



HAL
open science

Lipids in benthic diatoms: A new suitable screening procedure

Eva Cointet, Gaëtane Wielgosz-Collin, Vona Méléder, Olivier Gonçalves

► To cite this version:

Eva Cointet, Gaëtane Wielgosz-Collin, Vona Méléder, Olivier Gonçalves. Lipids in benthic diatoms: A new suitable screening procedure. *Algal Research - Biomass, Biofuels and Bioproducts*, 2019, 39, pp.101425. 10.1016/j.algal.2019.101425 . hal-02333689

HAL Id: hal-02333689

<https://hal.science/hal-02333689>

Submitted on 22 Oct 2021

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

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



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 Lipids in benthic diatoms: a new suitable screening procedure

2 Short title: Screening and rapid selection of lipid rich benthic diatoms

3 Eva Cointet¹, Gaëtane Wielgosz-Collin¹, Vona Méléder¹, Olivier Gonçalves²□

4 ¹ Université de Nantes, Laboratoire Mer Molécules Santé, EA 21 60, BP 92208, 44322 Nantes,
5 France

6 ² Université de Nantes, GEPEA, UMR CNRS-6144, Bât.CRTT, 37 Boulevard de l'Université,
7 BP406, F-44602 Saint-Nazaire Cedex, France

8 Abstract

9 The selection of suitable and indigenous microalgae species is a fundamental requirement
10 in developing added-value bioactive compounds recoverable in the food, health, and cosmetics
11 markets. In this work, an integrated screening approach was developed to characterize the lipid
12 rate of 33 diatom species (including 15 species studied for the first time) belonging to 16 genera
13 from the Nantes Culture Collection, with the main objective of discovering bioactive lipid
14 producers. For that purpose, a simple reliable method for establishing growth kinetics of strains
15 and semi-quantitative analysis of lipid rates was developed. Growth kinetics measurements
16 were achieved by daily minimal measurement fluorescence (F0) whereas lipid rate analyses
17 were performed by high-throughput Fourier Transform Infrared spectroscopy on entire cells
18 and lipid extracts. Results indicated that the method could be used directly on entire cells in
19 spite of the presence of silica for the FTIR approach (due to frustule). The total lipid rate was
20 species-dependant and ranged from 3.7% to 30.5% DW. Six strains out of 33 were found to
21 present a higher total lipid rate superior to 15% DW, and 11 showed medium lipid rates ranging
22 from 10% to 15% DW. The results revealed that five diatom species i.e. *Amphora* sp. NCC169,
23 *Nitzschia* sp. NCC109, *Nitzschia alexandrina* NCC33, *Opephora* sp NCC366 and *Staurosira*
24 sp. NCC182 presented interesting growth capabilities and should be further investigated as
25 potential sources for their original lipid rate.

26

27

28 **Keywords:** FTIR spectroscopy, growth kinetics, benthic Diatoms, lipid rate, bioactive fatty

29 acids.

30

31 **Highlights**

32 - Minimal fluorescence measurement characterized growth capabilities

33 - FTIR spectroscopy on whole cells characterized lipid rates

34 - Lipid rates of 15 strains were highlighted for the first time

35 - Five strains of benthic diatoms were selected for further biotechnology applications

36

37 *Corresponding authors

38 ^[L]_[SEP]Email: olivier.goncalves@univ-nantes.fr (Olivier Gonçalves), eva.cointet@univ-
39 nantes.fr (Eva Cointet), vona.meleder@univ-nantes.fr (Vona Méléder), Gaetane.Wielgosz-
40 Collin@univ-nantes.fr (Gaëtane Wielgosz-Collin).

41 ^[L]_[SEP]Preprint submitted to Algal Research September 09, 2018

42

43 **Conflict of interest**

44 The authors declare that there is no conflict of interest. No conflicts, informed consent, human
45 or animal rights are applicable to this work.

46

47 **Author's contribution**

48 Eva Cointet, Gaëtane Wielgosz-Collin, Vona Méléder and Olivier Gonçalves designed and
49 supervised the research. Eva Cointet, Gaëtane Wielgosz-Collin, Vona Méléder and Olivier
50 Gonçalves conducted experiments. Eva Cointet, Gaëtane Wielgosz-Collin, Vona Méléder and

51 Olivier Gonçalves analyzed and interpreted the data and drafted the manuscript. Eva Cointet,
52 Gaëtane Wielgosz-Collin, Vona Méléder and Olivier Gonçalves critically reviewed the
53 manuscript. All authors read and approved the final version of the manuscript.
54

55 Acknowledgements

56 This work was supported by the regional Atlantic Microalgae research program (AMI), funded
57 by the Pays de la Loire region. We also express our sincere thanks to GEPEA staff in particular
58 Remy Coat and Delphine Kucma for support and advice on the FTIR spectrometer.

59 Abbreviation

60 ATR: Attenuated Total Reflectance

61 AW: Ash weight

62 DW: Dry weight

63 DHA: Docosahexaenoic acid

64 EPA: Eicosapentaenoic acid

65 FA: Fatty acid

66 F0: Minimum chlorophyll fluorescence

67 HTSXT-FTIR: Fourier-transform infrared spectroscopy high-throughput screening extension

68 LED: Late exponential day

69 LED biomass: Biomass corresponding to culture harvest days

70 MUFA: Monounsaturated fatty acid

71 NCC: Nantes Culture Collection

72 NDVI: Normalized Difference Vegetation Index

73 N: Nitrogen

74 PFD: Photon flux density

75 PAM: Pulse amplitude modulation

76 PAR: Photosynthetically active radiation

77 PUFA: Polyunsaturated fatty acid

78 ρ : Reflectance

79 SD : Standard deviation

80 Si : Silica

81 SFA: Saturated fatty acid

82 TAG: Triacylglycerol

83 μ_{\max} : Maximum growth rate

84

85 1 Introduction

86 In the last two decades, a large body of research has focused on finding new strains of
87 microalgae capable of producing high lipid content for a wide range of applications including
88 pharmaceutical, cosmetics and alternative biofuels [1–4]. In the kingdom of microalgae,
89 diatoms are very accessible resources, since they are ubiquitously found in most aquatic
90 environments (rivers, oceans, coastal areas). They constitute a unicellular eukaryotic group with
91 a typical species-specific siliceous cell wall (also known as frustule). They present different
92 life-forms that could be benthic (microphytobenthos) or planktonic (phytoplankton). Marine
93 diatoms can grow quickly and store large amounts of lipids [5]. Their lipids are mainly
94 composed of a neutral fraction with traces of sterols and polar lipids [6]. Neutral lipids
95 constitute the reserve fraction, with triacylglycerol (TAG) accounting for more than 60% of the
96 total lipids [7]. Their PUFAs are mainly composed of eicosapentaenoic acid (EPA, C20:5 n-3)
97 [8] but some strains were also found to present docosahexaenoic acid (DHA, C22:6 n – 3) [9].
98 The biosynthesis of the lipids varies within the different diatom species, their growth stages,
99 and environmental parameters [10,11]. Previous studies [5–9] have demonstrated their ability
100 for lipid production, more specifically for the PUFA fraction (DHA and EPA), recognized for
101 its broad spectrum bioactivities (anti-carcinogenic, immune modulation, anti-diabetic, anti-
102 obesity and anti-thrombotic properties) [12]. Unfortunately, the Diatoms group is poorly
103 studied and constitutes therefore an underexploited resource [13] even if the number of their
104 genera and species is estimated to be between 250 and 100,000 [14]. Bioprospecting efforts
105 should therefore be encouraged in order to assess this potential. Basically, bioprospecting would
106 be achieved if the identified microalgae could be exploited at an industrial scale for their
107 biomass or their high value lipid compounds [1]. Therefore, during the screening approach,
108 specific focus should be performed on the efficient identification of the appropriate microalgae
109 strains, i.e. those characterized by high productivity (biomass and lipids), high resistance to

110 contamination and high tolerance to a wide range of environmental parameters [15–18]. Native
111 species adapt to local environmental changes and should be thus resilient and competitive
112 enough regarding these criteria. However, systematic estimation of the growth rates requires a
113 time-consuming series of measurements in order to estimate biomass evolution. The cell count
114 approach is among the most widely used. Alternatively, other parameters could be measured as
115 proxy for the cell numbers, if it could be shown to be linearly correlated. Typical proxy
116 measurements are *in vivo* fluorescence [19], optical density and biomass direct estimation (such
117 as dry weight, pigment content) [20] and *in vivo* or *in vitro* spectroscopy [21]. The
118 concentrations of protein, carbohydrates and lipids in cultures could also be used as proxy
119 measures depending of the robustness of their linear relationship with either cell numbers or
120 biomass. Finally, the cell number is recognized as being a robust reference method if “counting
121 methods” are easily available and could be applied to a given biological system. Only *in vivo*
122 fluorescence and spectroscopy seem to be faster than cell counting and are easy to use. For
123 these reasons, one of the objectives of this study is to test both techniques as an alternative to
124 counting in the screening process. Conventional methods used for lipid determination,
125 systematically require solvent extraction and gravimetric determination. These methods are
126 time consuming, need extensive manipulations, use high amounts of biomass (10-15 mg [22])
127 and have a low throughput screening rates. Consequently, a faster measurement of the lipid
128 content is needed [23]. As for counting methods, alternative approaches exist, that can be
129 roughly classified into invasive and non-invasive techniques. Fluorescent based technologies
130 are the most commonly used and require fluorescent dyes like Bodipy. However, it is an indirect
131 measure that has several issues such as sample preparation before staining, careful choice of
132 dye especially in microalgae due to the presence of chlorophylls within chloroplasts, leading to
133 non-quantitative information [24]. Vibrational spectroscopy such as Raman [25] could be a
134 good alternative for investigating lipid content since it can image and chemically identify the

135 lipids without labelling. However, autofluorescence signals from the chloroplast will
136 overwhelm the Raman signal. Increasing the acquisition time or incident power can solve this
137 issue, although irreversible photo damage may result, leading to loss of the semi-quantitative
138 information, limiting therefore its use in the context of high-throughput screening. Infrared
139 spectroscopy has advantages over the above techniques, since it limits photodamage, is not
140 influenced by autofluorescence, and presents robust systems such as high-throughput screening
141 for chemical spectra acquisition for large quantities of samples. Coat *et al.* demonstrated [26]
142 that it was robust and sensitive enough to quantify the lipids and that directly on entire
143 microalgae cells. The second objective of this study was to test this technique as an alternative
144 to gravimetric determination of lipid content.

145 In the present work, we propose to explore the potential for original lipid sourcing of the benthic
146 diatoms hosted in the Nantes Culture Collection (NCC) bank. First – to identify the NCC strains
147 that could potentially produce interesting fatty acids – a bibliographic inventory of current and
148 past knowledge of the principal genus and species hosted in the NCC bank was conducted. It
149 focused on the families of molecules with high added value for pharmacology, health, nutrition,
150 and cosmetics such as PUFAs, taking also into account the influence of cultural conditions
151 [27,28]. The selected species were then, investigated to gain basic knowledge on their biological
152 characteristics, to highlight their chemo diversity (in terms of protein, carbohydrates, silica and
153 lipid rates), and finally on their potential for high-value fatty acid molecule biosynthesis. To
154 select the most promising NCC strains, an original workflow was developed, integrating
155 different steps including the analysis of the NCC strain growth and their rapid biochemical
156 profiling through HTSXT-FTIR. The results obtained in the present study are discussed
157 hereafter.

158 2 Material and methods

159 2.1 Strains cultivation

160 Diatom strains from the NCC were selected and cultivated in 250 mL Erlenmeyer flasks filled
161 with 150 mL of F/2 culture medium, enriched in silica using 0.47 μm filtered natural sea water
162 [29, 30]. Vitamins and carbonates were added before autoclave sterilization (Autoclave Vertical
163 Lequeux AUV 100L), salinity was adjusted at 28 and pH was fixed at 7.8 to reduce nutrient
164 precipitation. Inoculation of each strain was performed from the stock cultures with an initial
165 concentration of 30 000 cells.mL⁻¹. The diatom strains were thus grown at 16°C with a photon
166 flux density (PFD) of 127 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ provided under continuous light to the bottom
167 of flasks by flat led pannel (LP EPURE-Chateaugiron, France).

168 2.2 Growth parameter estimations

169 To retrieve the growth parameters on the screened strains, the growth rate (μ_{max} in day⁻¹) and
170 the end of the exponential phase were identified. The growth rate identifies the faster growing
171 strains. The end of the exponential phase, known to be the period of lipid accumulation [31],
172 identifies the harvest day for lipid analyses. These parameters were retrieved using daily
173 biomass estimation. All erlen flasks and stir bars were autoclaved. Before sampling, flasks were
174 agitated by magnetic stirring for 2 minutes allowing cell and nutrient homogenization and
175 aggregate destruction.

176 Alternative techniques were tested to replace the cell counting by hemocytometer (here
177 Neubauer; $n \geq 300$) which is a time-consuming technique.,

178 - Minimum chlorophyll fluorescence measurement (F0) by fluorometry PAM (Water-PAM,
179 Waltz, Germany). This parameter, proportional to Chl *a* content, was used as a proxy of the
180 vegetal biomass [32]. Measurements were made directly on the microalgal suspension.

181 - Reflectance (ρ) of the cells by spectroradiometry (Jaz, Ocean Optics, USA) in the PAR domain
 182 (400-700 nm). Reflectance values were used to calculate the Normalized Difference Vegetation
 183 Index (NDVI) following the equation. (1), known to be proportional to Chl *a* content and used
 184 as a biomass proxy [21]. Measurements were made on filtered microalgae using CML
 185 microfiber filters with a 25 mm diameter and a 0.7 μm pore diameter and a 25mm Whatman
 186 filter funnel.

$$187 \quad NDVI = \frac{\lambda_{750} - \lambda_{675}}{\lambda_{750} + \lambda_{675}} \quad \text{eq. (1)}$$

188 With λ_{750} : maximal reflectance wavelength, and λ_{675} : Chl *a* absorption wavelength

189

190 Daily measurements were performed on each sample of culture to follow the growth kinetics
 191 (at the beginning of the growth phase and during three days, between 3 mL and 5 mL of cultures
 192 were sampled, and then 2 mL until the end of the growth). To extract and compare the growth
 193 parameters from the alternative techniques and the cell count approach, the Gompertz model
 194 [33] was used to fit the growth data (equation. 2) using MATLAB software. It consisted of a
 195 latency phase followed by an exponential phase and a stabilization of the curve at its maximum
 196 phase.

$$197 \quad f(x) = A \times e^{-e(\mu_{max} \times \frac{e^1}{A} \times (\lambda - x) + 1)} \quad \text{eq. (2)}$$

198 With A : maximum cell concentration in the natural logarithm of the biomass (number of cells.
 199 mL^{-1} , F0 or NDVI); μ_{max} : Maximum growth rate (day^{-1}); λ : Latency (days).

200 The A parameter allows calculation of the biomass ($LED_{biomass}$) obtained at the late exponential
 201 day (LED) (equation. 3). This corresponded to the harvest day for the lipid analyses.

$$202 \quad LED_{biomass} = \exp(A + \log(Bmin)) \quad \text{eq. (3)}$$

203 With $Bmin$ = Initial biomass (number of cells. mL^{-1} , F0 or NDVI); $LED_{biomass}$ = biomass of the
 204 late exponential day (cell. mL^{-1} , F0 or NDVI)

205

206 The comparison of the three techniques, and the selection of the fastest one to estimate the
207 growth parameters was done using a panel of 6 very different strains, chosen for their different
208 growth kinetics and their different morphology (aggregate of cells, cells in chains or solitary
209 cells) : *Amphora* sp. 1 NCC260, *Entomoneis* sp. 1 NCC350, *Entomoneis* sp. 6 NCC335,
210 *Entomoneis paludosa* NCC18.2, *Extubocellulus cribriger* NCC229 and *Navicula* sp. 2
211 NCC226.

212 2.3 Diatom strain characterisation

213 At the end of the exponential phase (LED), when the cells were harvested for HTSXT-FTIR
214 analyses (see §2.4), several analysis were performed to collect more information about the
215 strains:

216 - the cell length and the width were estimated using the image of light microscopy (OLYMPUS
217 CH40, ×400; n=150).

218 - the dry weight (DW) of the biomass was estimated by filtering 150 mL of algal suspension
219 through a microfiber filter, Whatman 47 mm diameter, 0.7 μm pore. The filters containing the
220 cells were washed using 10 mL of ammonium formate (3%) to remove salt. The wet filters
221 were frozen at -80°C and freeze-dried under a vacuum. DW (g.L⁻¹) and μ_{max} (day⁻¹) were then
222 used to estimate the strain productivity (Px) in g.L⁻¹.day⁻¹ (equation. 4).

$$223 \quad Px = \mu_{max} \times DW \quad \text{eq. (4)}$$

224 - the strains total lipid rate was assessed by gravimetric assay, to compare it to the infrared semi-
225 quantitative measurements (HTSXT-FTIR, see § 2.4). Biomass filtered, washed and freeze-dried
226 for dry weight estimation were used for lipid content estimation. The filters were macerated in
227 flasks using 100 mL of solvent per gram of biomass (dichloromethane-methanol (1:1 V/V))
228 [34]. Maceration at ambient temperature was performed for 24H on a vibrating tray (Edmund
229 Bühler GmbH, SM-30). After maceration, the mixture was filtered on pleated filters, 190 mm
230 diameter, 10μm pore, to remove the filter debris and the silica fragments. The filtrates were

231 transferred into a separatory funnel with 20 mL of distilled water and shaken for 5 min. The
232 lipid fraction (organic phase) was then separated from the separatory funnel, dried using an
233 anhydrous sodium sulfate salt, filtered, evaporated and weighed to obtain the crude lipid extract
234 (CLE) value. Total lipid rate (TLR) was finally expressed in % of the DW (equation. 5).

$$235 \quad TLR = \frac{CLE}{DW} \times 100 \quad \text{eq. (5)}$$

236 - the silica content of the cells was also determined to ensure the normalization of the FTIR
237 semi-quantitative results (see § 2.4). Cultures were harvested by filtering new 150 mL of the
238 algal suspension using the same recovery cell procedure used for dry weight estimation. Filters
239 were then freeze-dried, weighed (DW) and heated at 400°C for four hours in a muffle oven and
240 weighed (AW) and the silica proportion evaluated in % of DW by equation. 6 :

$$241 \quad \textit{Silica content} = \frac{AW}{DW} \times 100 \quad \text{eq. (6)}$$

242 2.4 Molecular profiles measured by infrared spectroscopy

243 The FTIR spectra acquisition on entire cells was performed according to Coat et al. 2014
244 recommendations [26]. It consisted in concentrating the cells up to $10^6 - 10^8$ cells.mL⁻¹ by
245 centrifugation (10 000g for 5 min) using Sigma 3K30 centrifuge. The supernatant was removed,
246 and the pellet resuspended in an ammonium formiate isotonic solution (68 g.L⁻¹). This process
247 was repeated twice to wash out the cells from the growth medium in order to avoid medium
248 contribution on the FTIR spectra. The cells were thus resuspended in 1mL in an isotonic
249 solution. Ammonium formiate solution prevents cell lyses during washing [35]. For the FTIR
250 spectra acquisition, a Bruker tensor 27 FTIR spectrometer equipped with a HTSXT plate reader
251 module coupled to OpusLab v 7.0.122 software (Bruker Optics, Germany) was used. Rinsed
252 cell aliquots of 5 µL were deposited on a 384 well microplate, void-dried in a vacuum desiccator
253 for at least 24 hours. FTIR spectra were then recorded in transmission mode directly on the
254 microplate loaded with the dried samples. This method was chosen since it is fast and non-

255 invasive on intact diatom cells, their biochemical signatures expressed in term of total lipids,
 256 total proteins and total carbohydrates superimposed partially with the silica signal of the diatom
 257 frustul. To confirm the whole cells FTIR spectra results, estimation of the lipid rate was also
 258 performed on crude lipid extract, using attenuated total reflectance sampling system (ATR),
 259 that was more suitable for crude organic extracts. For the acquisition of the ATR spectra, 10 μ L
 260 of crude lipid extract (see §2.3. [34]) were deposited directly on the Bruker tensor 27 FTIR
 261 spectrometer lens. The absorbance spectra were all collected between 4 000 cm^{-1} and 700 cm^{-1}
 262 with 30 scans and averaged. The spectra were analyzed by relatively straightforward methods
 263 such as peak ratios or integral ratios [36]. The lipid signature was associated to the $\text{CH}_2\text{-CH}_3$
 264 signal ($\sim 3000 - 2800 \text{ cm}^{-1}$) and the ester bond (Eb) signal ($\sim 1740 \text{ cm}^{-1}$). The carbohydrate
 265 signature was associated to the C-O-C signal of the polysaccharides ($\sim 1200\text{-}900 \text{ cm}^{-1}$) [37]. The
 266 protein signature was associated to the amide II band ($\sim 1540 \text{ cm}^{-1}$) of the N-H of the amides
 267 associated to the proteins. The silica signature was associated to the Si-O signal of the silicate
 268 frustule ($\sim 1068 \text{ cm}^{-1}$) [38]. To estimate the relative content of the lipids, carbohydrates and
 269 proteins, their respective peak heights (i.e. ester bond + ($\text{CH}_2 + \text{CH}_3$), $\sim 1159 \text{ cm}^{-1}$ and amide II)
 270 were standardized to the silica peak [39]. (equation 7). For the crude lipid extract spectra, the
 271 ratio used for that purpose was the ester bond and the $\text{CH}_2 + \text{CH}_3$ signals standardized with the
 272 total spectrum area (equation 8). FTIR and ATR ratio were expressed in arbitrary units
 273 abundance (a.u).

$$274 \quad \text{FTIR} = \frac{\text{Peak height (S)}}{\text{Silica peak high}} \quad \text{eq. (7)}$$

275 $S = \text{lipids (eb+(CH}_2\text{+CH}_3\text{)) or carbohydrates (}\sim 1159 \text{ cm}^{-1}\text{) or amide II (}\sim 1540 \text{ cm}^{-1}\text{)}$

$$276 \quad \text{ATR} = \frac{\text{Peak area (eb+CH}_2\text{+CH}_3\text{)}}{\text{Total spectra area}} \quad \text{eq. (8)}$$

277 2.5 Data processing

278 The Pearson product-moment correlation was carried out to test the positive correlation between
279 the growth curves obtained with the cell count (Cells.mL⁻¹), fluorescence (F0) and reflectance
280 (NDVI).

281 Comparison of the growth rate and LED estimated from the three techniques for the selected
282 species were performed using ANOVA when the data presented normal distribution or the
283 Kruskal – Wallis test when the data distribution was not normal. It was systematically followed
284 by the Tuckey post hoc test. The ANOVA was also performed on the growth rate (μ_{max}),
285 production (Px) and lipid rate (TLR) results to identify the strains with the highest performance.

286 A multivariate correspondence analysis was performed on the FTIR ratio normalized by silica
287 for lipids, protein and carbohydrates to assess the dispersion of the biochemical information
288 across the screened species and to identify if strains could be classified according to this
289 information. This method was chosen since it analyzes binary, ordinal and nominal data without
290 distributional assumptions (unlike traditional multivariate techniques) and also to preserve the
291 categorical nature of the variables. The correspondence analysis provided a unique graphical
292 display showing how the variable response categories were related [40].

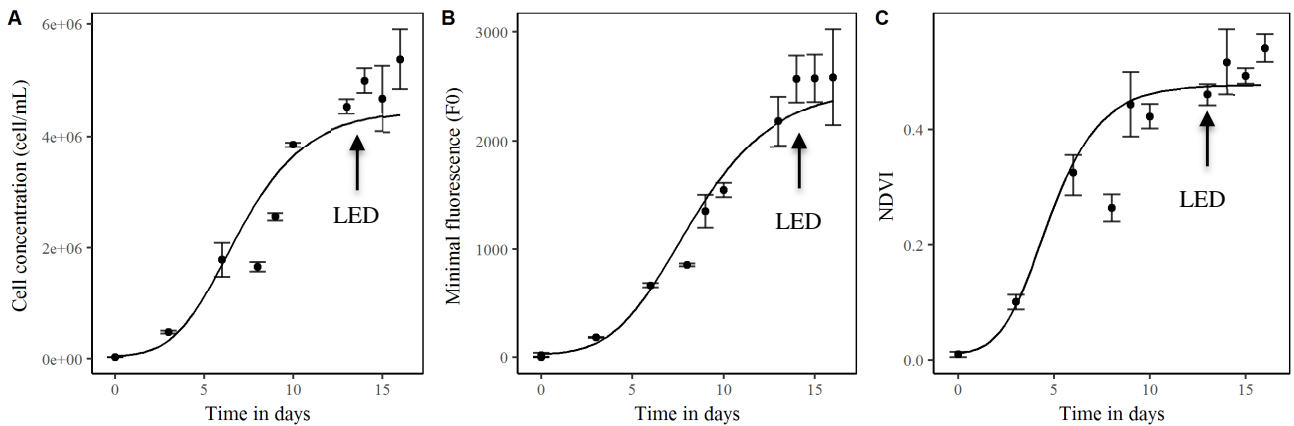
293 The Pearson product-moment correlation was used to test the correlation between the
294 calculation methods using the FTIR spectra semi-quantitative information with the total lipid
295 quantification reference method (the gravimetric approach). The Pearson product-moment
296 correlation and the comparison of the growth parameters were carried out using SigmaStat 3.1
297 software. The Past3 software was used for the correspondence analysis approach. All
298 experiments were performed in triplicate.

299 Finally, all the information and tests were combined and used for the validation of the screening
300 methodology to characterize the diatom strains of the NCC and the selection of candidates
301 presenting the best potential in terms of growth capabilities and lipid rates.

302 3 Results

303 3.1 Determination of the growth parameters

304 The latency, exponential and stationnary phases were observable whatever the technique used
305 to establish the growth curve for the 6 strains selected for their different kinetic behaviour (Fig.
306 1). Moreover, the cell count, NDVI and F0 data were all correlated (Tab.1) confirming the same
307 growth pattern. Cell count and NDVI were positively correlated ($p<0.001$) as cell count and F0
308 ($p<0.001$). R values for the NDVI varied from 0.70 to 0.90 and for F0 measurements from 0.60
309 to 0.93.



310 **Figure1.** Example of growth curves measured for *Amphora* sp 1 NCC260 by (A) cell count,
311 (B) fluorometry PAM (F0) and (C) radiometry (NDVI). Points corresponding to cell
312 concentration according to time. Line curves corresponding to the Gompertz model fitted to
313 cell concentration as a function of time. n=3, vertical bar = SD. The arrows indicate the late
314 exponential day estimated by the Gompertz model.

315

316

317 **Table 1.** Pearson product correlation between radiometry (NDVI) and cell count and
 318 fluorometry PAM (F0) and cell count. Correlation was significant ($p < 0.001$)

Species	NDVI	F0
<i>Amphora</i> sp. 1 NCC260	0.90	0.93
<i>Entomoneis paludosa</i> NCC18.2	0.86	0.82
<i>Entomoneis</i> sp. 1 NCC335	0.86	0.81
<i>Entomoneis</i> sp. 6 NCC350	0.80	0.60
<i>Extubocellulus cf cribriger</i> NCC229	0.70	0.90
<i>Navicula</i> sp. 2 NCC226	0.80	0.78

319
 320 There were no significant differences ($p > 0.05$) between the technique used to estimate the
 321 growth rate (μ_{max}) and the late exponential day (LED) for the six tested species, except for the
 322 μ_{max} value estimated for *Entomoneis paludosa* (Tab. 2). For example, *Amphora* sp. 1 reached
 323 the late exponential phase at the day c.a. 13 (LED, Fig. 1, Tab. 2) with a mean μ_{max} of c.a.
 324 0.81 (Tab. 2). At this time, the maximum biomass was reached by the cell count indicator with
 325 $4.5 \pm 0.2 \times 10^6$ cells.mL⁻¹ by PAM fluorometry with F0 values 2785 ± 609 , and by radiometry
 326 with NDVI reaching 0.47 ± 0.05 . Because the fluorimetry is used extensively for the
 327 measurement of extracted Chl *a*, for the estimation of the phytoplankton Chl *a in vivo* [19,41]
 328 and do not need filtration of high amounts of culture (contrary to NDVI), it was selected as an
 329 fast and reliable alternative approach to the cell count to determine the growth rates of all
 330 selected diatom strains in this study. Regarding *Entomoneis paludosa*, a new experiment using
 331 only PAM measurements, in the same culture conditions but without counting and NDVI
 332 estimation were conducted. μ_{max} values obtained were 0.52, 0.55, 0.55, leading to an average
 333 value of 0.54 ± 0.016 . A second statistical test performed on these new data concluded no
 334 significant difference for μ_{max} and LED ($p = 0.47$).

335 **Table 2.** Values of the growth parameters (maximum growth rate, μ_{max} (in d⁻¹) and late
 336 exponential day, LED (in days)) retrieved from the Gompertz model using six species employed
 337 for the comparison of the cell count approach with two alternative techniques: Fluorometry

338 PAM and radiometry. The calculated P value corresponded to the ANOVA test except for the
 339 (*) values that were obtained using the Kruskal-Wallis test. N=3, independent measurements,
 340 \pm SD.

Species	Techniques	μ max (day ⁻¹)	LED (day)	P value (μ max)	P Value (LED)
<i>Amphora</i> sp. 1 NCC260	Cell count	0.76 \pm 0.10	13 \pm 1	P=0.8	P=0.16*
	Pam	0.78 \pm 0.30	14 \pm 1		
	Spectroradiometer	0.88 \pm 0.25	12 \pm 2		
<i>Entomoneis</i> <i>paludosa</i> NCC18.2	Cell count	0.50 \pm 0.03	12 \pm 1	P \leq 0.05 ¹	P=0.07*
	Pam	0.95 \pm 0.23 ¹	10 \pm 0		
	Spectroradiometer	0.47 \pm 0.01	10 \pm 0		
<i>Entomoneis</i> sp. 6 NCC335	Cell count	0.53 \pm 0.10	14 \pm 1	P=0.27	P=0.14*
	Pam	0.54 \pm 0.12	12 \pm 1		
	Spectroradiometer	0.32 \pm 0.14	12 \pm 2		
<i>Entomoneis</i> sp. 1 NCC350	Cell count	0.24 \pm 0.01	16 \pm 2	P=0.29	P=0.09
	Pam	0.37 \pm 0.12	14 \pm 1		
	Spectroradiometer	0.35 \pm 0.06	14 \pm 1		
<i>Extubocellulus</i> <i>cf cribriger</i> NCC229	Cell count	0.68 \pm 0.04	14 \pm 1	P=0.10	P=0.14
	Pam	0.81 \pm 0.01	9 \pm 0		
	Spectroradiometer	0.83 \pm 0.07	10 \pm 3		
<i>Navicula</i> sp. 2 NCC226	Cell count	0.91 \pm 0.04	8 \pm 0	P=0.42*	P=0.78*
	Pam	1.02 \pm 0.23	8 \pm 2		
	Spectroradiometer	0.87 \pm 0.02	9 \pm 2		

341 ¹ *Entomoneis paludosa* μ max value for the second run was 0.54 and associated P value was 0.47

342 3.1.1 Diatom strain characteristics

343 Among the 68 screened strains, 33 were cultivated successfully and 36 did not grow
 344 (supplementary data 1). This information is reported in Tab. 3, (with the NCC reference, the
 345 sampling location, the cell size, the LED, the growth rate, the silica content, the DW and lipid
 346 rate). Globally the cell length varied from 58 \pm 14 μ m (*Craspedostauros* sp. 2 NCC218) to 4.8
 347 \pm 0.7 μ m (*Extubocellulus cf cribriger* NCC229). The cell width varied from 22.6 \pm 4.1 μ m
 348 (*Lithodesmium sp* NCC353) to 2.9 \pm 0.4 μ m (*Fallacia* sp. 1 NCC303). The end of the
 349 exponential phase, corresponding to the harvest day (LED), varied from day 6 for *Surirella* sp.
 350 1 NCC270 to day 16 for *Craspedostauros britannicus* NCC228 and *Navicula* sp. 1 NCC113
 351 (ANOVA, $p < 0.001$). Three groups were therefore identified depending on their growth rate:
 352 group A, included *Nitzschia alexandrina* NCC33 and *Entomoneis* sp. 5 NCC302 with
 353 respective growth rates of 1.89 \pm 0.10 and 1.50 \pm 0.18 day⁻¹, group B with 13 species (i.e 40%
 354 of the total number of strain) presented a growth rate ranging from 1.19 \pm 0.08 to 0.81 \pm 0.09

355 day⁻¹ and group C with 18 species (54%) had a growth rate below 0.8 day⁻¹ with a minimum at
356 0.22 ± 0.05 day⁻¹.

357 To ensure the standardization of the FTIR results for inter species comparison purposes, the
358 total silica rate was evaluated for the 33 species. It showed a significant difference (ANOVA,
359 $p < 0.01$), where five species (i.e. 15% of the total number of strains) had a silica content of c.a.
360 40% of the dry weight with a maximum of 46.11 ± 4.58% for *Entomoneis* sp 1 NCC350. Three
361 species (i.e. 9% of the total number of strains) presented a silica content of c.a. 23% of the dry
362 weight with a minimum of 22.9 ± 2.6% for *Conticriba weissflogii* CCMP1336. In the other
363 species, representing more than 75% of the total number of strains, the differences were not
364 significant (ANOVA, $p = 0.055$), supporting a stable silica content situated at around 35% DW.
365 The dry weight (DW) biomass values of the assessed strains ranged from 0.502 ± 0.039 g.L⁻¹
366 (*Craspedostauros britannicus* NCC195) to 0.065 ± 0.001 g.L⁻¹ (*Entomoneis* sp. 3 NCC351).
367 For three strains (9% of the total number of strains) the DW was greater than 0.30 g.L⁻¹. For 27
368 strains (81%) the biomass was greater than 0.10 g.L⁻¹ but lower than 0.30 g.L⁻¹. For the three
369 remaining strains (9%) the biomass was lower than 0.10 g.L⁻¹.

370 The lipid rate estimated with the gravimetric method ranged from 30.5 ± 0.7% DW (*Nitzschia*
371 sp. 5 NCC109) to 3.7 ± 1.1% DW (*Brockmaniella brockmanii* NCC161). For three strains (9%)
372 the lipid rate was greater than 20% DW. For 14 strains (42%) the lipid rate was greater than
373 10% DW. In the remaining 16 strains (48%) the lipid rate was lower than 10% DW.

374 Biomass productivity varied substantially among the tested strains (Fig. 2A) and ranged from
375 0.36 ± 0.02 g.L⁻¹.day⁻¹ (*Nitzschia alexandrina* NCC33) to 1.4 ± 0.3 × 10⁻² g.L⁻¹.day⁻¹
376 (*Entomoneis* sp. 3 NCC351). The results were found to be statistically different among those
377 strains (ANOVA, $p < 0.001$). Finally, the strains were clustered into three groups according to
378 the following parameters (Fig. 2): group 1 with productivity between 0.35 ± 0.02 and 0.19 ±
379 0.04 g.L⁻¹.day⁻¹; group 2 with productivity ranging from 0.17 ± 0.08 to 0.09 ± 0.02 g.L⁻¹.day⁻¹;

380 group 3 with lower productivity ranging from $7.5 \pm 0.4 \times 10^{-2}$ to $1.4 \pm 0.3 \times 10^{-2}$ (g.L⁻¹.day⁻¹).

381 The most productive strains were those exhibiting the highest μ_{max} associated to the highest

382 DW. But they did not correspond to the richest in terms of total lipid rate (Fig. 2B).

383

384

385

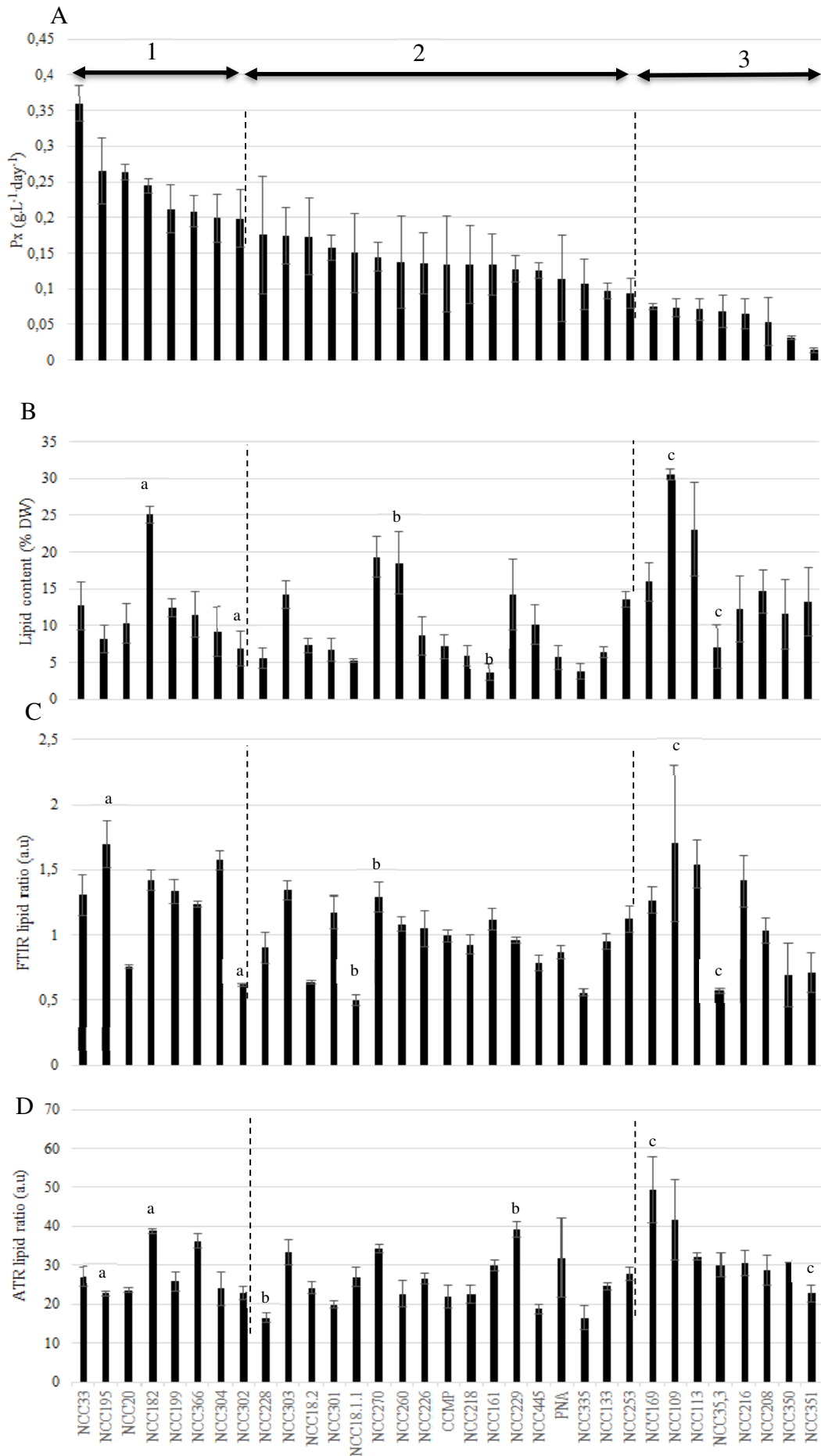
386
387Table 3. Characteristics and sampling locations of the investigated diatoms species. All data were obtained by experimental measurements. N=3, independent measurement, \pm SD.

Species	NCC strain identification	Sampling location	Cell size (μm)		LED (day)	μmax	Silica content (%)	Biomass ($\text{g}\cdot\text{L}^{-1}$)	TLR (% DW)	Ref
			Length	Width						
<i>Nitzschia alexandrina</i>	NCC33	France, NW Atlantic coast	11.0 \pm 1.2	3.7 \pm 0.4	11 \pm 0	1.89 \pm 0.10	27.2 \pm 3.8	0.189 \pm 0.004	12.7 \pm 3.2	[15,42–49]
<i>Entomoneis</i> sp. 5	NCC302	France, NW Atlantic coast	22.4 \pm 3.5	15.0 \pm 2.2	9 \pm 0	1.50 \pm 0.18	38.9 \pm 0.7	0.132 \pm 0.025	6.9 \pm 2.4	[50–54]
<i>Entomoneis</i> sp. 2	NCC20	France, NW Atlantic coast	30.5 \pm 4.6	18.5 \pm 3.2	8 \pm 0	1.19 \pm 0.08	34.3 \pm 2.8	0.244 \pm 0.042	10.3 \pm 2.7	[50–54]
<i>Staurosira</i> sp.	NCC182	France, NW Atlantic coast	6.9 \pm 1.4	5.5 \pm 1.2	8 \pm 2	1.17 \pm 0.09	28.9 \pm 5.1	0.262 \pm 0.084	25.1 \pm 1.2	[16]
<i>Fallacia</i> sp. 1	NCC303	Greenland, east coast	6.4 \pm 0.7	2.9 \pm 0.4	12 \pm 2	1.12 \pm 0.14	37.4 \pm 4.5	0.154 \pm 0.016	14.2 \pm 1.9	-
<i>Fallacia</i> sp. 2	NCC304	Greenland, east coast	6.4 \pm 0.8	3.1 \pm 0.4	12 \pm 0	1.07 \pm 0.15	31.8 \pm 4.9	0.265 \pm 0.028	9.2 \pm 3.3	-
<i>Entomoneis</i> sp. 7	NCC445	France, NW Atlantic coast	10.4 \pm 1.5	4.1 \pm 0.7	9 \pm 2	1.02 \pm 0.05	41.4 \pm 2.9	0.122 \pm 0.005	10.2 \pm 2.8	[50–54]
<i>Entomoneis paludosa</i>	NCC18.1.1	France, NW Atlantic coast	18.4 \pm 3.3	10.7 \pm 1.9	11 \pm 1	0.96 \pm 0.03	37.8 \pm 4.0	0.130 \pm 0.016	5.3 \pm 0.3	[50–54]
<i>Entomoneis</i> sp. 4	NCC301	France, NW Atlantic coast	23.9 \pm 3.4	17.3 \pm 2.2	7 \pm 1	0.92 \pm 0.14	26.7 \pm 3.9	0.171 \pm 0.012	6.7 \pm 1.5	[50–54]
<i>Pseudonitzschia americana</i>	PNA06 KER	France, NW Atlantic coast	6.6 \pm 0.8	4.5 \pm 0.5	9 \pm 2	0.89 \pm 0.09	36.4 \pm 0.6	0.099 \pm 0.019	5.7 \pm 1.6	-
<i>Surirella</i> sp. 1	NCC270	France, NW Atlantic coast	6.8 \pm 1.3	5.0 \pm 1.6	6 \pm 1	0.85 \pm 0.31	33.7 \pm 3.5	0.173 \pm 0.033	19.3 \pm 2.8	-
<i>Navicula</i> sp. 2	NCC226	France, NW Atlantic coast	9.1 \pm 1.3	5.6 \pm 0.7	8 \pm 2	0.82 \pm 0.05	41.4 \pm 1.5	0.131 \pm 0.015	8.7 \pm 2.5	[15,45,47,49,55,58]
<i>Opephora</i> sp. 1	NCC366	France, NW Atlantic coast	5.4 \pm 0.8	3.7 \pm 0.8	14 \pm 1	0.82 \pm 0.16	37.8 \pm 8.4	0.233 \pm 0.010	11.5 \pm 3.1	-
<i>Entomoneis paludosa</i>	NCC18.2.1	France, NW Atlantic coast	18.8 \pm 1.8	11.2 \pm 2.5	12 \pm 1	0.81 \pm 0.05	36.0 \pm 3.9	0.150 \pm 0.003	7.3 \pm 1.0	[50–54]
<i>Extubocellulus cf cribriger</i>	NCC229	France, NW Atlantic coast	4.8 \pm 0.7	3.9 \pm 0.5	9 \pm 0	0.81 \pm 0.09	33.7 \pm 2.3	0.155 \pm 0.003	14.2 \pm 4.8	[52,59,60]
<i>Amphora</i> sp. 1	NCC260	France, NW Atlantic coast	9.8 \pm 1.3	3.7 \pm 0.5	13 \pm 0	0.78 \pm 0.30	36.2 \pm 2.6	0.176 \pm 0.041	18.5 \pm 4.2	[49,55,58,61,62]
<i>Conticribea weissflogii</i>	CCMP1336	USA, NE Atlantic coast	15.6 \pm 2.5	10.8 \pm 1.4	12 \pm 0	0.59 \pm 0.15	22.9 \pm 2.6	0.277 \pm 0.035	7.2 \pm 1.7	-
<i>Brockmaniella brockmanii</i>	NCC161	France, NW Atlantic coast	9.2 \pm 2.4	4.3 \pm 0.6	10 \pm 0	0.58 \pm 0.01	37.5 \pm 0.2	0.185 \pm 0.025	3.7 \pm 1.1	-
<i>Craspedostauros britannicus</i>	NCC195	France, NW Atlantic coast	37.8 \pm 8.8	11.0 \pm 2.2	9 \pm 0	0.53 \pm 0.09	44.7 \pm 0.1	0.502 \pm 0.039	8.2 \pm 1.8	-
<i>Conticribea weissflogii</i>	NCC133	Morocco, Oum R'bia estuary	14.3 \pm 1.7	12.1 \pm 1.7	7 \pm 0	0.53 \pm 0.08	34.6 \pm 1.2	0.181 \pm 0.015	6.4 \pm 0.7	-
<i>Licmophora</i> sp. 1	NCC253	France, NW Atlantic coast	20.0 \pm 3.9	12.4 \pm 1.6	11 \pm 0	0.51 \pm 0.02	37.8 \pm 5.7	0.161 \pm 0.049	13.5 \pm 1.1	-
<i>Amphora</i> sp. 2	NCC169	France, NW Atlantic coast	9.0 \pm 1.2	5.8 \pm 1.2	13 \pm 3	0.48 \pm 0.14	33.8 \pm 4.8	0.157 \pm 0.024	16.0 \pm 2.6	[49,55,58,61,62]
<i>Craspedostauros</i> sp. 1	NCC228	France, NW Atlantic coast	55 \pm 14	14.9 \pm 4.2	16 \pm 4	0.46 \pm 0.22	25.2 \pm 0.7	0.384 \pm 0.042	5.6 \pm 1.4	-
<i>Craspedostauros</i> sp. 2	NCC218	France, NW Atlantic coast	58 \pm 14	13.6 \pm 3.4	12 \pm 0	0.45 \pm 0.18	26.0 \pm 4.1	0.178 \pm 0.047	5.9 \pm 1.4	-
<i>Cymatosira belgica</i>	NCC208	France, NW Atlantic coast	5.6 \pm 0.6	4.7 \pm 0.7	12 \pm 0	0.44 \pm 0.12	35.5 \pm 3.4	0.118 \pm 0.043	14.6 \pm 3.0	-
<i>Amphora acutiuscula</i>	NCC216	Viet Nam, South east coast	10.0 \pm 1.0	5.6 \pm 0.8	11 \pm 0	0.43 \pm 0.07	32.4 \pm 0.7	0.144 \pm 0.041	13.7 \pm 5.1	[49,55,58,61,62]
<i>Lithodesmium</i> sp	NCC353	Australia, Moreton bay	29.3 \pm 5.4	22.6 \pm 4.1	9 \pm 1	0.43 \pm 0.02	34.1 \pm 6.1	0.129 \pm 0.006	7.1 \pm 2.9	-
<i>Navicula</i> sp. 1	NCC113	Morocco, Oum R'bia estuary	17.1 \pm 2.9	5.1 \pm 1.2	16 \pm 0	0.42 \pm 0.02	25.0 \pm 1.0	0.168 \pm 0.035	23.1 \pm 6.4	[15,45,47,49,55,58]
<i>Entomoneis</i> sp. 6	NCC335	France, Mediterranean Sea	13.9 \pm 1.9	7.7 \pm 1.9	12 \pm 1	0.38 \pm 0.03	36.8 \pm 4.7	0.226 \pm 0.036	3.7 \pm 1.1	[50–54]
<i>Entomoneis</i> sp. 1	NCC350	France, Mediterranean Sea	29.0 \pm 4.0	17.5 \pm 2.7	14 \pm 0	0.36 \pm 0.13	46.1 \pm 4.6	0.098 \pm 0.048	8.8 \pm 0.1	[50–54]
<i>Craspedostauros britannicus</i>	NCC199	France, NW Atlantic coast	31.2 \pm 3.2	8.8 \pm 1.6	12 \pm 2	0.35 \pm 0.14	40.0 \pm 2.9	0.421 \pm 0.053	12.5 \pm 1.2	-
<i>Nitzschia</i> sp. 5	NCC109	France, NW Atlantic coast	33.1 \pm 4.1	10.1 \pm 1.5	12 \pm 1	0.31 \pm 0.05	23.0 \pm 1.5	0.239 \pm 0.026	30.5 \pm 0.7	[15,42–49]
<i>Entomoneis</i> sp. 3	NCC351	France, Mediterranean Sea	16.6 \pm 3.7	12.3 \pm 3.0	13 \pm 0	0.22 \pm 0.05	38.6 \pm 5.8	0.065 \pm 0.001	13.3 \pm 4.6	[50–54]

388

389

390



392 **Figure 2.** Values of the parameters measured for the screened strains includes strain
393 productivity (A), lipid rates as measured with the gravimetric method (B), lipid ratio measured
394 semi-quantitatively by the FTIR approaches, [(eb+CH₃+CH₃)/si] (C) and [area
395 eb+CH₂+CH₃]/Total area] multiplied by 100 for scaling purposes (D). Notations a, b, and c
396 correspond to the maximum and minimum values for groups 1, 2 and 3. N=3, independent
397 measurements, ± SD.

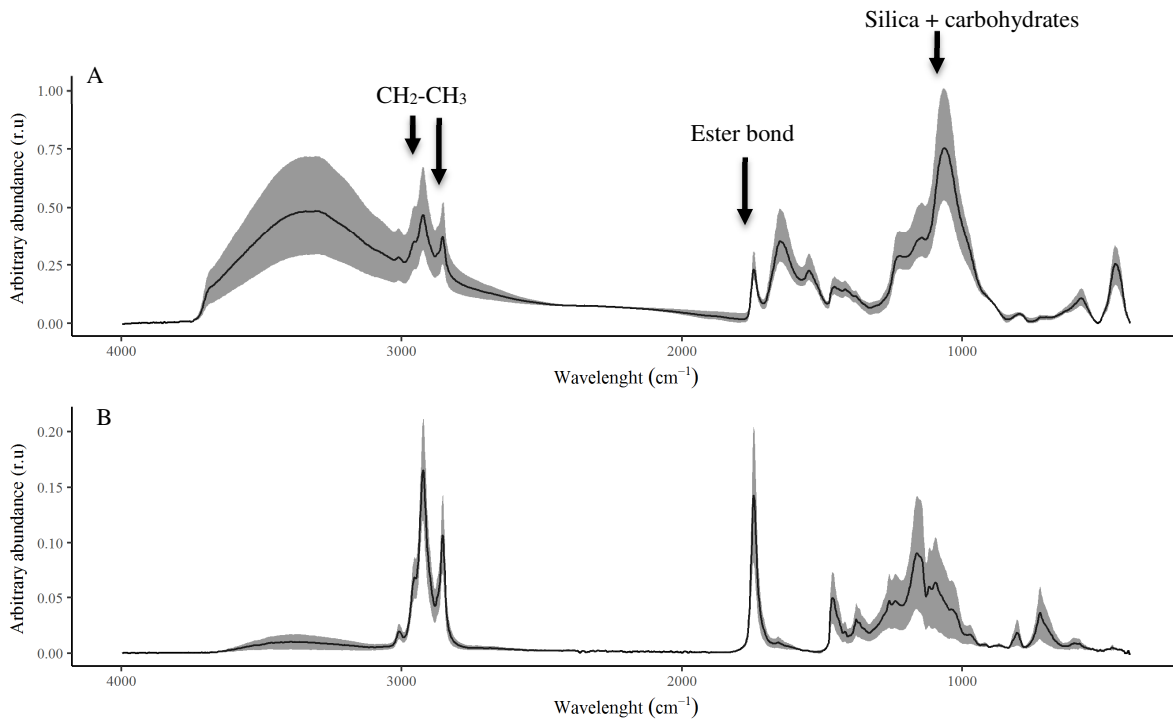
398 3.2 FTIR analysis

399 3.2.1 FTIR spectrum interpretation

400 HTSXT-FTIR analysis was performed on the 33 assayed species (Tab. 3). The lipid
401 rate was associated to three main signals on the recorded spectra (Fig. 3A), i.e. through the
402 two vibrations of the fatty acid carbon chains (ν CH₂ and ν CH₃) (ν C-H ~ 2923 and 2852 cm⁻¹
403 ¹) [26] and of the ester bond function (Eb) (ν C=O ~ 1750 cm⁻¹) [63]. The other major bands
404 corresponded to the principal cellular components such as the proteins (the amide I band ν
405 C=O ~1650 cm⁻¹; the amide II band δ N=H ~1540 cm⁻¹), the nucleic acids (ν P=O ~1230 cm⁻¹
406 ¹) and the carbohydrates band superimposed on the silica band (~900 – 1200 cm⁻¹). For details
407 see Wagner et al., 2010 [64]. Whereas the infrared signature obtained on the whole cells
408 showed superimposed bands of silica and carbohydrates at 1078 cm⁻¹ (Fig. 3A), the signature
409 obtained on the crude lipid extract (Fig. 3B) did not exhibit this band, but a well defined ester
410 bond (Eb) band at 1750 cm⁻¹ and well-defined bands for the CH₂-CH₃ signature at 3000-2800
411 cm⁻¹.

412 The lipid ratio estimated from the FTIR data measured on the entire cells (Fig.2C)
413 ranged from 1.70 ± 0.59 (*Nitzschia* sp. 5 NCC109) to 0.49 ± 0.04 (*Entomoneis paludosa*
414 NCC18.1.1). It was thus possible to cluster the assessed strains into three groups. Group 1
415 showed a maximum ratio at 1.69 ± 0.18 for *Craspedostauros britannicus* NCC195 and a
416 minimum at 0.61 ± 0.01 for *Entomoneis* sp 5 NCC302, group 2 showed a maximum ratio at
417 1.34 ± 0.07 for *Fallacia* sp 1 NCC303 and a minimum at 0.49 ± 0.04 for *Entomoneis paludosa*

418 NCC18.1.1 and group 3 showed a maximum ratio at 1.70 ± 0.59 for *Nitzschia* sp NCC109
419 and minimum at 0.56 ± 0.02 for *Lithodesmium* sp NCC35.3.



420

421 **Figure 3.** Example of averaged FTIR spectra recorded on entire cells or on the corresponding
422 lipid extract. (A) *Staurosira* sp NCC182 FTIR signature recorded on the entire cells and (B)
423 *Staurosira* sp NCC182 FTIR signature recorded on a crude lipid extract. The grey area
424 corresponds to the variation of the FTIR signal associated to the standard deviation for n=3
425 independent measurements.

426

427 The lipid ratio estimated from the FTIR data recorded on the lipid extracts (Fig.2D) ranged
428 from 49.3 ± 8.5 (*Amphora* sp. 2 NCC169) to 16.5 ± 3.1 (*Entomoneis* sp. 6 NCC335). Three
429 groups were also proposed regarding this criterion, group 1 showed a maximum ratio at 38.84
430 ± 0.63 for *Staurosira* sp. NCC182 and a minimum ratio at 22.8 ± 1.8 for *Entomoneis* sp. 5
431 NCC302, group 2 a maximum ratio at 39 ± 2 for *Extubocellulus* sp. NCC229 and a minimum
432 at 16.5 ± 1.2 for *Craspedostauros britannicus* NCC228 and group 3 with a maximum ratio at
433 49.3 ± 8.3 for *Amphora* sp. 2 NCC169 and a minimum at 23 ± 2 for *Entomoneis* sp. 3 NCC351.

434 The silica amount did not appear to significantly impact on the lipid FTIR signature except in
435 three species for which the FTIR lipid ratio fell between the two FTIR signature sampling
436 methods (i.e. on entire cells or on lipid extract). These were *Craspedostauros brittanicus*
437 NCC195, *Navicula* sp. 4 NCC113 and *Conticriba weissflogii* CCMP1336. For the other
438 species with different silica content, the lipid signature remained stable in both FTIR sampling
439 methods.

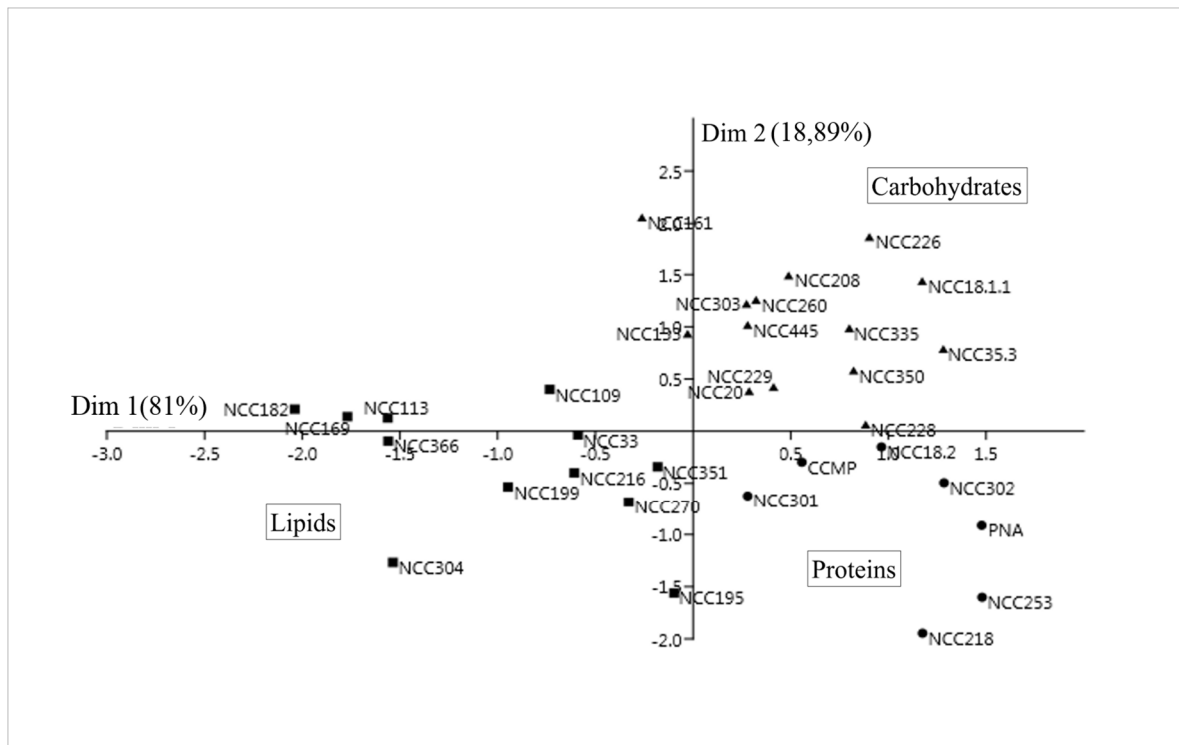
440

441 3.2.2 Multivariate analysis of the FTIR spectra recorded on entire cells

442 In order to assess the main differences in terms of biochemical composition among the 33
443 screened strains, a correspondence analysis approach using the lipid, protein and carbohydrate
444 bands normalized to the silica amounts was performed. The resulting map (Fig. 4) is a
445 classification of the data on two main dimensions. Dimension 1 represents 81% of the initial
446 information and could be associated to the variation of the lipid composition of the assessed
447 strains, ranging from the lowest to the highest amount of total lipids from right to left on the
448 map. Two strains, *Staurosira* sp. NCC182 and *Amphora* sp. 2 NCC169 presented the highest
449 amount in total lipids. Along dimension 2, representing 19% inertia, the strains associated to
450 that dimension were mainly opposed on the basis of their protein and carbohydrate content.
451 *Craspedostauros* sp. 2 NCC218 was rich in proteins, whereas in *Brockmaniella brockmanii*
452 NCC161 the main fraction was associated to the carbohydrates.

453 These results indicate that the 2 most notable differences or largest deviations in the sample
454 were observed first between *Staurosira* sp. NCC182, *Amphora* sp. 2 NCC169 and the other
455 species for their lipid rates, and secondly between the strains *Craspedostauros* sp. 2 NCC218,
456 *Brockmaniella brockmanii* NCC161 and the other species by their respective protein and
457 carbohydrate composition. This analysis summarizes the main biochemical characteristics of
458 the strains hosted in the NCC bank in a single step. Although the distance between the

459 macromolecular content and the species were not mathematically defined, their closeness on
 460 the map could be used as a guideline to interpret their biochemical characteristics: the squares
 461 correspond to the strains particularly rich in lipids, the triangles, the strains rich in
 462 carbohydrates and the dots, the strains particularly rich in proteins.
 463



464
 465 **Figure 4.** Correspondence analysis map calculated on the basis of the macromolecular content
 466 as evaluated by FTIR on all the assayed strains of the NCC. N=3 independent measurements.

467

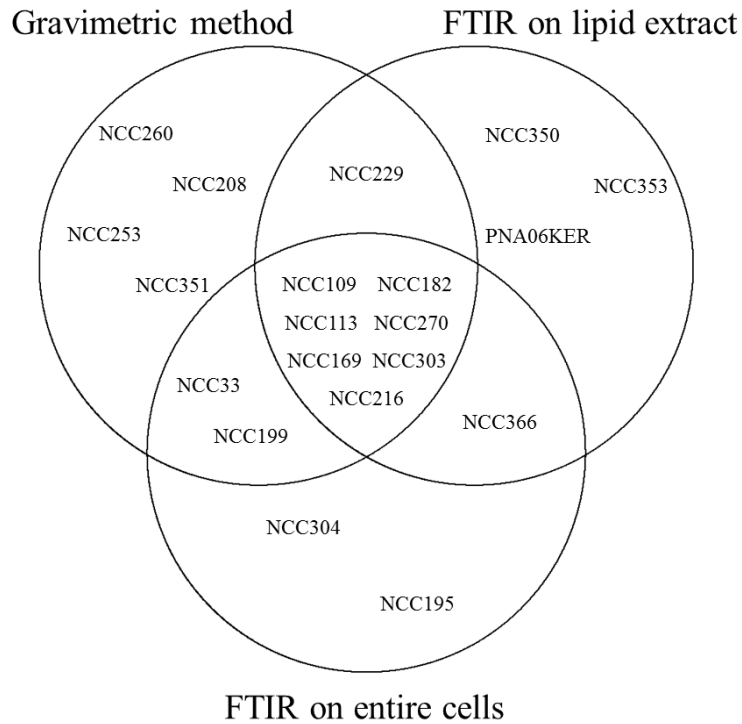
468 3.3 Comparison of the lipid amounts estimated by the gravimetric 469 method and the FTIR approaches.

470 A significant positive correlation was found for all the techniques (gravimetric and
 471 FTIR) ($p < 1.10^{-7}$) with a Pearson correlation score R superior to 0.53. For the 33 analyzed
 472 species, 14 presented a high lipid ratio ($> 12\%$ DW) using the gravimetric measurements:
 473 *Nitzschia alexandrina* NCC33, *Staurosira* sp. NCC182, *Fallacia* sp. 1 NCC303, *Surirella* sp.

474 1 NCC270, *Extubocellulus cf cribriger* NCC229, *Amphora* sp. 1 NCC260, *Licmophora* sp.1
475 NCC253, *Amphora* sp. 2 NCC169, *Cymatosira belgica* NCC208, *Amphora acutiuscula*
476 NCC216, *Navicula* sp. 1 NCC113, *Craspedostauros britannicus* NCC199, *Nitzschia* sp. 5
477 NCC109 and *Entomoneis* sp.3 NCC351 (Fig. 2B, Fig. 5). The FTIR method applied on entire
478 cells identified 12 species rich in lipids (FTIR lipid ratio > 1.20) with nine species identified
479 in common with the gravimetric method: NCC33, NCC182, NCC303, NCC270, NCC169,
480 NCC216, NCC113, NCC199 and NCC109 (Fig. 2C, Fig. 5). FTIR for the lipid extract
481 analyses identified 12 species rich in lipids (ATR lipid ratio > 30) with eight species identified
482 in common with the gravimetric approach: NCC182, NCC303, NCC270, NCC229, NCC169,
483 NCC216, NCC113 and NCC109 (Fig 2D, Fig. 5).

484 The correspondence analyses (Fig.4) performed on the FTIR profiles obtained in the entire
485 cells gave supplementary information and identified seven species particularly rich in lipids:
486 NCC182, NCC270, NCC169, NCC113, NCC366, NCC109 and NCC33. These species were
487 also identified by the gravimetric method with the exception of NCC366, only identified by
488 the FTIR method on crude lipid extract.

489



490

491 **Figure 5** Venn diagram showing the degree of overlap among the different approaches
 492 used to identify the lipid rich diatoms. In the gravimetric method circle, 14 strains were
 493 identified as rich in lipids: 4 only with this method (NCC260, NCC208, NCC253, NCC351),
 494 one strains were also identified as rich and lipid by the FTIR on lipid extract (NCC229), and
 495 two strains were also identified as rich in lipids by the FTIR on entire cells (NCC33, NCC199).
 496 One strain was identified by FTIR on lipid extract and FTIR on whole cells (NCC366) and
 497 seven strains were identified as rich in lipids by all three methods (NCC109, NCC182,
 498 NCC113, NCC270, NCC169, NCC303, NCC216).

499

500

501

502

503

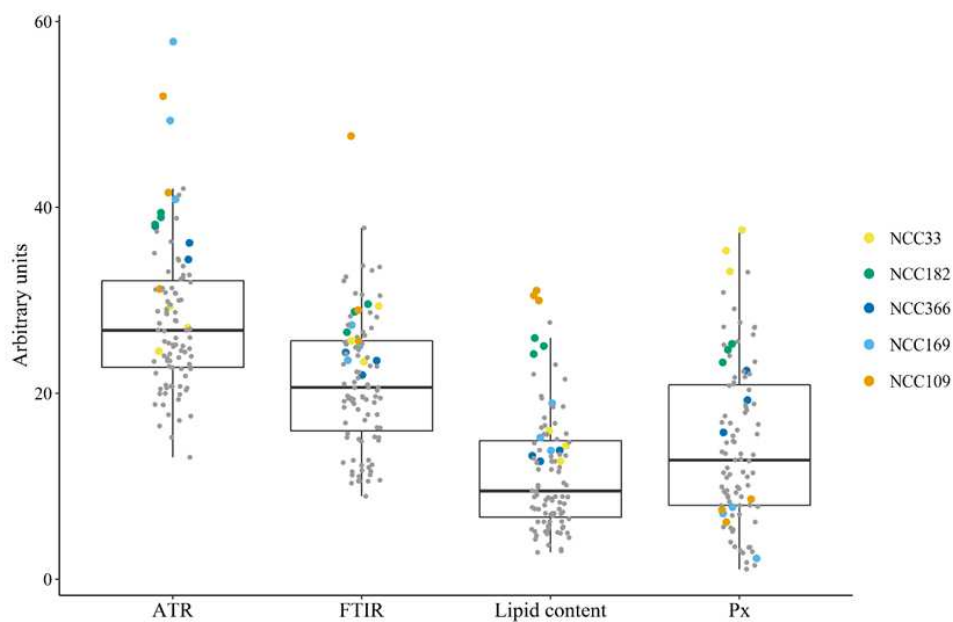
504

505 3.4 Strain selection

506 The selection of the strains exhibiting both high biomass productivity and high lipid rates, was
507 performed using whole sample distribution based on the lipid rates as estimated by the FTIR
508 approaches, gravimetry and strain biomass productivity as estimated by the fluorometry. On
509 the boxplots summarizing this data (Fig. 6), the colored dots represent the species with the
510 highest potential for biotechnology applications based on lipid molecules: *Nitzschia*
511 *alexandrina* NCC33, *Staurosira* sp NCC182 and *Opephora* sp NCC366 presented a lipid ratio
512 whatever the considered technique and productivity above the median. *Amphora* sp 2 NCC169
513 and *Nitzschia* sp 5 NCC109 were also identified with an above median lipid ratio, but with
514 lower productivity. Both these species were finally selected for their high lipid rate, even
515 though their productivity needs to be improved.

516

517



518 **Figure 6.** Boxplots summarizing the sample distribution criteria as measured with FTIR
519 methods, gravimetry for the lipid rate and fluorimetry for productivity. FTIR data was
520 expressed in arbitrary units. Lipid rate in %DW and Px in $\text{g.L}^{-1}.\text{day}^{-1}$ FTIR results were

521 multiplied by 20 for scaling purposes. The Px was multiplied by 100 for scaling purposes. 33
522 strains were assayed in independent biological triplicates.

523 4 Discussion

524 4.1.1 Determination of growth by fluorimetry

525 In this study a rapid and precise method for the estimation of cell abundance was needed to
526 efficiently establish the growth parameter of over 60 benthic diatom species. The use of
527 microscope counting chambers is slow, tedious and imprecise. The use of particle counters is
528 feasible only for those phytoplankton species within a certain size range that do not form
529 chains and have long appendages [65]. *In vivo* fluorescence has been successfully used in the
530 past to monitor the growth of phytoplankton culture [19, 20, 32, 41, 66]. *In vivo* fluorescence
531 measurement is rapid, sensitive and can be used with all types of diatoms cell structures, i.e.
532 on solitary cells such as *Entomoneis* and *Navicula* genera and on chain-forming diatoms, such
533 as *Extubocellulus* genus or aggregates of cells such as *Amphora* genus. However, *in vivo*
534 fluorescence is a measurement of the increase in chlorophyll and chlorophyll per cell varies
535 greatly with light intensity and cellular nutritional status so it can only be used to measure
536 growth when culture conditions are constant. Changes in light and other culture conditions
537 can lead to a change in chlorophyll a content. In this case F0 will also change, introducing a
538 bias in growth rate estimation. Reflectance can also be used to monitor the diatom growth
539 rate, as demonstrated in this study; the NDVI was highly correlated with the cell count as *in*
540 *vivo* fluorescence ($R > 0.70$). Méléder et al, 2003 [21] have already demonstrated this
541 correlation on monospecific benthic diatom cultures with *Entomoneis paludosa* and *Navicula*
542 *ramosissima*. Compared to fluorometry this technique required more culture sampling and a
543 filtration step which took longer to set up. Fluorometry was more sensitive than
544 spectroradiometry and required less material, which made it ideal for very low amounts of

545 biomass. *In vivo* fluorometry demonstrated several advantages: ease of use, real-time
546 measurement, non-destructive sampling.

547 4.1.2 Determination of the lipid rate by FTIR

548 The total lipid evaluation by gravimetry has been used for more than 50 years [67], but this
549 technique is clearly not compatible with screening efforts involving large numbers of samples
550 because of processing time required particularly in the extraction phase. Furthermore, large
551 volumes of samples are systematically needed for the measurements. It also requires sufficient
552 amounts of dried biological sample, thus making it unsuitable for high frequency monitoring
553 of small scale microalgal cultivation. Feng et al., (2013) [68] suggested that the presence of
554 chlorophyll could also affect the accuracy of the method. In comparison to conventional
555 chemical analysis, FTIR spectroscopy presented striking advantages due to its high reliability,
556 sensitivity and speed of the measurement [64]. An IR spectrometer coupled to a microtiter
557 plate reader open the possibility of high throughput analysis of a few nanograms of cell
558 material [69]. Coat et al, (2014) [26] demonstrated that the repeatability of FTIR signal
559 reached excellent values (10%), but only for a limited range of analyzed quantities of matter
560 (10^6 to 10^8 cells. mL⁻¹).

561 In the present study we have demonstrated that FTIR was a suitable technique to evaluate the
562 lipid ratio of diatoms. The use of the FTIR technique was even more rapid, due to the use of
563 direct fresh biomass. The results obtained on entire cells and lipid crude extracts were similar,
564 suggesting that measurement on entire cells did not improve the lipid quantification.
565 Measurement on whole cells could be considered sufficient to get a first idea on their
566 intracellular lipid rate. In addition, removing the silica did not seem necessary in view of the
567 present results to improve the lipid semi-quantification.

568 However, there are some limitations with this non-invasive technique. Even though the FTIR
569 results between entire cells and crude extracts were similar, we did not obtain a perfect linear

570 correlation with the traditional lipid extraction. Nevertheless, the FTIR method has the
571 advantage of simultaneously detecting the relative amount of lipids, carbohydrates and
572 proteins, even though those components overlapped, implying a certain degree of inaccuracy.
573 The different steps required for the lipid extraction can also produce a bias in the evaluation
574 of the lipid quantity, added to the inaccuracy of weighing of the lipid matter, making it less
575 precise than the spectrometric approach. The heterogeneity observed in the present FTIR
576 results could be also associated to the utilization of different strains (33 strains in the current
577 study). These are not described in the previous studies evaluating lipid quantification with
578 FTIR where only one species was used [26, 39, 70]. FTIR did not differentiate polar/apolar
579 lipids, or the different types of fatty acids produced by the strains. Lipid analysis by
580 chromatographic techniques coupled to mass spectrometry will be needed to further identify
581 the presence of interesting molecules like EPA or DHA. Nevertheless, the present study opens
582 the way to rapid and reliable semi-quantification of total amounts of intracellular lipids in
583 diatoms using a fast, non-invasive approach.

584

585

586

587

588

589

590

591

592

593

594 4.1.3 Method assessments

595 This study highlighted the efficiency of PAM and FTIR measurements as fast techniques to
 596 characterize both the growth and lipid content of microalgae (Tab.4). The use of PAM
 597 fluorometry was three times faster than cell counting. Where 15 minutes were necessary to
 598 count 3 samples (8.25 hours for 99 samples) and 10 minutes to measure reflectance by
 599 spectroradiometry (5.5 hours for 99 samples), only 5 minutes were necessary to measure F0
 600 by PAM fluorometry (2.75 hours for 99 samples).

601 **Table 4** Evaluation of three methods for algal growth kinetic determination

Evaluation index	Counting	NDVI	PAM
Cultivation scale	Large	Large–small	Large–small
Monitoring frequency	High	Medium	High
Sample state	Liquid	Filtered	Liquid
Volume consumed	5 - 1 mL	5- 3 mL	>1 mL
Time consumed	8.25 h	5.5 h	2.75 h
Equipment	Counting chambers	Spectroradiometer	PAM fluorometer
Reliability	Low	High	High

602

603 For lipid extraction, dichloromethane methanol solvent is usually used [34, 71–73]. Even
 604 though extraction was very effective, this method is known to have environmental and health
 605 risks [71]. Moreover, it is expensive: for 99 extractions, 10 L of each solvent was necessary
 606 costing 234 € (9.68 €/L of methanol and 13.72 €/L of dichloromethane (Fischer Scientific).
 607 The time necessary for the lipid extraction using Bligh and Dyer method was time-consuming
 608 (Tab. 5) due to the time needed for maceration and lipid separation. 5.5 weeks were necessary
 609 to perform the extraction of 99 samples (ca. 7920 hours). Nile red, a lipid soluble fluorescent
 610 dye, is commonly used to evaluate the lipid content of animal cells, microorganisms, and
 611 especially microalgal strains. Required time to perform spectrophotometric measurements
 612 with Nile red and Bodipy was evaluated at 10 mins per sample (measurement before and after
 613 Nile red or Bodipy application) and 25 to 40 mins incubation time was necessary after

614 application of Nile red or Bodipy before reading spectrophotometric values [68,74]. In
 615 comparison FTIR is a fast and eco-friendly technique. Two plates of 384 well were necessary
 616 to evaluate the lipid content for the 33 strains (each strain needed 15 wells with 5 wells per
 617 replicate). It took 30 seconds to read each well to obtain the FTIR spectra leading to ca. and 4
 618 hours to read an entire plate.

619 **Table 5** Evaluation of three methods for algal lipid content determination

Evaluation index	Gravimetric	FT-IR	Nile Red / Bodipy 505/515
Cultivation scale	Large	Large–small	Large–small
Monitoring frequency	Low	Medium	High
Sample state	Dried	Liquid	Liquid
Biomass consumed	100 mg	5 μ L of 1.0 mg mL ⁻¹	3 mL of OD ₆₈₀ 0.4
Time consumed	>50 h	8 h	25 h
Equipment	Nitrogen evaporator	FT-IR	FS & FM ^a
Accuracy	Total lipid	Total lipid	Neutral lipid

620 ^a Fluorescence spectrophotometer and Fluorescence microscopy

621 Nile red and Bodipy 505/515 staining are powerful quantification tools in terms of time and
 622 cost of biomass [75, 76], high throughput quantification method of lipids with Nile red or
 623 Bodipy 505/515 fluorescence can hardly been seen as a method for screening different species
 624 of microalgae, as the staining protocol is species specific. The significant disadvantages of
 625 Nile red were its limited photostability, interference with chlorophyll [77, 78], and difficulty
 626 of permeation for some species. Bodipy 505/515 produced a better marker than Nile red for
 627 visualizing neutral lipid content in fluorescence microscopy [75, 79] but some authors have
 628 reported disadvantages with these techniques such as background fluorescence of the dye in
 629 the medium and failure to quantify neutral lipids between rich and low oil strains. When
 630 microalgae were cultured on a large scale with a low-frequency monitoring requirement, any
 631 of the three methods could be adopted, although gravimetric determination might be
 632 preferable as it was an absolute method for quantification of both crude and neutral lipids
 633 without the need of specialized equipment. For the general laboratory culture of microalgae,

634 the FT-IR method for simultaneous characterization of total lipid, carbohydrate and protein
635 content and the Nile Red method for both neutral lipid content and location can be used, both
636 of which are relative quantification methods, but require special equipment.

637 Although these analyses demonstrate that FTIR and Nile Red were equally effective at
638 measuring lipid accumulation, FTIR was likely to be a more efficient tool for this purpose
639 because of its much faster analysis time and high reproducibility of results [80]. Furthermore,
640 FTIR may also be more suitable than Nile Red for efficiently detecting large increases in lipid
641 concentration. Nile Red does not appear to be efficient at accurately quantifying lipid
642 concentration above 20 mg/ml [77] while FTIR can efficiently detect linear lipid
643 concentration changes up to at least 250 mg [81]. Measurements with FTIR were more precise
644 because of the technical quintuplicate performed for the acquisition of spectra.

645 4.2 Screening for lipid rich benthic diatom strains

646 The aim of this study was to investigate the growth characteristics and the lipid rates of the
647 benthic marine diatom species hosted in the NCC bank in order to evaluate their potential for
648 original lipid bioproducts of potential economic interests. First, a selection based on the genera
649 identified in the literature was applied. Among the 134 strains hosted by the NCC,
650 corresponding to 40 genera and 101 species, 23 genera (corresponding to 77 species, 105
651 strains in the NCC) were largely studied [15, 16, 42, 43, 46,49, 52, 55, 58, 60, 82]. Among
652 these 23 genera, only 13 genera (42 species, 47 strains) with high productivity and/or
653 producing high lipid quantities were selected for the current study. Among these 47 strains,
654 only 18 strains, (6 genera, 17 species) grew successfully. Among the NCC's 40 genera, 17
655 genera (corresponding to 24 species, 29 strains) were not previously reported in the literature
656 in terms of productivity or ability to produce lipids or added value molecules. Thus, the 29
657 strains corresponding to these 17 genera were also selected and assayed for the current study.
658 Among these strains, 15 strains (13 species, 10 genera) did successfully grow. Finally, at the

659 end of this first screening step based on the growth rates, 33 strains (18 previously described
660 and 15 that have never been described before in the literature) were selected for the second
661 step of the screening process. As reported in the supplementary data, 43 strains failed to grow
662 which may be associated to shear stress; some species like *Rhizosolenia setigera* cells were
663 broken during agitation.

664 The second step consisted in the determination of the lipid rate on the selected strains. In
665 microalgae, it can typically vary from 1 to 85% of the dry weight under adverse conditions
666 [15, 83, 84]. Factors such as temperature, irradiance and most markedly nutrient availability
667 have been shown to affect both lipid composition and lipid rate [17, 57, 85]. In general, high
668 irradiance stimulates TAG accumulation [57], while under low irradiance, the polar lipids
669 (phospholipids and glycolipids), structurally and functionally associated to cell membrane,
670 are preferentially synthesized [17]. The lipid rate found in the current study for the *Amphora*
671 genus was similar to that estimated by Renaud et al., (1999) [46]. The authors proposed 19%
672 DW vs. 14 to 19% DW in the current study. This similarity could be explained by the similar
673 culture conditions for light and nutrient although the temperature was different (25°C vs 16°C
674 in the current study). According to Chtourou et al., 2015 [61], the temperature could be an
675 important factor for *Amphora* genus lipid rate which can achieve lipid rates up to 24% of DW
676 at 20°C. Media enrichment could also be an important factor: *Amphora* cells grown in media
677 enriched with macronutrients and trace metal can also achieve lipid rates up to 32% under low
678 light conditions (11.4 $\mu\text{mol photon m}^{-2} \text{s}^{-1}$) [62] or under nitrogen deficiency [55]. These
679 factors could be used as guideline to improve lipid production of the selected *Amphora* strains
680 hosted in the NCC. The biomass measurements for the *Amphora* genus in Zhao et al study
681 [49] were in accordance with our results. The authors found 0.13 g.L⁻¹ vs. 0.16 g.L⁻¹ in the
682 current study. According to their measured lipid productivity, the *Amphora* genus appeared
683 to be a good candidate for lipid based potential applications.

684 The lipid rate found in our study for the *Nitzschia* genus was similar to the results found by
685 Renaud et al., (1999) for one species [46] but very different for *Nitzschia sp 5* NCC109 in the
686 current study. They found a lipid rate of between 13 and 16% DW for the *Nitzschia* genus. In
687 the present study, *Nitzschia alexandrina* NCC33 showed a lipid rate of 13% DW but for
688 *Nitzschia sp 5* NCC109 it was estimated up to 30% DW. However, this rate could only be
689 found for *Nitzschia* genus under nitrogen or silica deficiency increasing up to 45-47% DW
690 [44,45,47]. Since the amount of nitrogen was not monitored in the culture, it is possible that
691 nitrogen depletion occurred, explaining the high lipid rate measured for *Nitzschia sp 5*
692 NCC109. Further analysis for this strain is necessary to establish whether this was due to the
693 absolute lipid richness of the strain, or to a bias in culture conditions.

694 The biomass found for the *Nitzschia* genus in Zhao et al (2016) [49] was lower than the
695 measurements in the present study. The biomass estimated in the current study ranged from
696 0.07 to 0.14 g.L⁻¹ vs. 0.19 to 0.23 g.L⁻¹. This difference could be due to the light/dark cycle
697 culture conditions, suggesting that under continuous light the biomass production would be
698 more significant for this specific genus [86].

699 It has been demonstrated that the species with a lipid rate of 30% DW and productivity under
700 non-optimized conditions of around 0.30 g.L⁻¹ could be potential strains for lipid production
701 [15, 87]. In the present study *Nitzschia sp 5* NCC109 had the highest lipid rate, 30.51% DW,
702 *Nitzschia alexandrina* NCC33 with the highest productivity 0.35 g.L⁻¹. Both strains were thus
703 selected as candidates for further analyses to assay their potential for lipid-based applications.

704 The *Navicula* genus also presented a lipid rate for the 2 tested species of between 9% and 23%
705 DW. Zhao et al, (2016) [49] found a lipid range from 5 to 30% for 3 *Navicula* species
706 (*Navicula ramosissima*, *Navicula molli* and *Navicula halophia*) and Scholz et al, (2013) [58]
707 between 18-25% for 6 *Navicula* species (*Navicula digito-radiata*, *Navicula forcipata*,
708 *Navicula gregaria*, *Navicula perminuta*, *Navicula phyllepta* and *Navicula salinicola*). The

709 data were estimated in spite of very different culture conditions (dark night cycle at 600 μmol
710 $\text{photon m}^{-2} \text{s}^{-1}$). This observation is important, since for that specific genus, the observed
711 variability in the range of lipid rate for the different tested species is the same regardless of
712 the culture conditions used. The biomass measurements found by Zhao et al., (2016) for the
713 *Navicula* genus are in accordance with the present study, ranging from 0.10 to 0.17 g.L^{-1} vs.
714 0.13 to 0.17 g.L^{-1} suggesting that this genus could produce the identical biomass quantity
715 under light/dark cycle or continuous light.

716 The lipid rate found for the *Extubocellulus* genus in Slocombe et al., (2015) [60] was 23%
717 DW and the biomass equalled 0.06 g. L^{-1} . The lipid rate for this genus, 15% DW in the actual
718 study was lower but the biomass obtained was higher 0.16 g.L^{-1} . The culture was grown under
719 light/dark cycle in the Slocombe et al. study suggesting that this genus could grow better under
720 continuous light but produce more lipids under light/dark cycle. Wahidin et al., (2013) [88]
721 found the same trend for the microalgae genus *Nannochloropsis*.

722 Knuckey et al, (2002) [52] found 27% DW of lipid rate for the *Entomoneis* genera. In the
723 present study 9 strains of *Entomoneis* presented lipid rates ranging from 3 to 13% DW. In the
724 Knuckey et al, (2002) study the pH and the nutrients were monitored. The culture conditions
725 were identical to the present study (continuous light, 120 $\mu\text{mol photon m}^{-2} \text{s}^{-1}$). It is therefore
726 possible, that for this genus a nutrient limitation occurred and instead of enhancing the lipid
727 production, it has diminished it. It has been reported that this species produces EPA which is
728 interesting for nutraceutical products [52].

729 In this study the species with the highest lipid rate (~20 %) and the greatest biomass
730 productivity (up to 0.24 $\text{g.L}^{-1}.\text{day}^{-1}$) was *Staurosira* sp NCC182. This species was similar in
731 productivity to the marine microalgae *Nannochloropsis* sp. with productivity higher than 0.21
732 $\text{g.L}^{-1}.\text{day}^{-1}$ and reached a total lipid rate of 30% DW, when cultivated in batch mode under
733 continuous light and reached 68% DW of lipid production under nitrogen deprivation [89].

734 *Nannochloropsis* sp. was investigated for algal biofuel production due to its ease of growth
735 and high oil rate. *Staurosira* sp was grown in raceway ponds by Huntley et al., (2015) [16]
736 and the lipid quantification demonstrated that these strains could reach a lipid rate of 45.5%
737 DW under low N content. In the present study, *Staurosira* sp. NCC182 presented all the
738 characteristics (good productivity and oil content) to be produced on a large scale and was
739 considered to be one of the most promising candidates for lipid-based applications.

740 Even if the microalgae oil yield is strain-dependent it is generally superior to other vegetable
741 crops [15, 87]. Oil content in Corn, Hemp, Soybean, Sunflower or Palm oil varied from 18
742 (Soybean) to 44% DW (Corn). Christie et al., (2007) [15] demonstrated that microalgae with
743 a lipid rate up to 30% of DW could produce 58 L/Ha of oil and microalgae up to 70% of DW
744 could produce 136 L/Ha of oil. Corn can produce 172 L/Ha, soybean 446 L/Ha and oil palm
745 5950 L/Ha but require a lot of land for production: 1540 M/Ha for Corn, 594 M/Ha for
746 Soybean and 45 M/Ha for oil palm. Microalgae production only requires 2 to 5 M/Ha. In the
747 NCC collection *Nitzschia* sp 5 NCC109 and *Staurosira* sp NCC182 had all the mandatory
748 features to be grown on a large scale: good productivity and high lipid rate. In the present
749 study, *Nitzschia alexandrina* NCC33 had the highest productivity in the NCC collection (0.36
750 g.L⁻¹.day⁻¹) and presented a lipid rate superior to 10% DW. Despite its low productivity of
751 0.08 g.L⁻¹.day⁻¹, *Amphora* sp 2 NCC169 was chosen for its lipid rate, superior to 15% DW.
752 These 4 strains were selected for further analyses and to improve their productivity and lipid
753 rate with the objective of supplying new resources for lipid based applications. *Entomoneis*
754 *paludosa* was also selected despite its low lipid rate since it was already characterized for its
755 ability to produce EPA.

756 Among the strains not described in the literature, the measured lipid rate ranged from 3 to
757 20% DW and productivity from 0.05 to 0.27 g.L⁻¹.day⁻¹. Only one genus was selected for
758 further study: *Opephora* sp 1 NCC366. This species showed a mandatory balance between

759 relatively high biomass productivity (0.23 g.L^{-1}) and a high intracellular lipid rate (above 10%
760 DW).

761 5 Conclusion

762 In this work, we focused on developing an easy to use screening method to explore the
763 NCC bank for diatom strains with the highest relative lipid content. The experimental results
764 showed that the combined use of water-PAM to estimate strain growth kinetics and FTIR on
765 whole cells to estimate the semi-quantitative strain macromolecular content and more
766 specifically lipids, could be rapid, reliable and accessible techniques. The developed
767 methodology opens the way to a systematic, fast, and convenient screening of microorganisms
768 (microalgae in this proof of concept). Moreover, the sensitivity and specificity of the method
769 makes it suitable for a reasonable amount of biomass. This method could also be used in
770 systematic studies for the optimization of culture conditions and to measure the influence of
771 the environment on the metabolic plasticity of the assessed organism. Using this screening
772 approach, 5 strains hosted in the NCC bank were selected for their high productivity and high
773 lipid rate: *Nitzschia alexandrina* NCC33, *Staurosira sp* NCC182, *Opephora sp 1* NCC366,
774 *Nitzschia sp 5* NCC109 and *Amphora sp 2* NCC169. The lipid rate achieved by these strains
775 reached a maximum of 30% DW in the assayed cultivation conditions.

776

777 In order to improve the lipid quantity, the selected strains could be grown under different
778 culture conditions. The impact of light, temperature and nutrients, especially nitrogen, could
779 be assayed both in terms of lipid productivity and ecophysiology to ensure the highest growth
780 rate possible. Once optimal conditions are found for those strains, the production in
781 photobioreactors could be tested and productivity and lipid rates evaluated, in order to

782 estimate quantitatively if the selected strains can compete with the best ones found in the
783 literature [90–92]. Finally, depending on their oil quality and original lipid activities, those
784 strains may constitute new and original genetic resources that could have potential interesting
785 applications (biodiesel, pharmaceutical, etc.).

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802 6 References

- 803 [1] T.M. Mata, A.A. Martins, N.S. Caetano, Microalgae for biodiesel production and other
804 applications: A review, *Renew. Sustain. Energy Rev.* 14 (2010) 217–232.
805 doi:10.1016/j.rser.2009.07.020.
- 806 [2] P. Spolaore, C. Joannis-Cassan, E. Duran, A. Isambert, Commercial applications of
807 microalgae, *J. Biosci. Bioeng.* 101 (2006) 87–96. doi:10.1263/jbb.101.87.

- 808 [3] M.A. Borowitzka, High-value products from microalgae—their development and
809 commercialisation, *J. Appl. Phycol.* 25 (2013) 743–756. doi:10.1007/s10811-013-9983-
810 9.
- 811 [4] M.I. Khan, J.H. Shin, J.D. Kim, The promising future of microalgae: current status,
812 challenges, and optimization of a sustainable and renewable industry for biofuels, feed,
813 and other products, *Microb. Cell Factories.* 17 (2018). doi:10.1186/s12934-018-0879-x.
- 814 [5] Y.-F. Niu, M.-H. Zhang, D.-W. Li, W.-D. Yang, J.-S. Liu, W.-B. Bai, H.-Y. Li,
815 Improvement of Neutral Lipid and Polyunsaturated Fatty Acid Biosynthesis by
816 Overexpressing a Type 2 Diacylglycerol Acyltransferase in Marine Diatom
817 *Phaeodactylum tricornutum*, *Mar. Drugs.* 11 (2013) 4558–4569.
818 doi:10.3390/md11114558.
- 819 [6] Z. Yi, M. Xu, X. Di, S. Brynjolfsson, W. Fu, Exploring Valuable Lipids in Diatoms,
820 *Front. Mar. Sci.* 4 (2017). doi:10.3389/fmars.2017.00017.
- 821 [7] E. Artamonova, J. Svenning, T. Vasskog, E. Hansen, H. Eilertsen, Analysis of
822 phospholipids and neutral lipids in three common northern cold water diatoms:
823 *Coscinodiscus concinnus*, *Porosira glacialis*, and *Chaetoceros socialis*, by ultra-high
824 performance liquid chromatography-mass spectrometry, *J. Appl. Phycol.* 29 (2017)
825 1241–1249.
- 826 [8] K.W. Chew, J.Y. Yap, P.L. Show, N.H. Suan, J.C. Juan, T.C. Ling, D.-J. Lee, J.-S.
827 Chang, Microalgae biorefinery: High value products perspectives, *Bioresour. Technol.*
828 229 (2017) 53–62. doi:10.1016/j.biortech.2017.01.006.
- 829 [9] G.A. Dunstan, J.K. Volkman, S.M. Barrett, J.-M. Leroi, S.W. Jeffrey, Essential
830 polyunsaturated fatty acids from 14 species of diatom (Bacillariophyceae),
831 *Phytochemistry.* 35 (1993) 155–161. doi:10.1016/S0031-9422(00)90525-9.
- 832 [10] Y.-C. Chen, The biomass and total lipid content and composition of twelve species of
833 marine diatoms cultured under various environments, *Food Chem.* 131 (2012) 211–219.
834 doi:10.1016/j.foodchem.2011.08.062.
- 835 [11] L. Chuecas, J.P. Riley, Component Fatty Acids of the Total Lipids of Some Marine
836 Phytoplankton, *J. Mar. Biol. Assoc. U. K.* 49 (1969) 97.
837 doi:10.1017/S0025315400046439.
- 838 [12] K. Nagao, T. Yanagita, Conjugated fatty acids in food and their health benefits, *J. Biosci.*
839 *Bioeng.* 100 (2005) 152–157. doi:10.1263/jbb.100.152.
- 840 [13] M. Gross, The mysteries of the diatoms, *Curr. Biol.* 22 (2012) R581–R585.
841 <http://www.sciencedirect.com/science/article/pii/S0960982212008664> (accessed
842 September 12, 2016).
- 843 [14] T. Lebeau, J.-M. Robert, Diatom cultivation and biotechnologically relevant products.
844 Part II: Current and putative products, *Appl. Microbiol. Biotechnol.* 60 (2003) 624–632.
845 doi:10.1007/s00253-002-1177-3.
- 846 [15] Y. Chisti, Biodiesel from microalgae, *Biotechnol. Adv.* 25 (2007) 294–306.
847 doi:10.1016/j.biotechadv.2007.02.001.
- 848 [16] M.E. Huntley, Z.I. Johnson, S.L. Brown, D.L. Sills, L. Gerber, I. Archibald, S.C.
849 Machesky, J. Granados, C. Beal, C.H. Greene, Demonstrated large-scale production of
850 marine microalgae for fuels and feed, *Algal Res.* 10 (2015) 249–265.
851 doi:10.1016/j.algal.2015.04.016.
- 852 [17] Q. Hu, M. Sommerfeld, E. Jarvis, M. Ghirardi, M. Posewitz, M. Seibert, A. Darzins,
853 Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and
854 advances, *Plant J.* 54 (2008) 621–639. doi:10.1111/j.1365-313X.2008.03492.x.
- 855 [18] T.A. Dempster, M.R. Sommerfeld, Effects of environmental conditions on growth and
856 lipid accumulation in *Nitzschia communis* (Bacillariophyceae), *J. Phycol.* 34 (1998)
857 712–721.

- 858 [19] V. Vyhnalek, Z. Fišar, A. Fišarová, J. Komarkova, In vivo fluorescence of chlorophyll
859 a: Estimation of phytoplankton biomass and activity in Římov Reservoir (Czech
860 Republic), Water Sci. Technol. 28 (1993) 29–33.
861 <http://wst.iwaponline.com/content/28/6/29.abstract> (accessed October 10, 2017).
- 862 [20] A.D. Steinman, G.A. Lamberti, P.R. Leavitt, D.G. Uzarski, Biomass and Pigments of
863 Benthic Algae, in: Methods Stream Ecol. Vol. 1, Elsevier, 2017: pp. 223–241.
864 doi:10.1016/B978-0-12-416558-8.00012-3.
- 865 [21] V. Méléder, L. Barillé, P. Launeau, V. Carrère, Y. Rincé, Spectrometric constraint in
866 analysis of benthic diatom biomass using monospecific cultures, Remote Sens. Environ.
867 88 (2003) 386–400. doi:10.1016/j.rse.2003.08.009.
- 868 [22] L. Akoto, R. Pel, H. Irth, A.T. Udo, R.J. Vreuls, Automated GC–MS analysis of raw
869 biological samples: Application to fatty acid profiling of aquatic micro-organisms, J.
870 Anal. Appl. Pyrolysis. 73 (2005) 69–75.
- 871 [23] K.E. Cooksey, J.B. Guckert, S.A. Williams, P.R. Callis, Fluorometric determination of
872 the neutral lipid content of microalgal cells using Nile Red, J. Microbiol. Methods. 6
873 (1987) 333–345.
- 874 [24] J. Koreivienė, Microalgae Lipid Staining with Fluorescent BODIPY Dye, (2017).
- 875 [25] D. Jaeger, C. Pilger, H. Hachmeister, E. Oberländer, R. Wördenweber, J. Wichmann,
876 J.H. Mussnug, T. Huser, O. Kruse, Label-free in vivo analysis of intracellular lipid
877 droplets in the oleaginous microalga *Monoraphidium neglectum* by coherent Raman
878 scattering microscopy, Sci. Rep. 6 (2016) 35340.
- 879 [26] R. Coat, V. Montalescot, E.S. León, D. Kucma, C. Perrier, S. Jubeau, G. Thouand, J.
880 Legrand, J. Pruvost, O. Gonçalves, Unravelling the matrix effect of fresh sampled cells
881 for in vivo unbiased FTIR determination of the absolute concentration of total lipid
882 content of microalgae, Bioprocess Biosyst. Eng. 37 (2014) 2175–2187.
- 883 [27] J. Fan, Y. Cui, M. Wan, W. Wang, Y. Li, Lipid accumulation and biosynthesis genes
884 response of the oleaginous *Chlorella pyrenoidosa* under three nutrition stressors,
885 Biotechnol. Biofuels. 7 (2014) 17. doi:10.1186/1754-6834-7-17.
- 886 [28] P. Metzger, C. Largeau, *Botryococcus braunii*: a rich source for hydrocarbons and related
887 ether lipids, Appl. Microbiol. Biotechnol. 66 (2005) 486–496. doi:10.1007/s00253-004-
888 1779-z.
- 889 [29] R.R. Guillard, Culture of phytoplankton for feeding marine invertebrates, in: Cult. Mar.
890 Invertebr. Anim., Springer, 1975: pp. 29–60.
- 891 [30] R.R. Guillard, J.H. Ryther, Studies of marine planktonic diatoms: I. *Cyclotella* Nana
892 Hustedt, and *Detonula Confervacea* (CLEVE) Gran., Can. J. Microbiol. 8 (1962) 229–
893 239.
- 894 [31] Z.-K. Yang, Y.-F. Niu, Y.-H. Ma, J. Xue, M.-H. Zhang, W.-D. Yang, J.-S. Liu, S.-H.
895 Lu, Y. Guan, H.-Y. Li, Molecular and cellular mechanisms of neutral lipid accumulation
896 in diatom following nitrogen deprivation, Biotechnol. Biofuels. 6 (2013) 67.
897 <https://biotechnologyforbiofuels.biomedcentral.com/articles/10.1186/1754-6834-6-67>
898 (accessed October 6, 2017).
- 899 [32] C. Honeywill, D. Paterson, S. Hagerthey, Determination of microphytobenthic biomass
900 using pulse-amplitude modulated minimum fluorescence, Eur. J. Phycol. 37 (2002) 485–
901 492. doi:10.1017/S0967026202003888.
- 902 [33] B. Gompertz, XXIV. On the nature of the function expressive of the law of human
903 mortality, and on a new mode of determining the value of life contingencies. In a letter
904 to Francis Baily, Esq. FRS &c, Philos. Trans. R. Soc. Lond. 115 (1825) 513–583.
- 905 [34] E.G. Bligh, W.J. Dyer, A RAPID METHOD OF TOTAL LIPID EXTRACTION AND
906 PURIFICATION, Can. J. Biochem. Physiol. 37 (1959) 911–917. doi:10.1139/y59-099.

- 907 [35] G. Breuer, W.A. Evers, J.H. de Vree, D.M. Kleinegris, D.E. Martens, R.H. Wijffels, P.P.
908 Lamers, Analysis of fatty acid content and composition in microalgae, *J. Vis. Exp. JoVE*.
909 (2013).
- 910 [36] E.S. León, R. Coat, B. Moutel, J. Pruvost, J. Legrand, O. Gonçalves, Influence of
911 physical and chemical properties of HTSXT-FTIR samples on the quality of prediction
912 models developed to determine absolute concentrations of total proteins, carbohydrates
913 and triglycerides: a preliminary study on the determination of their absolute
914 concentrations in fresh microalgal biomass, *Bioprocess Biosyst. Eng.* 37 (2014) 2371–
915 2380.
- 916 [37] W. Zeroual, C. Choisy, S.M. Doglia, H. Bobichon, J.-F. Angiboust, M. Manfait,
917 Monitoring of bacterial growth and structural analysis as probed by FT-IR spectroscopy,
918 *Biochim. Biophys. Acta BBA - Mol. Cell Res.* 1222 (1994) 171–178. doi:10.1016/0167-
919 4889(94)90166-X.
- 920 [38] D.H. Williams, I. Fleming, *Spectroscopic methods in organic chemistry*, McGraw-Hill,
921 1980.
- 922 [39] I. Schaub, H. Wagner, M. Graeve, U. Karsten, Effects of prolonged darkness and
923 temperature on the lipid metabolism in the benthic diatom *Navicula perminuta* from the
924 Arctic Adventfjorden, Svalbard, *Polar Biol.* (2017) 1–15. doi:10.1007/s00300-016-
925 2067-y.
- 926 [40] N. Sourial, C. Wolfson, B. Zhu, J. Quail, J. Fletcher, S. Karunanathan, K. Bandeen-
927 Roche, F. Béland, H. Bergman, Correspondence analysis is a useful tool to uncover the
928 relationships among categorical variables, *J. Clin. Epidemiol.* 63 (2010) 638–646.
929 doi:10.1016/j.jclinepi.2009.08.008.
- 930 [41] C.J. Lorenzen, A method for the continuous measurement of in vivo chlorophyll
931 concentration, in: *Deep Sea Res. Oceanogr. Abstr.*, Elsevier, 1966: pp. 223–227.
932 <http://www.sciencedirect.com/science/article/pii/0011747166911028> (accessed October
933 10, 2017).
- 934 [42] G. Chen, Y. Jiang, F. Chen, Fatty acid and lipid class composition of the
935 eicosapentaenoic acid-producing microalga, *Nitzschia laevis*, *Food Chem.* 104 (2007)
936 1580–1585. doi:10.1016/j.foodchem.2007.03.008.
- 937 [43] M.C. Dalay, S.S. Gunes, others, Biodiesel from Microalgae: A Renewable Energy
938 Source, *Middle East J. Sci. Res.* 22 (2014) 350–355.
939 [http://idosi.org/mejsr/mejsr21\(12\)14/15.pdf](http://idosi.org/mejsr/mejsr21(12)14/15.pdf) (accessed February 19, 2016).
- 940 [44] M.J. Griffiths, S.T.L. Harrison, Lipid productivity as a key characteristic for choosing
941 algal species for biodiesel production, *J. Appl. Phycol.* 21 (2009) 493–507.
942 doi:10.1007/s10811-008-9392-7.
- 943 [45] J. Johansen, P. Lemke, N. Nagle, P. Chelf, P. Roessler, R. Galloway, S. Toon, Addendum
944 to Microalgae Culture Collection 1986-1987, National Renewable Energy Lab.(NREL),
945 Golden, CO (United States), 1987.
- 946 [46] S.M. Renaud, L.-V. Thinh, D.L. Parry, The gross chemical composition and fatty acid
947 composition of 18 species of tropical Australian microalgae for possible use in
948 mariculture, *Aquaculture.* 170 (1999) 147–159.
- 949 [47] J. Sheehan, T. Dunahay, J. Benemann, P. Roessler, Look Back at the U.S. Department
950 of Energy’s Aquatic Species Program: Biodiesel from Algae; Close-Out Report,
951 National Renewable Energy Lab., Golden, CO. (US), 1998. doi:10.2172/15003040.
- 952 [48] Z.-Y. Wen, F. Chen, Production potential of eicosapentaenoic acid by the diatom
953 *Nitzschia laevis*, *Biotechnol. Lett.* 22 (2000) 727–733.
954 <http://link.springer.com/article/10.1023/A:1005666219163> (accessed February 18,
955 2016).

- 956 [49] F.-Y. Zhao, J.-R. Liang, Y.-H. Gao, Q.-Q. Luo, Y. Yu, C.-P. Chen, L. Sun, Variations
957 in the Total Lipid Content and Biological Characteristics of Diatom Species for Potential
958 Biodiesel Production, *J. Fundam. Renew. Energy Appl.* 2016 (2016).
959 [http://www.omicsonline.com/open-access/variations-in-the-total-lipid-content-and-](http://www.omicsonline.com/open-access/variations-in-the-total-lipid-content-and-biological-characteristics-of-diatom-species-for-potential-biodiesel-production-2090-4541-1000201.php?aid=66659)
960 [biological-characteristics-of-diatom-species-for-potential-biodiesel-production-2090-](http://www.omicsonline.com/open-access/variations-in-the-total-lipid-content-and-biological-characteristics-of-diatom-species-for-potential-biodiesel-production-2090-4541-1000201.php?aid=66659)
961 [4541-1000201.php?aid=66659](http://www.omicsonline.com/open-access/variations-in-the-total-lipid-content-and-biological-characteristics-of-diatom-species-for-potential-biodiesel-production-2090-4541-1000201.php?aid=66659) (accessed February 9, 2016).
- 962 [50] T. Jauffrais, S. Drouet, V. Turpin, V. Méléder, B. Jesus, B. Cognie, P. Raimbault, R.P.
963 Cosson, P. Decottignies, V. Martin-Jézéquel, Growth and biochemical composition of a
964 microphytobenthic diatom (*Entomoneis paludosa*) exposed to shorebird (*Calidris alpina*)
965 droppings, *J. Exp. Mar. Biol. Ecol.* 469 (2015) 83–92.
- 966 [51] T. Jauffrais, B. Jesus, V. Méléder, V. Turpin, A.D.P.G. Russo, P. Raimbault, V.M.
967 Jézéquel, Physiological and photophysiological responses of the benthic diatom
968 *Entomoneis paludosa* (Bacillariophyceae) to dissolved inorganic and organic nitrogen in
969 culture, *Mar. Biol.* 163 (2016). doi:10.1007/s00227-016-2888-9.
- 970 [52] R.M. Knuckey, M.R. Brown, S.M. Barrett, G.M. Hallegraeff, Isolation of new
971 nanoplanktonic diatom strains and their evaluation as diets for juvenile Pacific oysters
972 (*Crassostrea gigas*), *Aquaculture.* 211 (2002) 253–274.
973 doi:[http://dx.doi.org/10.1016/S0044-8486\(02\)00010-8](http://dx.doi.org/10.1016/S0044-8486(02)00010-8).
- 974 [53] A.T. Soares, B.F. Silva, L.L. Fialho, M.A.G. Pequeno, A.A.H. Vieira, A.G. Souza, N.R.
975 Antoniosi Filho, Chromatographic characterization of triacylglycerides and fatty acid
976 methyl esters in microalgae oils for biodiesel production, *J. Renew. Sustain. Energy.* 5
977 (2013) 053111. doi:10.1063/1.4821289.
- 978 [54] T. Viriyayingsiri, P. Sittplangkoon, S. Powtongsook, K. Nootong, Continuous
979 production of diatom *Entomoneis* sp. in mechanically stirred tank and flat-panel airlift
980 photobioreactors, *Prep. Biochem. Biotechnol.* 46 (2016) 740–746.
981 doi:10.1080/10826068.2015.1135460.
- 982 [55] F.J. Fields, J.P. Kocielek, An evolutionary perspective on selecting high-lipid-content
983 diatoms (Bacillariophyta), *J. Appl. Phycol.* 27 (2015) 2209–2220. doi:10.1007/s10811-
984 014-0505-1.
- 985 [56] G.E. Fogg, D.M. Collyer, The accumulation of fats as a characteristic of certain classes
986 of algae, *Proc 8 Int Bot Congr.* 17 (1954) 28.
- 987 [57] P.G. Roessler, ENVIRONMENTAL CONTROL OF GLYCEROLIPID
988 METABOLISM IN MICROALGAE: COMMERCIAL IMPLICATIONS AND
989 FUTURE RESEARCH DIRECTIONS, *J. Phycol.* 26 (1990) 393–399.
990 doi:10.1111/j.0022-3646.1990.00393.x.
- 991 [58] B. Scholz, G. Liebezeit, Biochemical characterisation and fatty acid profiles of 25
992 benthic marine diatoms isolated from the Solthörn tidal flat (southern North Sea), *J.*
993 *Appl. Phycol.* 25 (2013) 453–465. doi:10.1007/s10811-012-9879-0.
- 994 [59] M. Islam, M. Magnusson, R. Brown, G. Ayoko, M. Nabi, K. Heimann, Microalgal
995 Species Selection for Biodiesel Production Based on Fuel Properties Derived from Fatty
996 Acid Profiles, *Energies.* 6 (2013) 5676–5702. doi:10.3390/en6115676.
- 997 [60] S.P. Slocombe, Q. Zhang, M. Ross, A. Anderson, N.J. Thomas, Á. Lapresa, C. Rad-
998 Menéndez, C.N. Campbell, K.D. Black, M.S. Stanley, Unlocking nature’s treasure-
999 chest: screening for oleaginous algae, *Sci. Rep.* 5 (2015) 9844.
- 1000 [61] H. Chtourou, I. Dahmen, A. Jebali, F. Karray, I. Hassairi, S. Abdelkafi, H. Ayadi, S.
1001 Sayadi, A. Dhoubib, Characterization of *Amphora* sp., a newly isolated diatom wild
1002 strain, potentially usable for biodiesel production, *Bioprocess Biosyst. Eng.* 38 (2015)
1003 1381–1392. doi:10.1007/s00449-015-1379-6.

- 1004 [62] M. De la Pena, Cell growth and nutritive value of the tropical benthic diatom, *Amphora*
1005 sp., at varying levels of nutrients and light intensity, and different culture locations, *J.*
1006 *Appl. Phycol.* 19 (2007) 647–655.
- 1007 [63] M. Giordano, M. Kansiz, P. Heraud, J. Beardall, B. Wood, D. McNaughton, Fourier
1008 transform infrared spectroscopy as a novel tool to investigate changes in intracellular
1009 macromolecular pools in the marine microalga *Chaetoceros muellerii*
1010 (*Bacillariophyceae*), *J. Phycol.* 37 (2001) 271–279.
- 1011 [64] H. Wagner, S. Dunker, Z. Liu, C. Wilhelm, Subcommunity FTIR-spectroscopy to
1012 determine physiological cell states, *Curr. Opin. Biotechnol.* 24 (2013) 88–94.
1013 doi:<https://doi.org/10.1016/j.copbio.2012.09.008>.
- 1014 [65] L.E. Brand, R.R. Guillard, L.S. Murphy, A method for the rapid and precise
1015 determination of acclimated phytoplankton reproduction rates, *J. Plankton Res.* 3 (1981)
1016 193–201. <http://plankt.oxfordjournals.org/content/3/2/193.short> (accessed April 5,
1017 2016).
- 1018 [66] M. Gilstad, E. Sakshaug, Growth rates of ten diatom species from the Barents Sea at
1019 different irradiances and day lengths., *Mar. Ecol. Prog. Ser. Oldendorf.* 64 (1990) 169–
1020 173.
- 1021 [67] C. Breil, M. Abert Vian, T. Zemb, W. Kunz, F. Chemat, “Bligh and Dyer” and Folch
1022 Methods for Solid–Liquid–Liquid Extraction of Lipids from Microorganisms.
1023 Comprehension of Solvation Mechanisms and towards Substitution with Alternative
1024 Solvents, *Int. J. Mol. Sci.* 18 (2017). doi:10.3390/ijms18040708.
- 1025 [68] G.-D. Feng, F. Zhang, L.-H. Cheng, X.-H. Xu, L. Zhang, H.-L. Chen, Evaluation of FT-
1026 IR and Nile Red methods for microalgal lipid characterization and biomass composition
1027 determination, *Bioresour. Technol.* 128 (2013) 107–112.
- 1028 [69] K. Stehfest, J. Toepel, C. Wilhelm, The application of micro-FTIR spectroscopy to
1029 analyze nutrient stress-related changes in biomass composition of phytoplankton algae,
1030 *Plant Physiol. Biochem.* 43 (2005) 717–726. doi:10.1016/j.plaphy.2005.07.001.
- 1031 [70] A.P. Dean, D.C. Sigee, B. Estrada, J.K. Pittman, Using FTIR spectroscopy for rapid
1032 determination of lipid accumulation in response to nitrogen limitation in freshwater
1033 microalgae, *Bioresour. Technol.* 101 (2010) 4499–4507.
1034 doi:10.1016/j.biortech.2010.01.065.
- 1035 [71] M. Mubarak, A. Shaija, T.V. Suchithra, A review on the extraction of lipid from
1036 microalgae for biodiesel production, *Algal Res.* 7 (2015) 117–123.
1037 doi:10.1016/j.algal.2014.10.008.
- 1038 [72] Y. Li, F.G. Naghdi, S. Garg, T.C. Adarme-Vega, K.J. Thurecht, W.A. Ghafor, S.
1039 Tannock, P.M. Schenk, A comparative study: the impact of different lipid extraction
1040 methods on current microalgal lipid research, *Microb. Cell Factories.* 13 (2014) 14.
- 1041 [73] S.J. Lee, B.-D. Yoon, H.-M. Oh, Rapid method for the determination of lipid from the
1042 green alga *Botryococcus braunii*, *Biotechnol. Tech.* 12 (1998) 553–556.
- 1043 [74] D. Xu, Z. Gao, F. Li, X. Fan, X. Zhang, N. Ye, S. Mou, C. Liang, D. Li, Detection and
1044 quantitation of lipid in the microalga *Tetraselmis subcordiformis* (Wille) Butcher with
1045 BODIPY 505/515 staining, *Bioresour. Technol.* 127 (2013) 386–390.
1046 doi:10.1016/j.biortech.2012.09.068.
- 1047 [75] H. De la Hoz Siegler, W. Ayidzoe, A. Ben-Zvi, R. Burrell, W. McCaffrey, Improving
1048 the reliability of fluorescence-based neutral lipid content measurements in microalgal
1049 cultures, *Algal Res.* 1 (2012) 176–184.
- 1050 [76] T. Mutanda, D. Ramesh, S. Karthikeyan, S. Kumari, A. Anandraj, F. Bux,
1051 Bioprospecting for hyper-lipid producing microalgal strains for sustainable biofuel
1052 production, *Bioresour. Technol.* 102 (2011) 57–70. doi:10.1016/j.biortech.2010.06.077.

- 1053 [77] W. Chen, C. Zhang, L. Song, M. Sommerfeld, Q. Hu, A high throughput Nile red method
1054 for quantitative measurement of neutral lipids in microalgae, *J. Microbiol. Methods*. 77
1055 (2009) 41–47. doi:10.1016/j.mimet.2009.01.001.
- 1056 [78] L.M. Laurens, E.J. Wolfrum, Rapid compositional analysis of microalgae by NIR
1057 spectroscopy, *NIR News*. 23 (2012) 9–11.
- 1058 [79] M.S. Cooper, W.R. Hardin, T.W. Petersen, R.A. Cattolico, Visualizing “green oil” in
1059 live algal cells, *J. Biosci. Bioeng.* 109 (2010) 198–201.
1060 doi:10.1016/j.jbiosc.2009.08.004.
- 1061 [80] J.N. Murdock, D.L. Wetzell, FT-IR Microspectroscopy Enhances Biological and
1062 Ecological Analysis of Algae, *Appl. Spectrosc. Rev.* 44 (2009) 335–361.
1063 doi:10.1080/05704920902907440.
- 1064 [81] I. Dreissig, S. Machill, R. Salzer, C. Krafft, Quantification of brain lipids by FTIR
1065 spectroscopy and partial least squares regression, *Spectrochim. Acta. A. Mol. Biomol.*
1066 *Spectrosc.* 71 (2009) 2069–2075. doi:10.1016/j.saa.2008.08.008.
- 1067 [82] S.M. Renaud, H.C. Zhou, D.L. Parry, L.-V. Thinh, K.C. Woo, Effect of temperature on
1068 the growth, total lipid content and fatty acid composition of recently isolated tropical
1069 microalgae *Isochrysis* sp., *Nitzschia closterium*, *Nitzschia paleacea*, and commercial
1070 species *Isochrysis* sp.(clone T. ISO), *J. Appl. Phycol.* 7 (1995) 595–602.
1071 <http://link.springer.com/article/10.1007/BF00003948> (accessed August 25, 2016).
- 1072 [83] M.A. Borowitzka, L.J. Borowitzka, *Micro-algal biotechnology*, Cambridge University
1073 Press, 1988.
- 1074 [84] H.A. Spoehr, H.W. Milner, THE CHEMICAL COMPOSITION OF CHLORELLA;
1075 EFFECT OF ENVIRONMENTAL CONDITIONS, *Plant Physiol.* 24 (1949) 120–149.
1076 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC437916/>.
- 1077 [85] I.A. Guschina, J.L. Harwood, Lipids and lipid metabolism in eukaryotic algae, *Prog.*
1078 *Lipid Res.* 45 (2006) 160–186. doi:10.1016/j.plipres.2006.01.001.
- 1079 [86] L. Brand, R. Guillard, The effects of continuous light and light intensity on the
1080 reproduction rates of twenty-two species of marine phytoplankton, *J. Exp. Mar. Biol.*
1081 *Ecol.* 50 (1981) 119–132.
- 1082 [87] P.J. le B. Williams, L.M. Laurens, Microalgae as biodiesel & biomass feedstocks: review
1083 & analysis of the biochemistry, energetics & economics, *Energy Environ. Sci.* 3 (2010)
1084 554–590.
- 1085 [88] S. Wahidin, A. Idris, S.R.M. Shaleh, The influence of light intensity and photoperiod on
1086 the growth and lipid content of microalgae *Nannochloropsis* sp., *Bioresour. Technol.* 129
1087 (2013) 7–11. doi:10/f4mb36.
- 1088 [89] L. Rodolfi, G. Chini Zittelli, N. Bassi, G. Padovani, N. Biondi, G. Bonini, M.R. Tredici,
1089 Microalgae for oil: Strain selection, induction of lipid synthesis and outdoor mass
1090 cultivation in a low-cost photobioreactor, *Biotechnol. Bioeng.* 102 (2009) 100–112.
1091 doi:10.1002/bit.22033.
- 1092 [90] E.S. Shuba, D. Kifle, Microalgae to biofuels: ‘Promising’ alternative and renewable
1093 energy, review, *Renew. Sustain. Energy Rev.* 81 (2018) 743–755.
1094 doi:10.1016/j.rser.2017.08.042.
- 1095 [91] M.P. de Souza, M. Hoeltz, P.D. Gressler, L.B. Benitez, R.C.S. Schneider, Potential of
1096 Microalgal Bioproducts: General Perspectives and Main Challenges, *Waste Biomass*
1097 *Valorization*. (2018). doi:10.1007/s12649-018-0253-6.
- 1098 [92] M.-H. Liang, J. Zhu, J.-G. Jiang, High-value bioproducts from microalgae: strategies
1099 and progress, *Crit. Rev. Food Sci. Nutr.* 0 (2018) 01–53.
1100 doi:10.1080/10408398.2018.1455030.
- 1101

1102

1103

1104

1105 7 Figure legends

1106 **Figure. 1** Example of growth curves measured for *Amphora* sp 1 NCC260 by (A) cell count,
1107 (B) fluorometry PAM (F0) and (C) radiometry (NDVI). Points corresponding to cell
1108 concentration according to time. Line curves corresponding to the Gompertz model fitted to
1109 cell concentration as a function of time. n=3, vertical bar = SD. The arrows indicate the late
1110 exponential day estimated by the Gompertz model.

1111 **Figure. 2** Values of the parameters measured for the screened strains. This includes strain
1112 productivity (A), lipid content as measured from the gravimetric method (B), the lipid ratio as
1113 measured semi-quantitatively by the FTIR approaches, $[(eb+CH3+CH3)/si]$ (C) or $[area$
1114 $eb+CH2+CH3)/Total\ area]$ multiplied by 100 for scaling purposes (D). The a, b, c notation
1115 corresponded to the maximum and minimum values for groups 1, 2 and 3. N=3, independent
1116 measurements, \pm SD.

1117 **Figure. 3** Example of averaged FTIR spectra recorded on entire cells or on the corresponding
1118 lipid extract. (A) *Staurosira* sp NCC182 FTIR signature recorded on the entire cells and (B)
1119 *Staurosira* sp NCC182 FTIR signature recorded on a crude lipid extract. The grey area
1120 corresponds to the variation of the FTIR signal associated to the standard deviation for n=3
1121 independent measurements.

1122 **Figure. 4** Correspondence analysis map calculated on the basis of the macromolecular content
1123 as evaluated by FTIR on all the assayed strains of the NCC. N=3 independent measurements.

1124 **Figure 5** Venn diagram showing the degree of overlap among the different approaches used
1125 to identify the lipid rich diatoms. In the gravimetric method circle, 14 strains were identified
1126 as rich in lipids: 4 only with this method (NCC260, NCC208, NCC253, NCC351), one strains
1127 were also identified as rich and lipid by the FTIR on lipid extract (NCC229), and two strains
1128 were also identified as rich in lipids by the FTIR on entire cells (NCC33, NCC199). One strain
1129 was identified by FTIR on lipid extract and FTIR on whole cells (NCC366) and seven strains
1130 were identified as rich in lipids by all three methods (NCC109, NCC182, NCC113, NCC270,
1131 NCC169, NCC303, NCC216).

1132 **Figure 6.** Boxplots summarizing the sample distribution criteria as measured with FTIR
1133 methods, gravimetry for the lipid rate and with fluorimetry for productivity. FTIR data was
1134 expressed in arbitrary units. Lipid rates in %DW and Px in $\text{g}\cdot\text{L}^{-1}\cdot\text{day}^{-1}$. FTIR results
1135 was multiplied by 20 for scaling purposes. The Px was multiplied by 100 for scaling purposes.
1136 33 strains were assayed in independent biological triplicates.

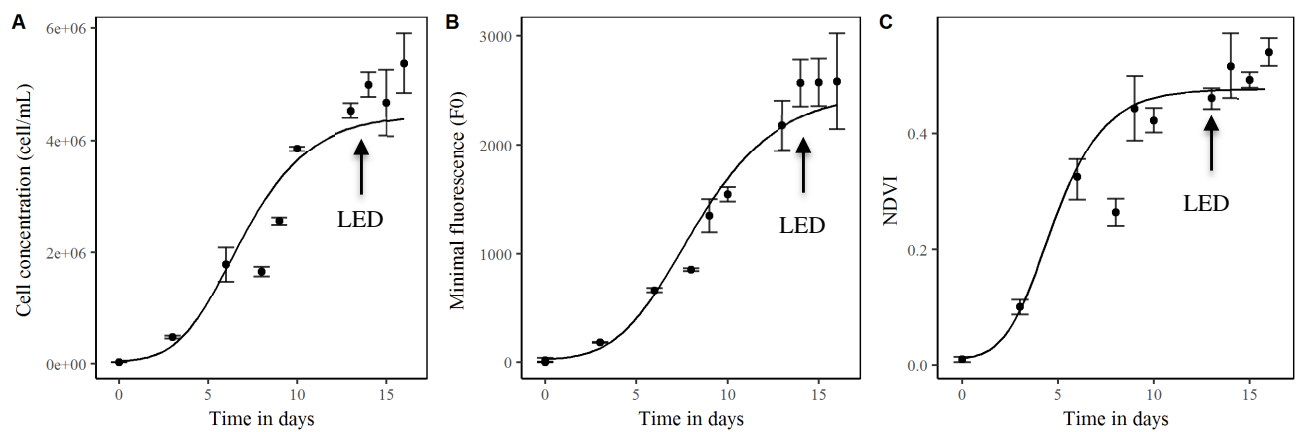
1137
1138
1139

1140 Figures

1141

1142 Figure 1.

1143



1144

1145

1146

1147

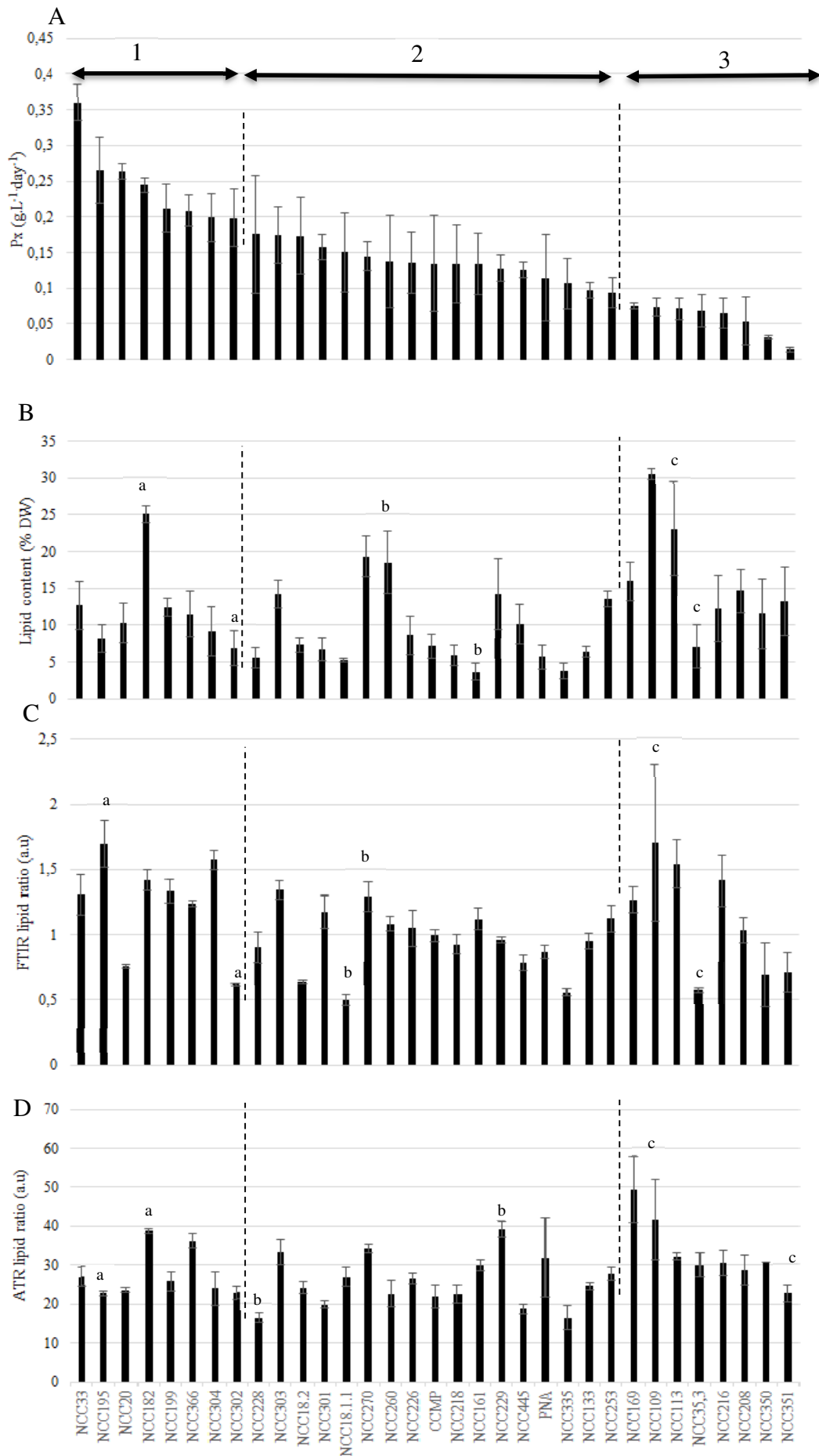
1148

1149

1150

1151

1152



1154

1155 Fig. 3

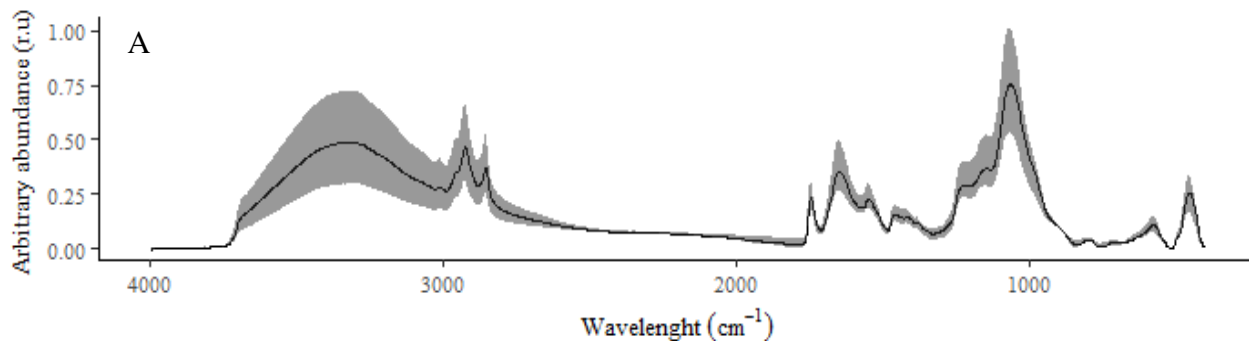
1156

1157

1158

1159

1160



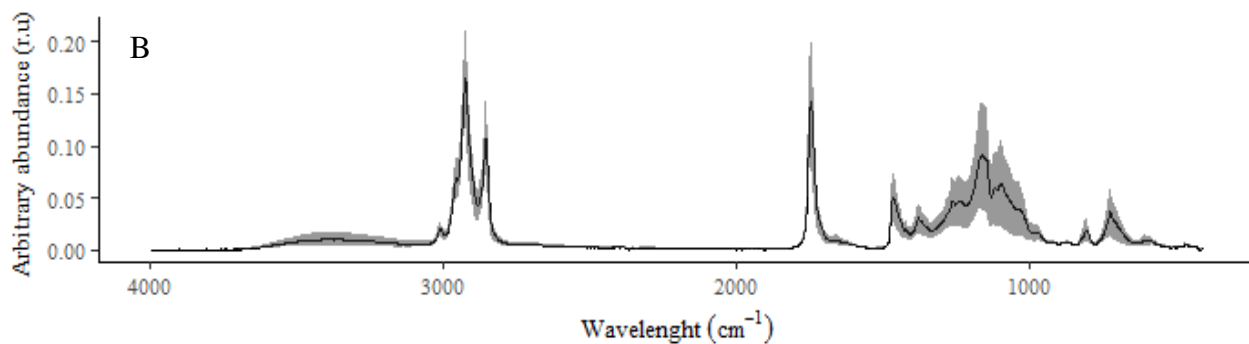
1161

1162

1163

1164

1165



1166

1167

1168

1169

1170

1171

1172

1173

1174

1175

1176

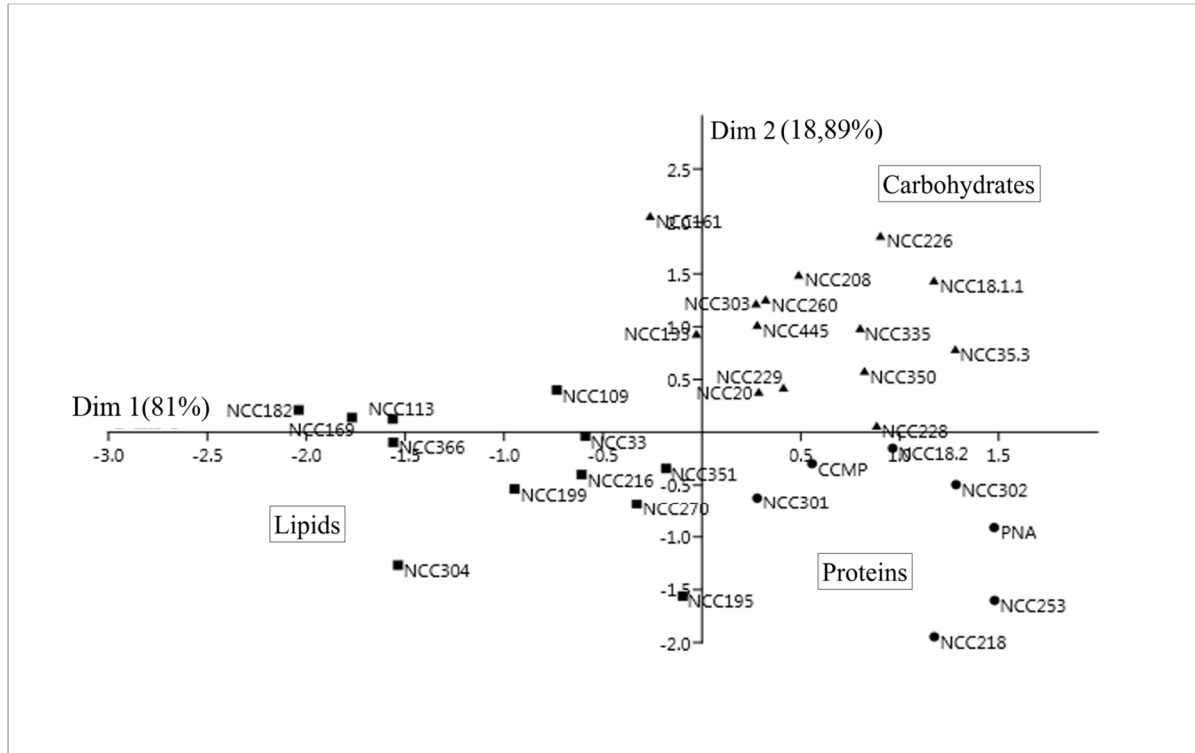
1177

1178

1179 Fig. 4

1180

1181



1182

1183

1184

1185

1186

1187

1188

1189

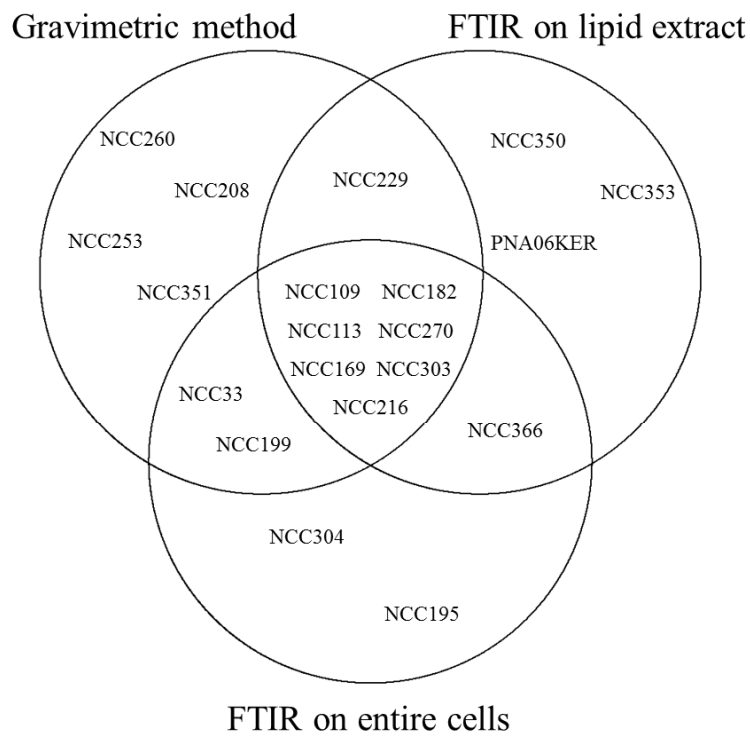
1190

1191

1192

1193

1194



1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208

1209

1210 Fig. 6

1211

1212

1213

1214

1215

1216

1217

1218

1219

1220

1221

1222

1223

1224

1225

1226

1227

1228

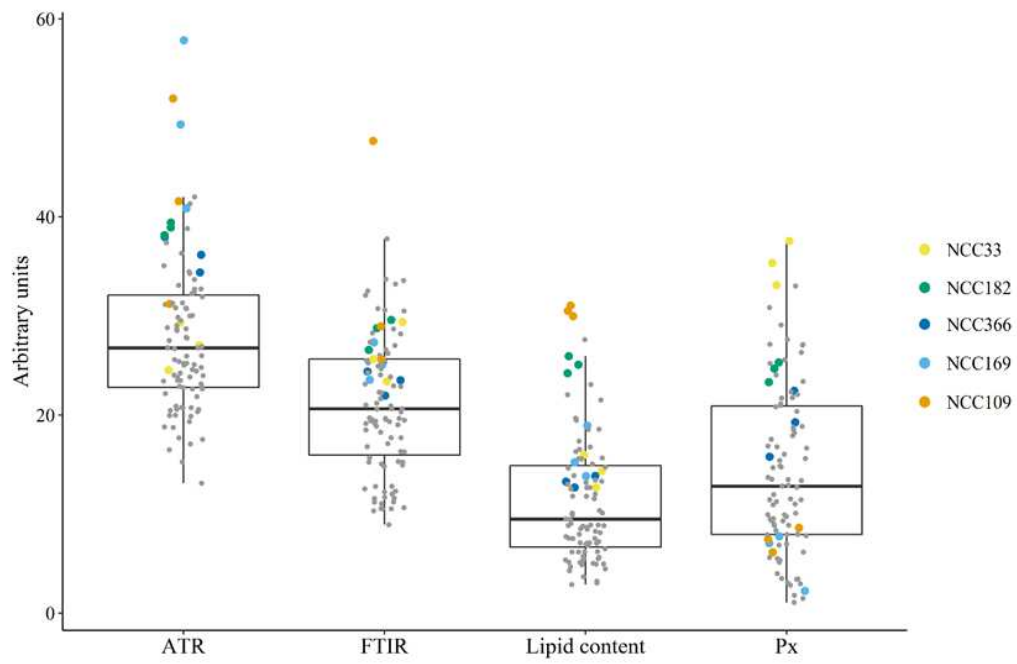
1229

1230

1231

1232

1233



1234

1235 Table S1. Supplementary data. Strains who did not grown.

1236

1237

1238

1239

1240

1241

1242

1243

Species	NCC strain identification	Sampling location
<i>Amphora</i> sp. B2	NCC261	France, NW Atlantic coast
<i>Amphora</i> sp. AC16	NCC410	France, NW Atlantic coast
<i>Amphora</i> sp. AE8	NCC413	France, NW Atlantic coast
<i>Berkeleya rutilans</i>	NCC210.2	France, NW Atlantic coast
<i>Berkeleya rutilans</i>	NCC309	France, Mediterranean sea
<i>Brockmaniella brockmanii</i> 2	NCC403	France, NW Atlantic coast
<i>Caloneis</i> sp. 1	NCC180	France, NW Atlantic coast
<i>Catcombas</i> sp. 1	NCC337	France, Mediterranean sea
<i>Cocconeis scutellum</i> 1	NCC209.1	France, NW Atlantic coast
<i>Cocconeis scutellum</i> 2	NCC209.2	France, NW Atlantic coast
<i>Cocconeis scutellum</i> 2	NCC209.3	France, NW Atlantic coast
<i>Craspedostauros</i> sp. 3	NCC57	France, NW Atlantic coast
<i>Craspedostauros</i> sp. 4	NCC58	France, NW Atlantic coast
<i>Craspedostauros</i> sp. 5	NCC204	France, NW Atlantic coast
<i>Entomoneis alata</i> 1	NCC16	France, NW Atlantic coast
<i>Entomoneis alata</i> 2	NCC448	Portugal, NW Atlantic coast
<i>Entomoneis</i> sp. BAB2	NCC415	France, NW Atlantic coast
<i>Gyrosigma</i> sp. 1	NCC411	France, NW Atlantic coast
<i>Gyrosigma</i> sp. 2	NCC412	France, NW Atlantic coast
<i>Gyrosigma tenuissimum</i>	NCC258	France, NW Atlantic coast
<i>Halamphora coffeaformis</i>	UTCC58	Canada, NW Atlantic coast
<i>Helicotheca tamesis</i> 1	NCC59	, France, Mediterranean sea
<i>Helicotheca tamesis</i> 2	NCC60	France, NW Atlantic coast
<i>Lampriscus</i> sp.	NCC347	France, Mediterranean sea
<i>Leptocylindrus danicus</i> 1	NCC205	France, NW Atlantic coast
<i>Leptocylindrus danicus</i> 2	NCC206	France, NW Atlantic coast
<i>Melosira nummuloïdes</i> 1	NCC25	France, NW Atlantic coast
<i>Melosira nummuloïdes</i> 2	NCC25.1	France, NW Atlantic coast
<i>Navicula</i> sp. Z4	NCC224	France, NW Atlantic coast
<i>Navicula</i> sp. e1	NCC269	France, NW Atlantic coast
<i>Navicula cf ramosissima</i>	NCC449	France, NW Atlantic coast
<i>Nitzschia laevis</i>	NCC39	France, NW Atlantic coast
<i>Nitzschia salinicola</i>	NCC41	France, NW Atlantic coast
<i>Nitzschia</i> sp. B4	NCC114	France, NW Atlantic coast
<i>Opehora</i> sp. 2	NCC365	France, NW Atlantic coast
<i>Paralia sulcata</i>	NCC177	France, NW Atlantic coast
<i>Pleurosigma</i> sp. K	NCC339	Ukraine, Black sea
<i>Pleurosigma</i> sp. LM	NCC404	France, NW Atlantic coast
<i>Pleurosigma</i> sp. BC1	NCC423	France, NW Atlantic coast
<i>Pleurosigma</i> sp. BC7	NCC425	France, NW Atlantic coast
<i>Pleurosigma</i> sp. BC15	NCC428	France, NW Atlantic coast
<i>Rhizosolenia setigera</i>	NCC127	France, NW Atlantic coast
<i>Tabularia tabulata</i>	NCC338	France, NW Atlantic coast

Table S2: Lipid quantification evaluated by FTIR and ATR method for all the species.

Species	code	FTIR					ATR	
		Carb/si	Prot/si	Eb/si	CH2/si	CH3/si	eb	CH2 + CH3
<i>Amphora acutiuscula</i>	NCC216	53.6 ± 8.2	45.8 ± 5.7	22.1 ± 4.9	65.7 ± 8.0	53.5 ± 7.8	6.2 ± 0.9	24.3 ± 2.4
<i>Amphora</i> sp. 1	NCC260	63.6 ± 2.9	40.0 ± 2.0	12.2 ± 2.3	57.3 ± 3.3	39.2 ± 0.6	4.5 ± 0.7	18.1 ± 3.1
<i>Amphora</i> sp. 2	NCC169	58.1 ± 4.3	42.4 ± 1.1	17.2 ± 3.4	59.7 ± 3.4	49.7 ± 3.0	7.1 ± 0.7	42.1 ± 11.0
<i>Brockmaniella brockmanii</i>	NCC161	51.2 ± 5.5	25.2 ± 6.4	12.8 ± 4.0	57.7 ± 5.5	41.8 ± 5.5	4.1 ± 1.6	21.1 ± 5.6
<i>Conticriba weissflogii</i>	NCC133	47.4 ± 0.5	40.2 ± 0.8	5.03 ± 1.5	48.4 ± 2.6	41.7 ± 1.9	4.7 ± 0.3	20.0 ± 0.7
<i>Conticriba weissflogii</i>	CCMP1336	48.3 ± 2.8	43.5 ± 3.3	9.8 ± 1.8	48.7 ± 3.0	40.9 ± 2.7	5.0 ± 0.7	16.9 ± 2.2
<i>Craspedostauros britannicus</i>	NCC195	60.7 ± 2.4	69.8 ± 11	23.2 ± 3.3	80.7 ± 8.0	66.0 ± 6.9	4.8 ± 0.4	19.2 ± 1.8
<i>Craspedostauros britannicus</i>	NCC199	55.1 ± 3.4	47.5 ± 5.7	26.9 ± 3.4	62.2 ± 2.7	44.2 ± 3.3	5.4 ± 0.6	20.5 ± 1.9
<i>Craspedostauros</i> sp. 1	NCC228	44.3 ± 0.1	37.8 ± 4.3	6.2 ± 0.4	46.3 ± 6.4	37.8 ± 6.0	3.9 ± 0.8	15.2 ± 3.4
<i>Craspedostauros</i> sp. 2	NCC218	54.6 ± 2.6	49.4 ± 3.5	7.5 ± 0.3	46.0 ± 3.8	39.0 ± 4.0	4.7 ± 0.3	17.9 ± 2.5
<i>Cymatosira belgica</i>	NCC208	41.6 ± 4.8	25.4 ± 3.3	6.7 ± 1.1	55.4 ± 5.1	41.5 ± 3.7	5.9 ± 0.7	22.7 ± 3.2
<i>Entomoneis paludosa</i>	NCC18.1.1	39.6 ± 0.3	26.0 ± 2.7	2.9 ± 0.1	27.3 ± 2.4	19.4 ± 1.8	3.9 ± 0.7	23.1 ± 1.8
<i>Entomoneis paludosa</i>	NCC18.2.1	45.7 ± 0.6	40.9 ± 1.4	6.1 ± 0.1	34.5 ± 0.9	22.5 ± 0.6	3.1 ± 0.4	21.6 ± 1.1
<i>Entomoneis</i> sp. 1	NCC350	39.9 ± 0.1	30.4 ± 0.8	5.5 ± 1.5	28.6 ± 0.9	21.2 ± 1.7	2.7 ± 0.1	30.6 ± 4.3
<i>Entomoneis</i> sp. 2	NCC20	42.5 ± 0.5	32.5 ± 0.3	9.5 ± 0.3	40.0 ± 0.4	26.0 ± 0.6	4.7 ± 0.3	18.6 ± 0.4
<i>Entomoneis</i> sp. 3	NCC351	45.5 ± 2.9	41.4 ± 4.2	14.4 ± 3.4	37.2 ± 0.8	28.3 ± 0.6	3.9 ± 1.7	15.8 ± 5.6
<i>Entomoneis</i> sp. 4	NCC301	62.5 ± 4.8	56.7 ± 6.4	12.5 ± 2.5	64.0 ± 7.9	46.7 ± 5.6	2.9 ± 0.1	17.0 ± 0.8
<i>Entomoneis</i> sp. 5	NCC302	37.6 ± 0.7	36.8 ± 0.8	3.43 ± 0.1	33.9 ± 0.7	23.9 ± 0.5	3.3 ± 0.3	19.6 ± 1.5
<i>Entomoneis</i> sp. 6	NCC335	50.5 ± 0.6	35.3 ± 2.7	6.6 ± 0.7	29.9 ± 1.9	19.0 ± 1.1	3.1 ± 0.2	13.4 ± 3.0
<i>Entomoneis</i> sp. 7	NCC445	40.6 ± 0.9	27.0 ± 2.7	8.5 ± 0.2	37.7 ± 0.4	29.2 ± 0.3	4.1 ± 0.9	17.5 ± 4.3
<i>Extubocellulus cf cribriger</i>	NCC229	49.8 ± 3.7	38.1 ± 2.6	10.1 ± 1.2	48.0 ± 1.9	38.0 ± 1.1	4.9 ± 0.3	34.2 ± 1.7
<i>Fallacia</i> sp. 1	NCC303	64.7 ± 1.8	74.4 ± 7.0	14.2 ± 2.5	64.5 ± 4.1	55.5 ± 5.4	6.9 ± 0.3	26.4 ± 2.9
<i>Fallacia</i> sp. 2	NCC304	46.0 ± 6.2	45.7 ± 3.2	31.3 ± 4.0	72.8 ± 2.7	53.3 ± 2.1	5.1 ± 1.0	18.9 ± 3.4
<i>Licmophora</i> sp. 1	NCC253	53.1 ± 2.7	65 ± 10	4.4 ± 1.1	57.9 ± 5.4	50.3 ± 5.2	3.6 ± 0.3	22.6 ± 0.6
<i>Lithodesmium</i> sp	NCC353	34.2 ± 2.0	25.8 ± 0.8	2.4 ± 0.3	31.1 ± 1.1	23.1 ± 0.4	2.1 ± 0.3	28.0 ± 3.2
<i>Navicula</i> sp. 1	NCC113	63.3 ± 4.7	43.8 ± 8.9	33.4 ± 4.7	68.2 ± 9.6	52.4 ± 5.6	3.4 ± 0.4	28.8 ± 1.2
<i>Navicula</i> sp. 2	NCC226	54.2 ± 4.3	27.0 ± 3.0	5.2 ± 0.9	53.1 ± 3.8	39.5 ± 2.8	4.2 ± 0.7	22.4 ± 0.7
<i>Nitzschia</i> sp. 5	NCC109	62.0 ± 2.8	42.4 ± 6.7	22.0 ± 2.3	72.7 ± 4.6	61.3 ± 4.8	6.9 ± 0.4	45.1 ± 28.9
<i>Nitzschia alexandrina</i>	NCC33	58.2 ± 7.2	44.0 ± 4.6	20.7 ± 3.3	56.1 ± 4.0	45.8 ± 0.7	5.7 ± 0.3	21.3 ± 2.3
<i>Opephora</i> sp. 1	NCC366	58 ± 12	40.1 ± 8.8	31.6 ± 6.1	50.6 ± 4.5	42.0 ± 2.5	6.9 ± 1.1	26.3 ± 4.4
<i>Pseudonitzschia americana</i>	PNA06 KER	56.6 ± 1.4	60.6 ± 3.9	3.9 ± 1.0	44.0 ± 2.8	39.1 ± 3.2	4.0 ± 0.3	38 ± 28
<i>Staurosira</i> sp.	NCC182	47.7 ± 6.6	30.7 ± 4.1	30.9 ± 5.4	61.6 ± 4.0	49.0 ± 2.3	7.3 ± 0.1	31.6 ± 0.6
<i>Surirella</i> sp. 1	NCC270	58.5 ± 2.4	54.4 ± 8.4	22.0 ± 4.6	56.1 ± 5.3	51.3 ± 5.4	7.1 ± 0.4	27.1 ± 0.6