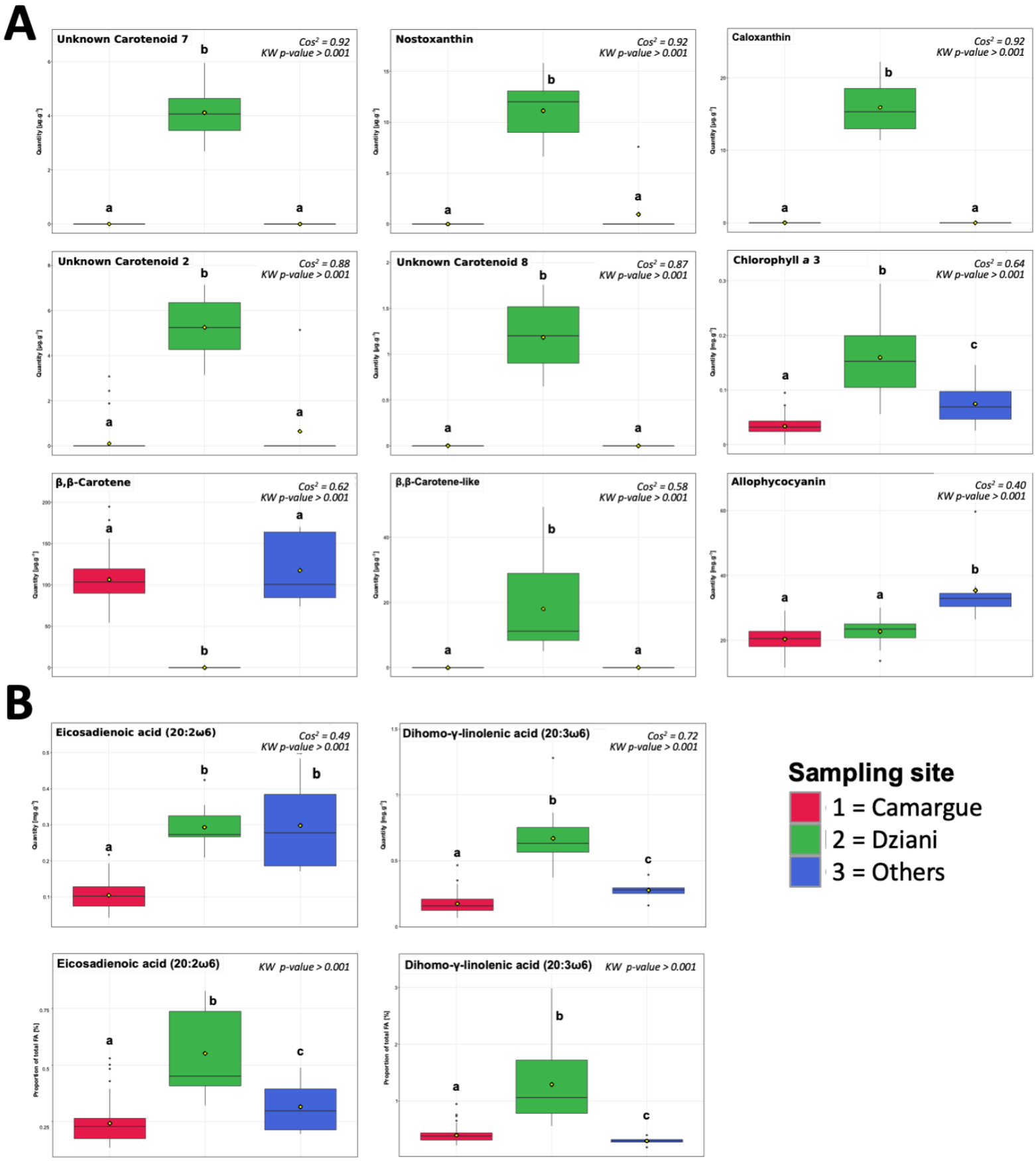
Figure 4. Most discriminant fatty acids and pigments ($cos^2 > 0.4$) were compared between sampling sites (Kruskal-Wallis test, pairwise Wilcoxon tests). Pigment concentrations are given in $mg.g^{-1}$ dw for phycobiliproteins and chlorophylls a, and $\mu g.g^{-1}$ dw for other pigments (A). Fatty acid concentrations are given in $mg.g^{-1}$ dw, and proportions in % of total fatty acids (B).

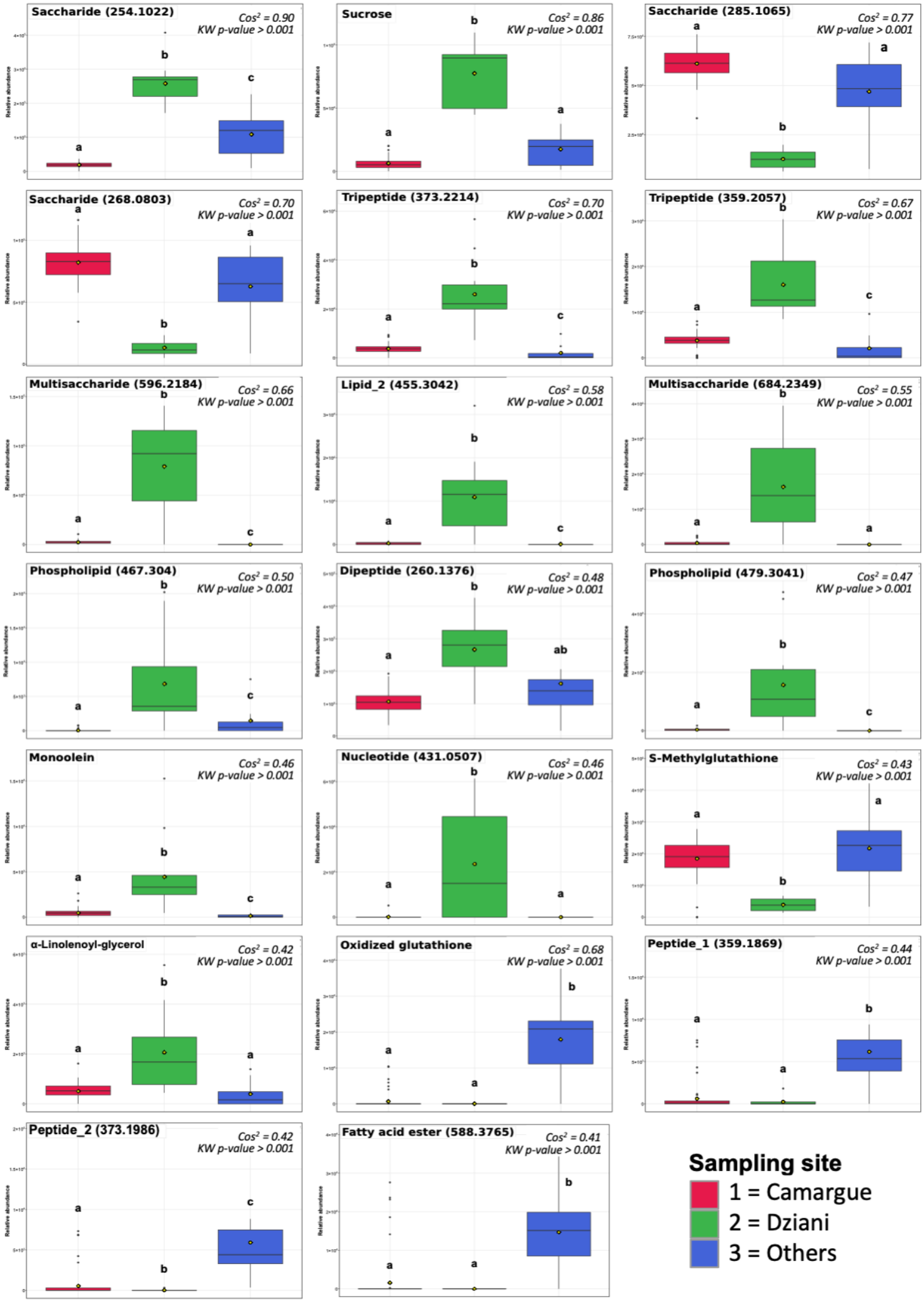
Figure 5. Relative abundances of most discriminant annotated ions ($cos^2 > 0.4$) were compared between sampling sites (Kruskal-Wallis and pairwise Wilcoxon tests).


A
Pigments
Sampling site
Salinity
Type of haibitat
Year of isolation
Flexible genes clade
β,β-carotene-like
Chlorophyll a3
Caloxanthin-like
UC7
UC2
Nostoxanthin
UC8
Chlorophyll a
β,ε-carotene-like
Oscillaxanthine
UC5
Myxoxanthophyll
Ucar1
C-phycocyanin
Allophycocyanin
Neoxanthin
UC1
Chlorophyll a1
Chlorophyll a2
Chlorophyll d-like1
Chlorophyll d-like2
β,ε-carotene-like
β,ζφ-carotene-like
Chlorophyll a2
Chlorophyll a2
UC3
UC4
UC6
Myxoxanthophyll-like1
Myxoxanthophyll-like2
UC9
Chlorophyll d-like3
UC10
Cantaxanthin
β,β-carotene-like
Zeaxanthin
Pheophytin a
Pheophytin a1
Chlorophyll a epi
Echinenone
5
0
-5
B
Lipids
Sampling site
Salinity
Type of habitat
Year of isolation
Flexible genes clade
1
2
14:0
15:0 iso
16:2 ω6
15:0
16:1 ω9
16:1 ω7
18:3 ω6
16:0
18:2 ω6
18:1 ω9
18:0
18:1 ω7
20:2 ω6
20:2 ω6
Pigment class
Phycobiliproteins
Chorophylls
Carotenoid - Carotenes
Carotenoid glycosides
Carotenoid - Xanthophylls
Carotenoid - Unknown
Sampling site
Camargue
Dziani
Others
Salinity [‰]
10–35
37–70
0
No data
pH
7-8
9–11
No data
Type of habitat
Brackish
Saline-alkaline
Fresh
Unk
Year of isolation
<2000
2010
2014
2018
2019
2020
Flexible genes clade
I.1
I.2
II
Unsequenced
C
Metabolites
Sampling site
Salinity
Type of habitat
Year of isolation
Flexible genes clade


A
ANOSIM p < 0.001
R = 0.90
HCPC Cluster
1 = Camargue
2 = Dziani
3 = Others
Axis 2 (14.23%)
Axis 1 (17.31%)
B
HCPC Cluster
1 = Camargue
2 = Dziani
3 = Others
C
Number of strains
Cluster 1
Cluster 2
Cluster 3
Sampling site
Camargue
Dziani
Others
pH
7-8
9-11
No data
Year of isolation
<2000
2010
2014
2018
2019
2020
Salinity [‰]
10-35
37-70
0-10
No data
Type of habitat
Brackish
Saline-alkaline
Freshwater
Unknown
Conding gene clade
I.1
I.2
II
Unsequenced
D
No data
No data
Clade II
No data
Freshwater
0‰
<2000
Other
Camargue
Clade I.1
10-35‰
pH 7-8
Brackish
2018
2020
2019
Clade I.2
Saline-alkaline
pH 9-11
Lake Dziani Dzaha
2014
2010
35-70‰
Factor
Sampling site
Salinity
pH
Habitat
Year of isolation
Genetic clade
Dim 2 (27%)
Dim 1 (29.1%)


A
MFA Scores (A)
Axis 2 : 14.23 %
Axis 1 : 17.31 %
Sampling site
1 = Camargue
2 = Dziani
3 = Others
B
BC-MFA Scores (B)
TIE = 25%
MC's test p-value = 0.001
Axis 2 : 36.89 %
Axis 1 : 63.11 %
Sampling site
1 = Camargue
2 = Dziani
3 = Others
C
BC-MFA Variables (C)
Fatty acid
Pigment
Annotated metabolite
Met_115
Met_62
Pig_40
Met_72
Met_38
FA_13
Met_67
Pig_10
Met_6
FA_14
Met_8
Pig_33
Met_56
Pig_23
Met_146
Pig_2
Met_52
Pig_6
Met_7
Met_136
Pig_28
Met_9
Pig_30
Met_137
Met_71
Met_126
Met_63
Pig_1
Met_119
Met_16
Met_17
Met_57
Axis 2 : 36.89 %
Axis 1 : 63.11 %


A
Unknown Carotenoid 7
Cos² = 0.92
KW p-value > 0.001
Nostoxanthin
Cos² = 0.92
KW p-value > 0.001
Caloxanthin
Cos² = 0.92
KW p-value > 0.001
Unknown Carotenoid 2
Cos² = 0.88
KW p-value > 0.001
Unknown Carotenoid 8
Cos² = 0.87
KW p-value > 0.001
Chlorophyll a 3
Cos² = 0.64
KW p-value > 0.001
β,β-Carotene
Cos² = 0.62
KW p-value > 0.001
β,β-Carotene-like
Cos² = 0.58
KW p-value > 0.001
Allophycocyanin
Cos² = 0.40
KW p-value > 0.001
B
Eicosadienoic acid (20:2ω6)
Cos² = 0.49
KW p-value > 0.001
Dihomo-γ-linolenic acid (20:3ω6)
Cos² = 0.72
KW p-value > 0.001
Eicosadienoic acid (20:2ω6)
KW p-value > 0.001
Dihomo-γ-linolenic acid (20:3ω6)
KW p-value > 0.001
Sampling site
1 = Camargue
2 = Dziani
3 = Others


Saccharide (254.1022)
Cos² = 0.90
KW p-value > 0.001
Sucrose
Cos² = 0.86
KW p-value > 0.001
Saccharide (285.1065)
Cos² = 0.77
KW p-value > 0.001
Saccharide (268.0803)
Cos² = 0.70
KW p-value > 0.001
Tripeptide (373.2214)
Cos² = 0.70
KW p-value > 0.001
Tripeptide (359.2057)
Cos² = 0.67
KW p-value > 0.001
Multisaccharide (596.2184)
Cos² = 0.66
KW p-value > 0.001
Lipid_2 (455.3042)
Cos² = 0.58
KW p-value > 0.001
Multisaccharide (684.2349)
Cos² = 0.55
KW p-value > 0.001
Phospholipid (467.304)
Cos² = 0.50
KW p-value > 0.001
Dipeptide (260.1376)
Cos² = 0.48
KW p-value > 0.001
Phospholipid (479.3041)
Cos² = 0.47
KW p-value > 0.001
Monoolein
Cos² = 0.46
KW p-value > 0.001
Nucleotide (431.0507)
Cos² = 0.46
KW p-value > 0.001
S-Methylglutathione
Cos² = 0.43
KW p-value > 0.001
α-Linolenoyl-glycerol
Cos² = 0.42
KW p-value > 0.001
Oxidized glutathione
Cos² = 0.68
KW p-value > 0.001
Peptide_1 (359.1869)
Cos² = 0.44
KW p-value > 0.001
Peptide_2 (373.1986)
Cos² = 0.42
KW p-value > 0.001
Fatty acid ester (588.3765)
Cos² = 0.41
KW p-value > 0.001
Relative abundance
Sampling site
1 = Camargue
2 = Dziani
3 = Others