Studies on the Metabolism of Lipids and Energy Storage Compounds to Elucidate the Regulation of Alkenone Biosynthesis in a Haptophyte alga *Emiliania huxleyi*

September 2017

BAKKU RANJITH KUMAR

Studies on the Metabolism of Lipids and Energy Storage Compounds to Elucidate the Regulation of Alkenone Biosynthesis in a Haptophyte alga *Emiliania huxleyi*

A Dissertation Submitted to

the Graduate School of Life and Environmental Sciences,

the University of Tsukuba

in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy

(Doctoral Program in Integrative Environment and Biomass Sciences)

BAKKU RANJITH KUMAR

Abbreviations

14:0	tetradecanoic acid
15:0	
16:0	palmitic acid
17:0	
18:0	stearic acid
18:1n-9	oleic acid
18:2n-6	α-linoleic acid
18:5n-3	octadecatetraenoic acid
20:4n-3	eicosatetraenoic acid
20:5n-3	eicosapentaenoic acid
22:5n-3	docosapentanoic acid
22:6n-3	docosahexanoic acid,DHA
AP	acid polysaccharide
BLL	beatine like lipids
CE	collision energy
CUR	curtain gas
CXP	collision exit potential
DAG	diacylglycerol

DGCC	diacylglyceryl-3-O-carboxyhydroxymethylcholine
DGDG	digalactosylglyceride
DGTS	diacylglyceryltrimethylhomoserine
DNPPE	dinitrophenyl-phosphatidylethanolamine
DP	de-clustering potential
Dw	deionized water
ESI	electron spray ionization
FAs	fatty acids
GC-FID	gas chromatography with flame ionization detection
hGSL	host specific glyco sphingo lipids
HPLC	high performance liquid chromatography
IST	internal standard
LC	liquid chromatography
LMCs	low molecular weight compounds
MCFAs	medium chain saturated, mono- and di-unsaturated fatty acids
MeOH	methanol
MGDG	monogalactosyldiglyceride
MPIS	multiple precursor ion scan
MS	mass spectrometry

–N	N-sufficient
+N	N-limitation
NL	neutral los scan
NLF	neutral lipid fraction
NP	neutral polysaccharide
PA	phosphatidic acid
PC	phosphatidyl choline
PDPT	phosphatidyl-S,S-dimethylpropanethiol
PE	phosphatidyl ethanolamine
PERL	practical extraction and reporting language
PG	phosphatidyl glycerol
Pi	inorganic phosphate
PI	phosphatidyl ionositol
PIS	precursor ion scan
PLF	polar lipid fraction
PS	phosphatidyl serine
PUFAs	polyunsaturated fatty acids
QTRAP	triple quadrapole
sGSL	monosaccharide head group with salicyllic

SQDG	sulfoquinovosyl diacylglycerols
TAGs	triacylglycerols
TEM	Source temperature
TIC	total ion chromatogram
$U^{K'}_{37}$	Unsaturation index
XIC	extracted ion chromatogram

Abstract

Emiliania huxleyi is a dominant bloom-forming coccolithophore. It is one of the five alkenone-producing haptophytes and is a major contributor to the global biogeochemical cycles in the cosmopolitan marine environment. Alkenones are unique lipids having long chain methyl ketones (C_{37}-C_{40}) with 2-4 *trans* double bonds. The numbers of the *trans* unsaturation bonds in the molecules indicate the temperatures at which the molecules were synthesized by the organisms, thus, it is very useful for estimating temperature in the paleoenvironment. These molecules are also used for estimating paleo pCO_2. For these reasons, alkenones are considered valuable biomarkers for geochemical studies. The long-chain compounds are also found to be useful as a feedstock for biofuels. Despite these applications, the biosynthesis pathways, physiological functions, and metabolic profile of the alkenones are not well-known. The *trans* unsaturation of alkenones is very similar to that of *cis* unsaturation of the membrane lipid fatty acids, with regard to changes in temperature. Therefore, alkenones are assumed to be structural components. However, recent studies reveal that the accumulation characteristics of alkenones are more similar to neutral lipids like triglycerides (TAGs). Also, as alkenone producers accumulate relatively less amounts of TAGs, alkenones are assumed to be storage lipids. In addition to alkenones, it is observed that *E. huxleyi* also distributes a major portion of its carbon into several photosynthetic components such as low-molecular-weight compounds (LMCs), β-glucan (a neutral polysaccharide, NP) and acid polysaccharides (AP). Having such a variety of carbon (C) storage components, limits the understanding of C allocation mechanism and lipid metabolism in *E. huxleyi*. Therefore, we are interested in (1) elucidating how alkenones and other lipids contribute to energy storage, carbon

fixation, and cellular carbon-partitioning in relation to other cellular components, and (2) establishing a comprehensive lipidomic analysis method for future exploration of the *E.huxleyi* lipid metabolism.

Part 1, Elucidation of carbon allocation mechanism into alkenones and other energy storage components under nitrogen limitation

The general C-metabolism in algae is highly regulated by nitrogen (N)-availability. Under the N-limitation, most of the algae redirect the fixed carbon to TAG biosynthesis or carbohydrate biosynthesis. Therefore, I expected to observe; (i) a clear difference in C-allocation into various metabolites when cells are exposed to sudden changes in N-nutrition, namely from +N to –N, and (ii) to know the mechanism by which *E. huxleyi* switches the direction of C-flow into various macromolecular pools functioning as the significant carbon sinks. Moreover, no report was available on C-allocation profiles in alkenone-producing haptophytes when cells were exposed to N-limiting conditions (–N) from N-sufficient conditions (+N) by changing nitrate availability in the culture. Also, no studies were reported to interpret how photosynthetic carbon metabolism on C-allocation is changed, especially, in the accumulation into alkenones and C-distribution among various cell components in non-coccolith-bearing strain of *E. huxleyi* (CCMP 2090). To understand the mechanism of energy storage, I analyzed the changes in C-allocation among various cell components like lipids, alkenones, proteins and polysaccharides between cells exposed to N-sufficient (+N) and N-limited conditions (–N) in *E. huxleyi* CCMP 2090. Finally, the alkenone was found to function as the main storage lipids and its accumulation was clearly increased by –N whereas triacylglycerols (TAGs) were barely detected under

any N-conditions. The mobilization of carbons into alkenones was stimulated by –N from 15% under +N to 27% under –N. However, photosynthetic C-allocation into other components was suppressed by –N, showing that percent C-allocation into fatty acids, proteins and polysaccharides were decreased from 9%, 46% and 6.8% under +N to 7%, 25% and 4.5% under –N, respectively. Although the plastid-located fatty acids such as palmitic acid (16:0), stearic acid (18:0), oleic acid (18:1) and linoleic acid (18:2) became dominant under –N, while octadecapentaenoic acid (18:5) became dominant under +N conditions, the contents of the major fatty acid in the non-plastid phospholipids, docosahexaenoic acid (DHA, 22:6), were not significantly altered. A large fraction of unallocated carbon "others" was also observed in this study. I hypothesize that this might be composed of LMCs which include mannitol. To test this, I used Fourier Transformed Infrared Spectroscopy (FTIR) to observe the overall changes in macromolecular carbon pools. The results showed increase in carbohydrate related signals indicating accumulation of LMCs. This finding supports my hypothesis. From these analyses, I also observed that novel FTIR absorption at 1705.5 cm^{-1} increases with alkenones accumulation especially under the –N conditions. Therefore, this peak can be used as a novel indicator of –N limitation in alkenone producers. Further, I am attempting to establish an FTIR method to estimate polysaccharides in *E. huxleyi* and semi quantification of major cellular components based on alkenone content.

Part 2, Establishment of lipidomics methods and lipidomics of *Emiliania huxleyi* CCMP 2090

To explore the lipid metabolism in detail, I focused on developing methods to examine complete lipidome of this organism. For this, I developed a shotgun electron spray ionization mass spectrometry (ESI-MS) approach for detecting all lipid classes including *E. huxleyi* specific lipids. Later, I achieved major lipid class separation in a single 50-60 min run using a normal phase BETASIL DIOL column. I analyzed the lipidome of the logarithmic and stationary phase cultures of *E. huxleyi* CCMP 2090 was analyzed and identified nearly 600 lipid species under both conditions. Qualitative and quantitative analyses for the *E.huxleyi* lipidome was performed using in-house build PERL-based computational algorithms. From the lipidomic analysis of the log and stationary phase batch cultures I observed that TAGs are produced in this strain. However, the contents of TAGs were very low (2%mol of the total lipidome) under both conditions. Therefore, I assume that the TAG accumulation was highly suppressed in this organism under all conditions. On the other hand, phosphatidyl choline (PC) was the major lipid (65-70%mol) in both conditions. Whereas, regarding to glycolipids, monogalactosyl diacylglycerol (MGDG) was around 12-17%, digalactosyl diacylglycerol (DGDG) and sulfoquinovosyl diacylglycerol (SQDG) were 2-3% each. In addition, the composition of other particular lipids like phosphatidyldimethylpropanethiol (PDPT) and betaine lipids, (Diacylglyceryl-3-O-carboxyhydroxymethylcholine (DGCC), Diacylglyceryl trimethylhomoserine (DGTS) and betaine-like lipids (BLL)) were only 1-2% each in the total lipidome. I also found very long-chain fatty acids like tetracosatetraenoic acid (24:4), tetracosaopentaenoic acid (24:5) and tetracosahexaenoic acid (24:6), indicating presence of the mammalian-type DHA biosynthesis pathway in *E. huxleyi*. These results indicate that *Emiliania* is more dependent on the phopsholipids rather than glycolipids.

Therefore, the polar lipid metabolism might be highly influenced by the availability of phosphates rather than nitrates as nutrients. This result explains why the polar lipid fraction under the N-limited condition did not change so significantly. However, a detailed analysis of lipidomics is needed to examine the dynamic changes in the fatty acid composition under the N-stress. In summary, the established lipidomic technique may be useful for analyzing lipidic components of almost all major lipid classes produced by marine algae. Especially, this study could be beneficial for future investigations on the lipids and alkenone biosynthetic pathways and their regulation in the alkenone-producing marine haptophyte.

General Introduction

Motivation

Global climate change is a subject of considerable public and scientific interest today. It is now widely agreed that climate change is directly related to increased carbon emissions caused by burning fossil fuels for every increasing energy demand (Richard 2012). Addressing climate change and transitioning into a more sustainable renewable fuel based economy are currently two major challenges facing humanity. Understanding paleo-climates is essential to address challenges posed by climate change. Photosynthetic marine microalgae play a role in studying paleo-climate as well as in bio-fuel research. They are major contributors in global biogeochemical cycles (Holligan et al 1993; Falkowski 1994). Algae utilize C from atmosphere and photosynthetically fix them in the form of various metabolites (Lloyd 1977; Blankenship and Hartman 1998; Bendall et al. 2008). Some of these algal products are used as biomarkers in biogeochemical and geophysical studies (Müller et al. 1998; Prahl and Wakeham 1987). Chemical signatures (isotope ratios, structure, and unsaturation etc.) of these molecules provide essential information on the type of environment in which the organism existed leading to understanding of paleo-climates. On the other hand, lipids produced by these organisms are applicable as alternatives to crude oils (Hannon et al. 2010). Therefore, the photosynthetic products from marine prokaryotic photoautotrophs such as cyanobacteria and eukaryotic algae are of high significance. Alkenones are one such compound produced by specific marine Haptophyta, useful as both biomarkers for climate studies as well as biofuels (Müller et al. 1998; Prahl and Wakeham 1987; Wu et al. 1999; O'Neil et al. 2015).

Alkenones and their distribution

Alkenones are lipid like compounds comprising of very long carbon chain (mainly C_{37}-C_{39}) with a methyl or ethyl ketone group (Volkman et al. 1980a, b; Marlowe et al. 1984). A unique feature of these compounds is the presence of 2 to 4 *trans*-type carbon-carbon double bonds (Fig. 1). Along with alkenones, similar types of such compounds such as alkenoates (ketone group of alkenones altered to methyl ester) were initially identified in marine sediments. All these compounds are known to be produced by only five genera (*Emiliania, Gephyrocapsa, Isochrysis, Tisochrysis* and *Chrysotila*) of marine haptophytes (Conte et al. 1994; Volkman et al. 1995; Sawada et al. 1996; Rontani and Volkman 2004; Nakamura et al. 2014). Several types of these alkenones (Fig. 1) like C_{37} methyl ketones and C_{38}-C_{39} ethyl ketones with two to three carbon double bonds were detected from algae such as *Emiliania* and *Gephyrocapsa*, waters and sediments in marine and lacustrine environments (Conte et al. 1994; Cranwell 1985; Marlowe et al. 1984; Sawada et al. 1996; Volkman et al. 1980a, b).

Furthermore, the distribution of some alkenones was identified to be specific to certain settings, like $C_{38:4}$ in cold water sediments (Marlowe et al. 1984), $C_{37:4}$ in haptophyte alga *Chrysotila*, waters and sediments in sulfate-rich lakes (Nakamura et al. 2014; Rontani and Volkman 2004; Theissen et al. 2005; Theroux et al. 2010; Toney et al. 2010), C_{35}-C_{36} in the Black Sea sediments (Xu et al. 2001), C_{40}-C_{41} in hyper-saline environments (Zhao et al. 2014), etc. These evidences indicate that the composition, distribution and the un-saturation number of alkenones and alkenoates are varied depending on algal species, habitats, growth temperature and so on.

Importance of alkenones as biomarkers and biofuels

Alkenones are found to be useful in two major research areas because of their unique features. Their application as biomarkers is due to their role in responding to temperature changes. A close correlation between the un-saturation degree of alkenones and sea surface temperatures was identified mainly in C_{37} alkenones (Brassell 1986; Müller et al. 1998; Prahl and Wakeham 1987). The changes in alkenone unsaturation are determined by unsaturation index. The C_{37} alkenone unsaturation indexes are defined by the equation given as U^{K}_{37} = ([$C_{37:2}$Me] – [$C_{37:4}$Me]) / ([$C_{37:2}$ Me] + [$C_{37:3}$Me] + [$C_{37:4}$Me]) and $U^{K'}_{37}$ = [$C_{37:2}$Me] / ([$C_{37:2}$Me] + [$C_{37:3}$Me]). These ratios seem to vary linearly with sea surface temperatures. Not only U^{K}_{37} but, various alternative unsaturation indexes were developed as the alkenone distribution in lakes differs from marine realm (Rontani and Volkman 2004; Theroux et al. 2010; Toney et al. 2010). The index has been used for estimating paleo-temperature when alkenone-producing haptophytes were distributed in geochemical studies (Brassell et al. 1993; Popp et al. 1998). In addition, the $^{13}C/^{12}C$-isotopic carbon partitioning value in alkenone molecules is found to be useful for estimating ancient CO_2 concentrations (Freeman and Hayes 1992; Jasper and Hayes 1990, Pagani et al. 2002; Riebesell et al. 2000). Alkenones have become the most important lipidic biomarkers in geochemical and geophysical researches.

Besides being widely accepted as biomarkers, alkenones are also assumed to be potential biofuel precursors. Initially, it was observed that 10−20% of cell C during the stationary phase accumulates into alkenones and more abundant than TAGs in these organism (Eltgroth et al. 2005; Volkman et al. 1989). Further, alkenones are found to

comprise nearly 75% of lipids in cytoplasmic lipid bodies (Shi et al. 2015) indicating their role as fuel bodies of cells. Wu et al. (1999) showed that, when *E. huxleyi* and *G. oceanica* are subjected to pyrolysis at 100°C to 500°C *n*-alkanes and *n*-alkenes are produced. Therefore, alkenones are expected to be molecular sources for these liquid saturated hydrocarbons. It is also observed that alkenones can be converted to jet fuel range hydrocarbons (C_{10}-C_{17}) by techniques like butenolysis (O'Neil et al. 2015, 2016). Having less unsaturation number and absence of glycerol backbone is an advantage of these molecules. On the other hand, common biodiesel precursors like TAGs are polyunsaturated and require treating byproducts like glycerol during biofuel production. Therefore, use of alkenones as biofuels could be advantageous over TAGs. However, an industry level production is yet to be developed from the knowledge of biosynthesis mechanisms of these alkenones.

Biological role and biosynthesis of alkenones

Despite many studies on such features and applications of alkenones for paleotemperature reconstruction in organic geosciences, the biological function of alkenones in haptophytes is still unclear. The property of temperature dependent changes in the number of *trans*-type unsaturation of alkenones is similar to that of *cis*-type unsaturation of fatty acids (FAs) in membrane lipids. Therefore, Brassell et al. (1986) speculated that similar to membrane lipids, alkenones are used to maintain fluidity and rigidity of cellular membranes. Several other studies on the characteristics of alkenones speculate their biological role as buoyancy regulators or structural backbones (Prahl et al. 2003; Volkman et al. 1980b; Fernández et al. 1996; Sawada and Shiraiwa 2004).

However, recent studies suggest that alkenones can be considered as storage lipids. For example, alkenones are observed to be accumulated under nutrient-deficient stresses, to form lipid droplets in the cells, and to be depleted under dark conditions (Eltgroth et al. 2005; Epstein et al. 1998; Prahl et al. 2003; Shi et al. 2015; Tsuji et al. 2015). Further, one of the haptophytes such as *E. huxleyi* contains more alkenones, but very low amount of TAGs which are known as storage lipids in many microalgae (Bell and Pond 1996; Hunter et al. 2015). Their role as storage lipids like triacylglycerol TAG is highly under debate and yet to be explored.

On the other hand, the composition and distribution of alkenones are highly influenced by genetic background of strains, non-thermal factors (ex. CO_2 concentration, nutrient limitation, light availability, salinity etc.) and physiological state of cells (Benthien et al. 2007; Conte et al. 1995; Conte et al. 1998; Epstein et al. 2001; Pan et al. 2014; Popp et al 1998; Prahl et al. 2003; Sorrosa et al. 2003; Ono et al. 2012; Yamamoto et al. 2000). Therefore, due to the ambiguity in their possible biological roles and the influence of several non-thermal factors on alkenone production, physiological functions and metabolic profiles of alkenones are still unclear and need to be elucidated.

Regulation of C-storage mechanism by nitrogen availability

In the photosynthetic organisms, under the standard growth conditions (+N, nitrogen sufficient) the assimilated N from NO_3^-, is utilized by the glutamine synthetase (GS) and glutamate synthase (GOGAT) mechanism (Fernandez and Galvan 2007; Kaffes et al. 2010; Rokitta et al. 2014; Turpin 1991; Weger and Turpin 1989). This incorporates N into nitrogenous organic compounds. The synthesis of amino acids and proteins requires carbon skeleton, which is acquired by CO_2 fixation (Johnson and Alric

2013; Weger and Turpin 1989). C fixed through photosynthesis is simultaneously used for carbohydrate, lipid synthesis and other metabolite synthesis. When there is sufficient N, more C is allocated into proteins and rest is distributed into lipids and carbohydrates (Johnson and Alric 2013; Kaffes et al. 2010) as shown in Fig. 2.

Under –N, with no sufficient N to make new proteins, cells degrade unwanted proteins and reallocate N for necessary proteins (Alipanah et al. 2015; Bai et al. 2016). Many photosynthetic organisms cannot continue functioning in this condition and C fixation in photosynthetic apparatus declines to some extent. This is due to decrease of several proteins associated with the photosynthetic apparatus (PSII) by light energy (photo-inactivation) and PSII cannot be repaired due to insufficient N (–N, nitrogen limited). This results in decrease of total protein and decrease in C content of proteins. During this condition as cell experiences –N, the C:N balance is shifted towards accumulated C (Talmy et al. 2014; Kaffes et al. 2010; Palmucci et al. 2011). This forces the cell to direct the accumulating C from C-fixation and degraded proteins into storage components. As a result components like carbohydrates/storage polysaccharides and neutral lipids like TAGs that do not contain N (Fig. 2) accumulate under –N (Ball et al. 1990; Benavente-Valdésa et al. 2016; Reitan et al. 1994).

Despite, alkenones being assumed to be similar to TAGs, such C-allocation mechanism is not yet understood in alkenone-producing marine Haptophytes. Part-I of this Ph. D. thesis is the major section, which mainly focus on this issue. In this study I subjected one of the strains of *E. huxleyi* to –N and examined C-allocation patterns into various storage components. This study provided a detailed map of C-regulation under +N and –N conditions. Further, this study also provided information on the alkenone

accumulation patterns and how their characteristic features are influenced by –N.

Methods for lipidomics and significance

Part-II of this Ph. D. thesis is related to establishment of methods for lipidome analysis using sensitive mass spectrometric techniques. In relation to studies on unique lipids like alkenones, there is an urgent need to identify their biosynthetic pathways. To detect such mechanisms, it is necessary to understand lipid molecular compositions and their dynamics during cell growth. Lipidome of organisms is affected significantly under several stress conditions. Especially, major changes occur in the membrane and neutral lipids which are responsible for structural stability and energy storage in the cells. To clearly examine these changes, highly sensitive and robust methods are needed. Lipidomics mainly focus on development of multiple techniques to identify various lipids, their constituents and finally quantify them. This is a rapidly expanding research field.

By recent advances in various mass spectrometric techniques lipid detection is made easy. Lipid detection is based on ionization of lipids, molecular ion generation and fragmentation. Hard ionization techniques which produce extensive lipid fragments were initially used for this purpose. However, this is a disadvantage while analyzing multi component lipid mixtures. Lipid detection through ionization techniques were greatly improved from hard ionization (chemical ionization, fast atom bombardment, Electron impact ionization etc.) to less fragment generating soft ionization techniques (electron spray ionization-ESI, Matrix assisted laser desorption-MALDI). ESI is an emerging and efficient technique mainly used for lipidomics (Han and Gross 2003, 2005; Taguchi et al. 2005). The basic principle of this method lies in ionization of lipids

by charge desorption technique (Fenn et al. 1989, 1990). Here, liquid sample is sprayed into fine charged droplets using a charged nozzle. This enables subsequent desolvation of liquid surface to form charged lipid molecular ions. This is more advantageous lipid ionization method as it causes fewer fragmentations in the lipids. These are further detected by various mass analyzers.

There are several types of detection methods for ionized lipids, popularly known as mass analyzers. MS analyzers such as ion trap (IT), quadrupole, time of flight (TOF), and Fourier transform ion cyclotron (FT-ICR) differ in their mass accuracy (Milne et al. 2006). Among the most widely used mass spectrometers, triple quadrupole (TQ) tandem-MS analyzers provide detection of both molecular and fragment ions (Brugger et al. 1997; Han and Gross 2003, 2005). The ionization and fragmentation of lipids using these techniques are extensively reported for major lipid classes (Brugger et al. 1997; Ejsing et al. 2006; Taguchi et al. 2005).

Direct injection of lipid mixture might affect detection of less abundant lipids due to ion suppression and ion enhancement effects (Han et al. 2011). In order to prevent this, lipid separation methods like gas chromatography (GC), thin layer chromatography (TLC), liquid-chromatography (LC) etc. were applied prior to the MS analysis. Among which, LC techniques are latest in lipidomics and most frequently used in combination with MS (Cjaka and Fiehn 20014; Ikeda et al. 2011; Hummel et al. 2011; Knittelfelder et al. 2014; Sturt et al. 2004). LC in combination with MS has allowed lipids to be studied with greater sensitivity and specificity. Several computations tools are also improved simultaneously to examine MS data. Softwares like LipidBlast, Lipid Data Analyzer, LipidXplorer, MZmine, SIM-Lipid, Lipidpro,

Lipidminer, LipidView etc. are most commonly used for this purpose (Kind et al. 2013; Hartler et al. 2011; Herzog et al. 2012; Pluskal et al. 2010; Premier Biosoft; Ahmed et al. 2015; Meng et al 2014; AB SCIEX Co.). All these tools mainly focus on annotation and support quantitative characterization of lipid species based on their MS information.

Availability of several methods on lipid detection using ESI-MS, lipid separation using HPLC and data analysis tools make it difficult for a naive user to work on lipidomics. In addition, as technology is advancing many novel lipid classes are being discovered, the available methods should also be improved to separate and detect the lipid from complex lipid mixtures. At the downstream of lipidome data analysis, computational methods to process the detected lipids for various studies are not available. Therefore, as part-II of my Ph.D. thesis, I tried to address these challenges by improving methods to separate and detect most of the major lipid classes (including novel lipids in *E.huxleyi*) in a single run by using ESI-MS in combination with Normal phase HPLC techniques. Further, I also worked on designing some computational algorithms that can process the detected lipidome data for qualitative screening, calculation of unknown fatty acids and visualization.

Part I

Elucidation of carbon allocation mechanism into alkenones and other energy storage components under nitrogen limitation

Introduction

The bloom-producing haptophyte alga *E. huxleyi* is known to be one of the dominant producers of alkenones and a cosmopolitan species widely distributed in the ocean (Conte et al. 1994; Volkman et al. 1995). According to whole genome analysis of various strains of *E. huxleyi*, such wide distribution among various stressful environments is due to the presence of pan genome present in nuclear genes (Read et al. 2013). This is due to their evolution through secondary endo-symbiosis mechanism. *E. huxleyi* with such a complex genome exhibits a wide range of features. It is considered as the major producers of Alkenones among other Haptophytes in global Oceans. It can also exhibit certain physiological features like haplodiplod life cycle and formation of $CaCO_3$ exoskeletons (coccoliths). These features are based on several physiological and transcriptomic differences between haploid (motile non-cocolith forming cells) and diploid (non-motile coccolith forming cells) phases (Dassow et al. 2009; Houdan et al. 2005)

In addition, as a photoautotrophic organism, *E. huxleyi* also produces various kinds of photosynthetic metabolites. Mainly it produces a water soluble neutral-polysaccharide (NP) composed of β-1, 3/1, 6-glucan which is a storage product in many microalgae (Varum et al. 1986) and acid-polysaccharides (AP) which are specially associated with coccolith morphogenesis in coccolithophores belonging to haptophytes (Fichtinger-Schepman et al. 1981). Studies have also demonstrated that *E. huxleyi* also consists low molecular weight compounds (LMCs) like mannitol. Having such diverse group of storage components limits the understanding on the role of alkenones as prime storage components.

Recent isotope studies have demonstrated that *E. huxleyi* distributes a major portion of carbon into low molecular weight compounds (LMCs, 35% Ca.) like mannitol, membrane lipids (20% Ca.), neutral lipids such as alkenones (15% Ca.) and neutral and acid polysaccharides (NP and AP, respectively) (Epstein et al. 1998; Fernández et al. 1996a; Obata et al. 2013; Tsuji et al. 2015). Mainly, Fernández et al. (1994, 1996a), Tsuji et al. (2015) and Pan et al. (2017) observed active C-allocation into alkenones/lipids and alkenones functions as storage components in the cells. Fernández et al. (1994, 1996a) studied C-allocation in the photosynthetic carbon metabolism in *E. huxleyi* cells obtained in coastal mesocosm experiments. They found that the coccolithophore *E. huxleyi* has high capacity to uptake dissolved inorganic carbons from sea water based on nitrogen availability. This is due to the observed lower C-incorporation into proteins and high C-incorporation into lipids in comparison with other phytoplankton species. They also found that reallocation of C into proteins takes place in dark conditions and are related to cellular growth rate. Further they observed that the shifts in C incorporation into LMCs are related to species composition associated with halocline. Tsuji et al. (2015) found the significance of alkenones, but not β-glucan, as storage components. Pan et al. (2014, 2017) found that (i) the carbon isotopic composition changes (heterogeneous or homogenous) along with the lipid composition (membrane lipids or storage lipids) in the logarithmic or stationary growth phase, respectively and (ii) during the stationary growth phase, the lipid composition is of high poly-, mono-unsaturated FAs and alkenones and low saturated FAs, phytol and sterols than in the logarithmic growth phase.

A strong association between genes involved in C and N metabolisms was

observed in many algae including *E. huxleyi* (Flynn et al. 1993a, b; McKew et al. 2013; Nunes-Nesi et al. 2010; Singh et al. 2008; Talmy et al. 2014). Nitrate (NO_3^-) availability was identified as the main factor to give a crucial impact on the amounts of alkenones as well as other C-storing compounds in *E. huxleyi* (Kafes et al. 2010). Despite such observations, very less information is available on the relationship of metabolic regulation between alkenones and other storage components (membrane lipids, polysaccharides, proteins, etc.). Moreover, no report is available on C-allocation profiles in alkenone-producing haptophytes when cells were exposed to N-limiting conditions (–N) from N-sufficient conditions (+N) by changing nitrate availability in the culture.

In addition, strains with genotypic differences are already found in *E. huxleyi* CCMP 1516 and CCMP 2090 as coccolith and non-coccolith-bearing strains, respectively (Dassow et al. 2009; Houdan et al. 2005, Read et al. 2013). However, no studies are reported to interpret how photosynthetic carbon metabolism on C-allocation is changed especially in the accumulation into alkenones and C-distribution among various cell components in non-coccolith-bearing strain of *E. huxleyi*.

There are several unresolved issues in alkenone research like (i) biosynthesis mechanism, (ii) biological role and (iii) influence of non-thermal factors on alkenone unsaturation. These pose a major challenge for understanding their potential as biofuels and biomarkers. Current work aims at addressing some of these issues by looking at the C-regulation into alkenones and other metabolites in *E. huxleyi*. From this I expect to observe; (i) a clear difference in C-allocation into various metabolites when cells are exposed to sudden changes in N-nutrition, namely from +N to –N, and (ii) to know the mechanism by which *E. huxleyi* switches the direction of C-flow into various

macromolecular pools functioning as the significant carbon sinks (iii) effect of N-limitation on alkenone unsaturation.

For this, I examined the effects of N-limitation on change in C-allocation into alkenones and other lipids such as membrane lipids, proteins, carbohydrates etc. in the alkenone-producing haptophyte *E. huxleyi* CCMP 2090. Finally, this study revealed that the alkenones (but not NP and membrane lipids) function mainly as storage components in the cells, indicating the presence of quite unique physiological profile of alkenone-producing Haptophytes in comparison with other non-alkenone-producing microalgae. Further, I presented a model of C allocation under +N and –N into various cellular components of *E. huxleyi* CCMP 2090. In addition, I also found that –N does not affect alkenone unsaturation in this stain.

Materials and Methods

Organism and culture conditions

The *E. huxleyi* CCMP 2090 was obtained from the National Centre for the Culture of Marine Phytoplankton (CCMP, Maine, USA). *E. huxleyi* CCMP 2090 cells were grown in artificial seawater Marine Art SF-1 (Osaka-Yakken, Osaka, Japan), enriched with Erd-Schreiber's seawater containing 10 nM sodium selenite instead of soil extracts contained as an original component (Danbara and Shiraiwa, 1999). The standard composition of nitrate and phosphate in control medium were 1.41 mM $NaNO_3$ and ca. 30 μM K_2HPO_4. The temperature, light intensity and air flow rate were maintained at 20°C, 100 μmol quanta m^{-2} s^{-1} (continuously illuminated by fluorescent lamps) and 75-80 mL min^{-1}, respectively.

For the experiment as shown in Fig. 3, the cell suspension was divided into two batches, namely batch-1 and batch-2 for N-sufficient (+N, which is standard medium) and N-limitation (–N medium containing 0.05 mM $NaNO_3$) conditions, respectively. For the analysis of cell components, algal cells were harvested at two stages namely, logarithmic (stage-I) and stationary growth phases (stage-II) in each batch. Totally, six of 1.5-L-flat oblong glass bottles were used to grow the cells under the +N conditions until the mid-log growth phase for 4 days.

To obtain the stage-II cells, the stage-I cells were harvested once by centrifugation at 2,000 × *g* for 10 min and then the cell pellets were re-suspended into two kinds of fresh medium, namely either the +N or –N medium, in three 1.5-L-bottles each. The initial concentration of cells was adjusted to 0.2–0.3 (OD_{750}). Cell number

and size were monitored daily by using a cell counter (Sysmex CDA-1000, Kobe, Japan). Cells were harvested every 24 h after starting the stage-II culture from day 4 to 8 during 96-h cultivation. The harvested cell samples were analyzed for several components using various methods as described below and shown in Fig. 4.

Analytical methods for quantification of cell components

For the analysis of cell components such as chlorophylls, proteins, NP, AP, one mL of cell suspension was withdrawn from the culture and cell pellets were obtained by centrifugation at 2,300 × *g* for 5 min for each analysis. Chlorophylls were extracted from cell pellets with 90% acetone according to Jeffrey (1972) and quantified by a spectrophotometer (UV-1700, Shimadzu, Kyoto, Japan). For the protein assay, the cell pellet was resuspended into 200 μL of distilled water and then sonicated for destroying cells. After obtaining supernatant, total protein content was determined by the Bradford assay method (Bio-Rad, Hercules, CA) with a series of 0-40 μg mL^{-1} of bovine serum albumin as a standard. Polysaccharides were extracted with 300 μL of 5% (w:v) trichloroacetic acid from *E. huxleyi* cell pellet by sonication. After sonication, the extract was centrifuged for 5 min at 2,300 × *g* and then 200 μL of supernatant was used for the estimation of either NP or AP. NP was estimated using the phenol-H_2SO_4 assay (Hodge and Hofreiter 1962) calibrated with a series of 0–90 μg mL^{-1} glucose as a standard. AP was estimated using the carbazole-H_2SO_4 assay (Bitter and Muir 1962) and calibrated with a series of 0–90 μg mL^{-1} glucuronic acid (Chugai Pharmaceutical, Tokyo, Japan) as a standard. For the analysis of P_i in the medium, one mL of the resultant supernatant of cell suspension after centrifugation from above was used. Further, the analysis was done according to molybdenum blue method (Murphy and Riley 1962).

Extraction, separation and analysis of lipid fractions

Lipid extraction was done according to Sawada and Shiraiwa (2004) with minor modifications. Lipids were separated into two fractions, namely neutral lipid fraction (NLF) and polar lipid fraction (PLF). NLF and PLF are generally composed of alkenones, alkenes and TAGs, and all membrane lipids, respectively. Whole lipids were extracted from *E. huxleyi* cell pellets with 4 mL of methanol by sonication.

As internal standards (IST) for the quantification of alkenones and FAs, 20 μg of *n*-triacontane and heptadecanoic acid (50 μL each of 1 mg mL^{-1} solution) were spiked into 4-mL of crude extracts, respectively. Then, the extracts were fractionated with 4 mL of 1:1 methanol:dichloromethane and then 4 mL of dichloromethane. Thereafter, water (25 mL) and saturated NaCl solution (5 mL) were added and then mixed vigorously. After separation, the lower organic phase containing lipids was further dehydrated by passing it through a column containing anhydrous Na_2SO_4 and dried using an evaporator. The dried lipid fraction was dissolved in 2 mL of hexane and applied onto a silica gel column and eluted sequentially.

At first, NLF was collected using hexane, hexane:ethyl acetate (95:5, v:v), hexane:ethyl acetate (9:1, v:v). All these eluates were combined, evaporated and re-dissolved in hexane for gas chromatography with flame ionization detection (GC-FID) analysis. At second, PLF were eluted with methanol:ethyl acetate (1:1, v:v). The PLF was further dried and re-dissolved in 0.5 M HCl in methanol and incubated at 100°C for 90 min for methyl esterification. The FA methyl esters were extracted using hexane and used for GC-FID analysis. The GC-FID analysis of alkenones was done according to the method of Nakamura et al. (2014). The NLF and methyl esterified PLF

were detected by a GC-FID (GC-2014AFSC; Shimadzu) with a CP-SIL 5CB column (50 m, 0.32 mm id, 0.12 μm; Agilent, Santa Clara, CA). The carrier gas used was He with a flow rate of 2.08 mL min^{-1} in split-less mode. The column temperature was set to 60°C for 1.5 min, heated at a speed of 20°C min^{-1} to 130°C, and then taken at 4°C min^{-1} to 300°C, and holding at 300°C for 25 min.

Detected peaks of NLF and PLF were quantified by normalizing with respective IST peak areas corresponding to the internal standards. The amount of individual FAs was summed up to calculate total FA (TFA) content which is regarded PLF content. The relative accumulation of individual FAs was also determined by taking the ratios of mol% of each component.

For detecting TAG, NLF was used for analysis by gas-chromatograph, GC-2010, equipped with a mass spectrometer, QP-2010 (Shimadzu) by following above method in which analysis was performed at an increased temperature from 230°C to 340°C. The sensitivity of GC-MS was determined by analyzing a series of standard mixtures of various concentrations of TAG 17:0/17:0/17:0 (TAG C17) (0-100 ng μL^{-1}) and triacontane (100 ng μL^{-1}).

Total organic carbon (TOC) analysis

Fifty mL of culture was harvested during mid-log phase (day 4) to a stationary growth phase (day 8). Half of the cell suspension (25 mL) was centrifuged (1,720 × *g*, for 5 min). The supernatant was filtered through a syringe driven Millex HA filter (pore size 0.45 μm, Merck Millipore, Darmstadt, Germany) to obtain a culture medium free of cells (culture medium fraction). TOC values of both cell suspension and culture media

fraction (25 mL each) were measured with a TOC-L analyzer (Shimadzu). C quota of cells (TOC of cells) was calculated by [(TOC in cell suspension (mg L^{-1})) – (TOC in the medium (mg L^{-1}))]. Then, TOC in a cell (pg L^{-1}) was calculated by [(TOC of cells)/(cell number (cells L^{-1})].The relative value of total carbons of each metabolite such as proteins, lipids, alkenones, NP and AP (expressed as %TOC) was individually calculated as the ratio of TOC in each metabolite to total TOC in a cell. For the purpose, I did an approximate estimation of organic C for proteins, NP and AP. The calculation was done assuming that 53% of total weight of proteins is composed of carbon (Laws 1991). Similarly, %C of NP and AP were calculated by determining carbon content from β-1, 3- and -1,6-glucan (40%) and AP structure (39%) determined by Fichtinger-Schepman et al. (1981), respectively (Table. 1). Such estimates to calculate carbon equivalents for various metabolites were previously applied successfully to study patterns of C metabolism in *E. huxleyi* (Fernández et al. 1996b).

Additionally, total nitrogen (TN) of cells was also calculated similar to TOC of cells, by the TOC-L analyzer (equipped with TN measuring unit). The measurement was based on the catalytic oxidation method and detected using an infrared gas analyzer. Here, TN is the sum of NO_3-N, NO_2-N, NH_3-N and organically bonded nitrogen in the compounds. Using this, C:N ratio was calculated to observe the effect of N on C allocation.

Fourier Transformed Infrared (FTIR) Spectroscopy

Twenty-five mL of cell culture was harvested from all conditions and dried on silicon windows. Spectral acquisition and band assignment was done as mentioned in Giordano et al. (2001), Pelusi et al. (2016) and Palmucci et al. (2011). Briefly, an

Affinity-1S IR (infrared) spectrometer (Shimadzu), with an excitation source fitted with He-Ne laser operating at wavelength of 632.8 nm was used. The FTIR spectrum for each sample was acquired over the wave number ranging from 4000 cm^{-1} to 500 cm^{-1}. Resolution is maintained at 4 cm^{-1} with 32 averaged scans. FTIR spectral bands/peaks at 962.5, 1057, 1655, and 1737 cm^{-1} were selected for relative quantification of alkenones, carbohydrates, proteins (amide-I), and lipids (except alkenones), respectively as shown in Table. 2. LabSolutions IR software (Shimadzu) was used to achieve peak separation and deconvolution. The de-convoluted spectral band areas (i.e. content) of each macromolecular compounds were used for calculating a relative quantity of each compounds (% of the sum of organic matters determined), according to Dean et al. (2010). In this study, semi-quantification of each macromolecule was performed by FTIR analysis according to Palmucci et al. (2011).

$$\text{Concentration of macromolecule (i.e lipid, alkenone, carbohydrate or protein)} =$$

$$\frac{\%\text{of macromolecule organic matter}}{\%\text{of protein organic matter}} \times (\text{actual amount of protein})$$

To observe the changes in LMCs alone using FTIR, fractionation of cellular components was done according to Fernandez et al. (1994) and Li. et al (1980). Based on this freeze dried cell pellet was subject to $CHCl_3$:MeOH:DW (0.5:0.5:0.45) extraction. In this, LMCs were extracted into the top MeOH:Dw fraction (polar fraction, F1), middle insoluble pellet (F2) consisted polysaccharides, nucleic acids and proteins, and lipids were extracted into bottom CHCl3 (non-polar fraction, F3). 80µl of this F1 fraction was dried on silicon windows and analyzed for LMCs.

Results

Effect of –N on growth and cell property

The coccolithophore *E. huxleyi* CCMP 2090 cells grown at 20°C in +N medium for 4 days (stage-I) were transferred to either the +N or –N conditions (stage-II) (Fig. 5). The cell growth was suppressed under the –N conditions whereas high growth rate was maintained at least 3 days under the +N conditions. Such effect of the –N was observed in change of Pi concentration in the medium (the initial concentration: 30 μM) which was nearly depleted to 1.2 μM under the +N conditions of stage-I, and the depletion was little suppressed under the –N conditions (final [Pi]: 1.4 and 4.4 μM under +N and –N conditions in stage-II), respectively.

Effect of –N on cellular components such as chlorophylls, proteins, polysaccharides and lipids

Effect of the –N on changes in cellular components was analyzed during stage-II (Fig. 6). The increase in cell number was greatly changed depending on N-conditions and the cell diameter of –N cells gradually increased about 20% for 4 days during stage-II whereas no obvious change was observed in the +N cells (Fig. 5 and Fig. 6a, b). According to these results, the contents of cellular components were expressed on the basis of a cell in Fig. 6c-f. The chlorophyll and protein content (per cell) decreased rapidly irrespective of N-conditions, but the extent of decrease was more obvious under the –N conditions than +N conditions (Fig. 6c, d). For data expressed on the basis of culture, see Fig. 7a, b. Similarly, the contents of proteins, NP and AP gradually decreased irrespective of N-nutrition conditions during stage-II (Fig. 6e, f; for

data per culture, see Fig. 7c, d).

Total lipid content is expressed as the sum of neutral lipids consisted of mainly alkenones and alkenes and polar lipids mainly composed of FAs. Percentage of each lipid fraction from total lipid content was measured as shown in Table. 3. Among them, alkenones, alkenes and FAs occupied 57.8%, 4.6% and 37.5% of total lipids and 77.3%, 5.9% and 16.8% of total lipids in cells grown under the +N and –N conditions, respectively, and harvested on day 8.

Effect of –N on NLF composed of alkenones and alkenes

Neutral lipids detected in *E. huxleyi* CCMP 2090 were alkenes (C_{31}, C_{33}) and alkenones involving $C_{37:2\text{-}3}$MK (methyl alkenones), $C_{38:2\text{-}3}$EK (ethyl alkenones) and $C_{38:2\text{-}3}$MK, $C_{39:2\text{-}3}$EK; while alkenoates were not detected here. The contents of alkenones and alkenes were calculated as the sum of those individual molecular species (Fig. 8). The contents of both alkenone and alkenes showed similar trend of changes under the +N and –N conditions during stage-II. Namely, both contents greatly increased under the –N conditions for 4 days in stage-II whereas no obvious change was observed under the +N conditions (Fig. 8a, b). The ratio of alkenones/alkenes in contents was constant at 13±1.5, irrespective of the N-nutrition conditions (Fig. 8c). For data per culture, see Fig. 9a, b.

Among various species of alkenones, molecules with two carbon double bonds such as $C_{37:2}$MK and $C_{38:2}$EK+MK were dominant (Fig. 10a). Furthermore, the ratio of $C_{37:2}/(C_{37:2} + C_{37:3})$ which is known as the C_{37} alkenone unsaturation index (U^{K}_{37}) was constant under the +N and –N conditions during stage-II (Fig. 10b). This data was also

confirmed by that the ratios of $C_{37:3}/C_{37:2}$, which were nearly constant during culture in stage-II and same under the +N and –N conditions (Fig. 10c). In addition, the ratio of C_{37}/C_{38} alkenones was not affected by N-nutrition conditions although the ratio gradually decreased during culture in stage-II (Fig. 10d).

While alkenones were greatly accumulated under the –N conditions, no production of storage neutral lipids like TAGs was observed in this strain of *E. huxleyi* CCMP 2090 (Fig. 11a), indicating that storage lipids in this alga are alkenones. In another analysis, the sensitivity of GC-MS on TAG was confirmed to be at least below 40 ng μL^{-1} (Fig. 11b).

Effect of –N on PLF composed of FAs

Changes in FA contents were examined after the transfer of 4-d-grown +N cells to either the +N or –N conditions in *E. huxleyi* CCMP 2090 (Fig. 12). No significant difference was observed in total contents of FAs as PLFs between the +N and –N cells in stage-II and the contents were constant at 0.5 to 0.4 ± 0.035 pg $cell^{-1}$ (Fig. 12a). However, the compositions of individual FAs were quite different between +N and –N conditions (Fig. 12b-d). The C_{16}-C_{18} medium chain saturated, mono- and di-unsaturated FAs (MCFAs) such as 16:0 (palmitic acid), 18:0 (stearic acid), 18:1n-9 (oleic acid) and 18:2n-6 (α-linoleic acid) were accumulated under the –N conditions while C_{18}-C_{22} polyunsaturated FAs (PUFAs) such as 18:5n-3 (octadecatetraenoic acid) as the major C_{18} FAs, 20:4n-3 (eicosatetraenoic acid) and 20:5n-3 (eicosapentaenoic acid) accumulated more under the +N conditions. However, the contents of 22:5n-3 (docosapentanoic acid) and 22:6n-3 (docosahexanoic acid, DHA) did not show a clear trend under any condition (Fig. 12d).

Effect of –N on C-allocation into macromolecular cell components

The C-allocation into various macromolecular cell components such as proteins, alkenones, FAs, AP and NP were analyzed using TOC system (Fig. 13). The amount of total organic carbons fixed by cells (per culture), namely C allocated into all organic compounds, was higher under +N conditions than that under the –N conditions (Fig. 13a). But, TOC per cell, namely C stored in a cell, was decreased under the +N conditions whereas the parameter was maintained constant under the –N conditions (Fig. 13b). In addition, significant differences in TN were also identified. During stage-II, the values of TN reached to 13.1 mg L^{-1} and 7 mg L^{-1} in +N and –N respectively on day 8. This along with TOC, indicated a significant difference in C:N ratio of 11 and 7 in –N and +N cells, respectively.

The C-allocation among major cellular components was analyzed by the TOC system in cells grown under the +N and –N conditions (Fig. 13c,d). Those components are proteins (ca. 46% and 25%), alkenones (ca. 15% and 27%), alkenes (ca. 1% and 2%), FAs (ca. 9% and 7%), NP (ca. 3.4% and 1.3%) and AP (ca. 3.4% and 3%) and others including unidentified compounds individually (ca. 21% and 35%) (data in parenthesis represent % of TOC in cells grown under the +N and –N conditions, respectively, and harvested at day 8). The differences in accumulations between both conditions during stage II (day 4 to 8) was also shown in Fig. 13e. According to data, FAs, proteins and NP are positively produced under the +N conditions and vice versa under the –N conditions. Whereas alkenones, alkenes and AP are positively produced under the –N conditions and vice versa (Fig. 13e). This figure shows the production of alkenones and proteins are conversely regulated by change in N-nutrition conditions. In order to

examine the significance of the storage components in C-storage, C-allocation patterns during day 5, 6 and 8 were also represented clearly in Fig. 14.

Changes in macromolecular carbon pools from FTIR spectra

The FTIR spectra of +N and –N cultures showed a clear difference in the accumulation of carbon content into major macromolecules (Fig. 15). Increasing alkenone content under –N can be clearly seen from the increasing alkenone specific peak (at 962.5 cm^{-1}) during day 6-8. Another peak at 1705.5 cm^{-1}, also increased under –N specifically. Also, a large difference in carbohydrate specific peak (1057 cm^{-1}) was observed between +N and –N conditions. From the semi quantification data, accumulation patterns of alkenones and lipids and carbohydrates were observed as shown in Fig. 16. This data also show similar trend in accumulation of alkenones and depletion of proteins under –N conditions in comparison to +N condition. However, the carbohydrate pool under –N condition (35 to 52 ± 3 μg mL^{-1}) was higher than +N (22 to 34 ± 3 μg mL^{-1}) condition. The signal at 1057 cm^{-1} is assumed to be same for polysaccharides and other carbohydrates. It occupies nearly 35 to 45% of the total cellular contents besides proteins. It appears to be increasing over time in both conditions. Further analysis of FTIR spectra for LMC fraction (Fig. 17) showed an increase in specific peaks at 1057 cm^{-1} under both conditions. However, the LMC specific peaks were much big in spectra of –N day 7-8, indicating stimulated production of LMC like sugars under –N condition.

Discussion

Changes in cell growth and carbon storage caused by –N

The analyses of C-allocation into major cell metabolites such as proteins, carbohydrates, alkenones, FAs, AP and NP were performed in *E. huxleyi* CCMP 2090 cells grown in a continuous light under the +N and –N conditions. It is very clear that the –N triggered suppression of algal growth and Pi-utilization together with the enlargement of cell volume (Figs. 5, 6). The suppression of Pi-utilization by the –N can be explained by suppression of photosynthesis and growth. In addition, the rapid decrease in Pi concentration in the medium also suppressed cell growth, but the effect of Pi-limitation seems not so serious since cell growth was not so greatly suppressed under the +N conditions (Fig. 5).

Increase in cell volume was also observed in other microalgae and explained by the suppression of cell division, namely the suppression of increase in cell number per culture, as reported previously (Flynn et al. 1993a, b; Msanne et al. 2012; Parrish et al. 1998).

The total organic carbon was determined in *E. huxleyi* CCMP 2090 under the –N (6.7 ± 0.9 pg $cell^{-1}$) and +N conditions (ca. 3.85± 0.3 pg $cell^{-1}$) at the end of stage-II in this study (Fig. 13). The value under the –N conditions is similar to the value (7.8 pg $cell^{-1}$) reported by Riegman et al. (2000) in *E. huxleyi* (strain L) grown in a continuous N-limited (0.025 mM $NaNO_3$) culture. These data suggest that N-limitation did not so greatly influence cell metabolism and total carbon accumulation per culture even after 4-day under the –N condition in this study (Figs. 6, 13a). Such results agree with

previous reports which showed *E. huxleyi* has an ability to continue photosynthesis at a low rate even under the –N conditions (Loebl et al. 2010; Raven et al. 2012; Rokitta et al. 2014).

Effect of –N on lipids in composition and unsaturation degree

When *E. huxleyi* CCMP 2090 cells were exposed to the –N conditions, photosynthetically fixed C was greatly allocated into alkenones (ca. 27% of TOC) as storage neutral lipids. The C contents of alkenones in *E. huxleyi* CCMP 2090 was 1.5-fold higher than that of previous reports in several coccolith forming strains under N- or P-limitation (Bell and Pond, 1996; Epstein et al. 1998; Epstein et al., 2001; Prahl et al. 2003; Pan et al. 2014). This indicates that the non-coccolith producing strains of *E. huxleyi* are adequate for their utilization as good alkenone producers in future study for biofuel production.

The relative amount of alkenones (neutral lipids) accumulated was 77% of total lipids in *E. huxleyi* CCMP 2090 cells when grown under the –N condition. This is similar to maximum accumulation of neutral lipids during stationary phase (73% of total lipids) in *E. huxleyi* CCMP 371 as reported by Pan et al. (2017). On the other hand, non alkenone producers like *Chlamydomonas reinhardtii* and *Nannochloropsis gaditana* were reported to accumulate high amount of neutral lipids like TAGs (38%-79% of total lipids) under –N conditions (Simionato et al. 2013; Fan et al. 2011; Msanne et al. 2012; Li et al. 2011). This is due to that most algae, when placed under –N conditions shifts their metabolic direction from carbohydrates to MCFAs and produce storage lipids like TAGs (Choi et al. 2011; Elton et al. 2016; Jia et al. 2015; Michel et al. 2010; Negi et al. 2016).

The haptophyte *E. huxleyi* was found to produce only very little amounts of TAGs; despite accumulating MCFAs under –N or during stationary phase (Reitan et al. 1994; Pan et al. 2017). Similarly, my results also showed increased MCFAs and absence of TAGs under the N-limitation. In addition, Garnier et al. (2016) showed that 28% of the total storage lipids in wild-type of *Tisochrysis lutea* (WTc1) consists alkenones, but their amount was decreased to <2% in TAG-accumulating mutant (2Xc1). These results suggest that the biosynthesis of alkenones and TAGs might be co-regulated in a competitive manner in the haptophyte alga. When *E. huxleyi* cells were kept under –N conditions, the tricarboxylate carriers were up-regulated to increase delivering of precursors from lipid producing organelles to cytoplasm (Rokitta et al. 2014).

On the contrary, the high accumulation of TAGs under –N conditions was observed in some strains of *E. huxleyi* and other alkenone producing haptophyte *Isochrysis galbana (*Maltisky et al. 2016; Parrish et al. 1998*).* According to these evidences, I cannot completely rule out the possibility that alkenone-producing haptophytes still maintain the ability or potential to produce TAGs and the metabolic potential is expressed under some conditions. Further study is required to elucidate the complete mechanism of alkenones and TAGs metabolism in alkenone-producing algae or phylogenetically related haptophyte species.

Recently, some attempts are being made to utilize alkenones as alternate precursors of hydrocarbons for jet fuels such as C_{11}-C_{18} alkanes (O'Neil et al. 2016). For the purpose, the development of metabolic engineering approach to manipulate the chain length of alkenones and alkenes as well as removing *keto*-group from the molecules will be strongly required. The information on the complete metabolic process

and the regulatory mechanism of alkenones and alkenes production will become more important to realize such metabolic engineering procedures.

The $C_{37:2}$MK ($C_{37:2}$), molecule was dominantly synthesized among alkenones (50-60% of total alkenones) while $C_{37:3}$MK ($C_{37:3}$) comprised only 8-7% in *E. huxleyi* CCMP 2090 grown at 20°C (Fig. 10a). The result is different from Pan et al. (2017) reporting that $C_{37:3}$ was dominant species occupying 28% of neutral lipids, namely alkenones, in *E. huxleyi* CCMP 371 grown at 17°C, suggesting species-dependent variation in alkenone composition.

In contrast to the effect of –N to increase the production/quantity of alkenones, the composition of alkenones was not affected by N-availability during culture under both the +N and –N conditions (Fig. 10a). Therefore, the $U^{K'}_{37}$ index was maintained constant (Fig. 10b). The result was also confirmed by showing no significant difference in the ratio of $C_{37:3}/C_{37:2}$ during culture under the +N and –N conditions (Fig. 10c). These results agree with results of Popp et al. (1998), showing no significant change in $U^{K'}_{37}$ under N-limited conditions. Whilst, Pan and Sun (2011) reported that $U^{K'}_{37}$ decreased with the increase in $C_{37:3}$ and decrease in $C_{37:2}$ and related to growth phase of *E. huxleyi* CCMP 371 grown at 17°C. The existence of inverse relationship between $U^{K'}_{37}$ and some factors such as NO_x status and cell division was reported by Epstein et al. (1998). Yamamoto et al. (2000) observed that the unsaturation degree of alkenones and FAs (mainly C_{18}) changed proportionally depending on growth stage during culture even in isothermal cultures. The present study showed significant changes in FA composition between +N and –N conditions without change in $U^{K'}_{37}$ throughout stage-II in *E. huxleyi* CCMP 2090 (Fig. 12, 10b) and the data agree with that $U^{K'}_{37}$ value is well

known to respond stably to temperature only but not to N-limitation (Popp et al. 1998). At this moment, there is no information on molecular mechanism how and why the alkenone unsaturation degree is responding to temperature change. Therefore, effect of other factors except temperature on unsaturation degree of alkenones needs to be carefully analyzed for which steps or genes are regulated and respond to temperature change in future study.

The present results show that the chain length of alkenones was maintained almost same but with slight increase in C_{38} alkenones, but without any effect of N-limitation in *E. huxleyi* CCMP 2090 (Fig. 10b, d). However, recent reports using *E. huxleyi* CCMP 371 showed the increase in the production of C_{38} and C_{39} alkenones (Pan et al. 2017) and the decrease in chain length ratio (C_{37}/C_{38}) depending on changes in respiration and vary in various growth phases at the stationary growth phase (Pan and Sun 2011). At this moment, there is no clear explanation of such difference in trend of change in alkenone composition.

Concerning effects of –N on polar lipids, no significant change was observed in total amount of FAs per cell when *E. huxleyi* cells were transferred from the +N to the –N conditions (Fig. 12a). However, the transfer of cells from +N to the –N triggered a remarkable change in the composition of FAs from C_{18}-C_{22} PUFAs to C_{16}-C_{18} saturated, mono- and di-unsaturated FAs (Fig. 12b-d). Similar results were previously reported in green microalgae and alkenone-producing haptophytes (Dunstan et al. 1993; Msanne et al. 2012 ; Reitan et al. 1994). Despite such remodeling of FAs composition, total amount was not changed. A minor difference in %C of FAs might be due to slight decrease in contents of long chain (>C_{18}) PUFAs under the –N. Such results indicate the presence of

strong regulatory mechanism of C-allocation in membrane lipids. The proportions of C_{18} PUFAs, mainly 18:5n-3 were high under the +N conditions and decreased under the –N conditions, indicating their significance in the +N conditions (Fig. 12b-d). The 18:5n-3 was identified as the major constituent occupying 40-60% of glucolipids in *E. huxleyi* (Bell and Pond 1996) and plays important role in photosynthesis (Pond and Harris 1996; Kotajima et al 2015). The decrease in 18:5n-3 might be decreased accompanied with decrease in chlorophylls (Figs. 6c, 12c). From this point of view, the decrease in 18:5n-3 may trigger thylakoid membrane remodeling, as suggested by occurrence of remodeling of chloroplast membranes under the –N conditions in the literature (Goodson et al. 2011). The proportions of moderately accumulated 18:3 and 18:4 and less accumulated 20:4n-3 and 20:5n-3 did not change significantly during culture, irrespective of N-conditions. This may be due to the fact that these PUFAs are involved in phospholipid biosynthesis (Pond and Harris 1996).

The level of DHA was constant at 17% of TFA content during culture, irrespective of N-nutrition conditions (Fig. 12) and the result is similar to Pan et al. (2014) reporting 18-19% of TFA content. It was reported that DHA is the major component (>60%) of phosphatidylcholine among other components such as 14:0, 16:0 and 18:1n-9 in *E. huxleyi* (Bell and Pond 1996). Recently, abundance of phosphatidylcholine in total lipids was reported to be nearly 30% in *E. huxleyi* CCMP 2090 (Shemi et al. 2016). In addition, current work showed that amounts of 16:0, 18:0, 18:1n-9 and 18:2n-6 FAs were synthesized more under the –N conditions than +N conditions (Fig. 12b-d). These results indicate that N-limitation promotes remodeling of membrane lipids to transfer FAs from chloroplast membranes such as thylakoids to

non-chloroplastic membranes.

Generally, 18:1n-9 is assumed to be involved in TAGs as energy storage or in membranes for regulating membrane fluidity (Brown et al. 1996; Iwai et al. 2014; Allen et al. 2015; Pond and Harris 1996). However, the haptophyte *E. huxleyi* CCMP 2090 produces alkenones as storage lipids, but no TAGs and therefore 18:1n-9 may play significant role as membrane lipids in the alga.

A model of metabolic regulation of C-allocation among various cell components by N-availability

When *E. huxleyi* CCMP 2090 cells grown under the +N conditions were transferred to the –N conditions, C-allocation into proteins was greatly diminished from 51% to 25% (Fig. 13b) during day 4 to 8. This data shows that, C fixed is allocated into non-N compounds for storing C under N-limited conditions where cellular metabolisms are depressed, as described in the literatures (Ball et al. 1990; Benavente-Valdésa et al. 2016; Klein 1987). An increase in C:N ratio also supports this phenomena (Reitan et al.1994). The data also agree with the evidence showing that cells reallocate N for necessary proteins by degrading unessential proteins under –N conditions in diatoms (Alipanah et al. 2015; Bai et al. 2016) as well as *E. huxleyi* 1516 (McKew et al. 2013).

The haptophyte alga *E. huxleyi* produces two kind of macromolecular polysaccharides, namely NP and AP, and those metabolisms are regulated by Pi availability (Kayano and Shiraiwa 2009). In previous reports, Obata et al. (2013) and Tsuji et al. (2015) identified that LMCs were mainly accumulated as metabolic carbon pools occupying nearly 35% of whole cellular TOC and mannitol is found to be one of

major compound in LMCs. However, the content of NP like β-glucan is low and does not function as energy storage and alkenones is the major energy storage compound in *E. huxleyi* (Tsuji et al. 2015). The present study confirmed the very small content of NP, but also clearly found that NP production was stimulated under the +N conditions and decreased by 2.5% under the –N conditions (Fig. 13c-e), whilst AP content was also small at ~3% TOC but clearly increased under the –N conditions (Fig. 13e). However, the exact reason for the increase in AP production under the –N conditions is yet unclear (Fig. 6f, 13e). One possible explanation is due to the stimulation of AP production used as the matrix for coccolith production under –N conditions even in *E. huxleyi* CCMP 2090, which genetically lost coccolith production ability. According to my data, strong stimulation of coccolith production can be observed under the –N in coccolith-producing *E. huxleyi* NIES-837 (data not shown). Also, in this work AP was surely confirmed to be produced (Fig. 13e) and therefore it is possible to consider that AP is still produced even in this strain. On the other hand FTIR data showed an increase in carbohydrate content from semi quantification of 1057 cm^{-1} peak despite lower contents of total polysaccharides in *E.huxleyi*.

Generally, microalgae can shift their carbon storage compounds from NP like β-glucan to neutral lipids like TAG under stress conditions (Elton et al. 2016; Jia et al. 2015). Also, in another haptophyte strain (*Tisochrysis sp.* CCAP 927⁄14) an increase in carbohydrates along with neutral lipids during –N was observed (Lacour et al. 2012). Although LMCs were not directly quantified, I observed an increase in other unidentified components under the –N conditions during culture (Fig. 13d). Further, similar trend was also observed under –N condition FTIR spectra of MeOH:Dw fraction

that consist LMCs (Fig. 17). Therefore, I assume that LMCs including mannitol may function as carbohydrate C pool under the –N condition also. In order to estimate the LMC fraction from FTIR, semi quantification of polysaccharides and other unknown components is needed. However, at present such methods are not available and I am working on this. Therefore LMC accumulation could only be speculated visually.

Using quantitative analytical methods of colorimetric spectrometry and GC-MS in combination with TOC system, changes in the contents of major cell components (pg $cell^{-1}$) were examined in *E. huxleyi* CCMP 2029 cells grown under the +N and –N conditions (Figs. 6-12). Individual amounts of all components proteins, NP, AP, FAs, alkenones and alkenes (expressed as percent of total carbons) were calculated for this purpose. Out of which, %C of proteins, AP and NP were considered as approximate measurements due to no availability of accurate compositional information. However, these results provide a brief understanding on patterns of C flow into various metabolites in *E. huxleyi* CCMP 2029 (Fig. 13).

Current results indicate that proteins are most major components (46% of total carbons in the cells), and then neutral lipids composed of alkenones and alkenes (16%), polar lipids composed of FAs (9%) and polysaccharides composed of NP and AP (6.8%) in +N cells (Fig. 13c). This indicates that carbons fixed were mainly used for protein synthesis in *E. huxleyi* CCMP 2090 cells grown under the +N conditions, as reported in other studies in *E. huxleyi* (Kaffes et al. 2010; Rokitta et al. 2014) and in green algae such as *Chalydomonas* and *Selenastrum* (Fernández and Galvan 2007; Johnson and Alric 2013; Turpin 1991; Weger and Turpin 1989).

In summary, a model in Fig. 18 represents the direction of C-allocation into

various cell components under the +N and –N conditions in the alkenone-producing haptophyte *E. huxleyi* CCMP 2090 which genetically lost coccolith production ability. The *E. huxleyi* cells exhibit a dynamic carbon directing and redirection mechanism to store the energy by changing N-availability. Alkenones are the only neutral lipids which are accumulated and immediately used as C sinks to store energy in this haptophyte alga. It might be an energy efficient mechanism to readily use stored C from alkenones than polysaccharides such as glucan. In addition, significant changes in acyl-group composition of polar lipid fraction show that this alga can effectively remodel its membrane lipids in response to variations in N availability. The detailed pathways including various intermediates involved in the carbon allocation mechanisms is yet to be discerned. All these results could be beneficial for understanding biological role of alkenones in haptophytes and further application of alkenones to geo-physical and geoscientific studies involving the alkenone paleothermometers and algal biofuel production. In addition, from FTIR results, we propose that a peak at 1075.5 cm^{-1} an alkenone related peak (Pelusi et al. 2016) can be used as a marker for N limiting condition in alkenone producing haptophytes.

Part II

Part II Establishment of lipidomics methods and lipidomics of *Emiliania huxleyi* CCMP 2090

Introduction

Lipids are a complex mixture of a variety of fatty acids and specific chemical structures known as head groups. Different head groups have unique physio-chemical characteristics by which the lipids are divided into various classes. Due to such varied properties lipids play important role in biological systems. For example, cellular membrane fluidity can be affected the degree of unsaturation of the acyl chains. Identification and quantification of all lipids from biological samples to understand lipid metabolism and lipid mediated cell signaling is lipidomics (Han and Gross 2003). This helps us to study biosynthesis of specific metabolites, storage of energy, cellular membrane dynamics, study of diseases such as obesity, diabetes, atherosclerosis and cancer (Wenk 2005; Watson 2006). As a consequence lipidomics is an emerging discipline in many fields of scientific research spanning from industry to medicine. Despite having such significant role, due to their large number and diversity, the detection and molecular characterization of lipids from biological samples has been difficult.

In general lipid mass spectrometry (MS), most of the lipid head groups are detected by positive ionization method and fatty acids (FA) are detected by negative ionization method (Brugger et al. 1997). Previously, several methods are established to detect major membrane lipid classes based on their head group fragmentation. Phospholipid, glycolipid classes are detected in both negative and positive ionization modes (Brugger et al. 1997; Ejsing et al. 2006; Milne et al. 2006; Taguchi et al. 2005; Welti et al. 2002; Popendorf et al. 2013). From these reports (as discussed in general introduction), it is indicative that for every type of lipid class detection, a shift between

positive to negative ionization might be required every time. This is very tedious job and might result in loss of necessary information. Also, detection based on multiple precursor ion scan (MPIS) for neutral lipids like DAG, TAG and polar lipid fatty acids are previously suggested (Stahlman et al. 2009; Ejsing et al. (2006). However, MPIS might result in acquiring less MS information as multiple experiments reduce search space. Further, several new lipid classes like Betaine lipids (BLs), phosphatidyldimethylpropanethiol (PDPT), glycosphingo lipids (GSLs) etc. are recently identified in marine haptophytes (Armada et al. 2013; Roche and Leblond 2010; Van mooy et al. 2009; Popendorf et al. 2013). Currently, very few methods are available to identify these new lipids on ESI/MS system. Therefore, there is a need to optimize the available methods to detect most of the lipid classes either in positive or negative modes. Also, optimization is a necessary process as ionization parameters vary with each MS-equipment.

Prior to lipid detection by MS, lipid separation is also necessary. With recent advances in lipid separation techniques like chromatography, several reports are available on high performance liquid chromatorgraphy (HPLC) separation techniques for lipidomics. Hummel et al. 2001 showed separation of plant membrane lipids in a 22 min time scale. Polar glycerolipid separation was developed by Sturt et al. (2004). Ikeda et al. (2011) showed separation of lipids species within lipid class in 90 min duration. Popendorf et al. (2013) improved the method of Sturt et al. 2004, to separate membrane lipids including marine Haptophyte specific lipids in a single 45 min run. All these methods focused on membrane lipid separation only. Recently, Knittelfelder et al. (2014) showed that by using ultra performance liquid chromatography (UPLC)

separation of neutral lipids and glycerol phospholipids in a single 50 min run. However, in most of these methods the separation of many lipid classes within phospholipids, neutral and glycol lipids together is not achieved completely. Also, more research is yet to be done on detection and separation of unique lipids like BLs (diacylglyceryl-3-O-carboxyhydroxymethylcholine, DGCC, diacylglyceryl trimethylhomoserine, DGTS and betaine like lipids, BLL), PDPT and GSLs (host specific glycosphingo lipids, hGSL and salicylic glycol sphingo lipid, sGSL) using LC-ESI/MS techniques.

Besides lipid detection, qualitative analysis by computations approaches is another challenge. Qualitative analysis requires comparison of hundreds of lipid classes between positive and negative methods. This helps to confirm the presence of each lipid and its fatty acyl groups (Lipid class (FA1/FA2)). However, this is not a simple task due to large data. There are no available tools to do this kind of data processing. Most researchers' use Excel based manual analysis, which is time taking and error prone. Also, as several lipid species are detected, representing those using simple bar graphs is also difficult. For a quick analysis of lipid classes and their relation with fatty acyl group's, network based visualization is a good way. Network visualization of biological data can be done by using well known software's like Cytoscape and BioLayout (Shannon et al. 2003, Theocharidis et al. 2009). These are most commonly used in genomic and proteomic analysis. But, they can also be used for lipidomics (Yetkuri et al. 2007). However, preparing the data for this kind of tools is difficult for a naïve biologist. Therefore, there is a need for development of computational methods for lipidomic data processing.

Thus, the main objectives of this work are (i) to setup an efficient lipidomics platform for the analysis of cellular lipids and (ii) to implement lipidomic method for E. huxleyi lipidome analysis. Currently, there is no single method that can detect all lipid classes (including haptophyte specific lipids) in one ionization mode. In addition, lipid separation methods using LC techniques are also highly diverse. Therefore, it is difficult to use a single approach to separate all lipid classes and study the lipidome in a single experiment. In this work I address these two issues by using the knowledge of lipid fragmentation and separation methods from previous reports and optimizing them to fit into a single experiment. I also developed computational methods for qualitative analysis of lipidome data. All these could be helpful in studying the lipid metabolism of biological samples easily in less time.

Materials and Methods

Internal standards (ISTs)

13 ISTs related to major lipid class were purchased from Avanti Polar Lipids, Inc. (Alabaster, AL), Matreya LLC (Pleasant Gap, PA) and Sigma-Aldrich (St. Louis, MO). Stock solutions of all these ISTs were prepared by dissolving in respective solvents separately as shown in detail in Table. 4. From the stock solutions, samples of ISTs were made ranging from 100 ng to 1 ng concentration by dissolving into 300:665:35 chloroform ($CHCl_3$):methano(MeOH):[Ammonium acetate(NH_4Ac)-300 μM + Sodium acetate(NaAc)-50 μM] according to Welti *et al.* (2002). NH_4 and Na salts were used to induce adduct formation and therefore increase ionization efficiency. Thirteen samples with 100 ng concentration of each internal standard and a sample-mixture of all 13 internal standards (each at 100 ng) were used for optimization of lipid detection and LC separation parameters.

Culture, sample collection and lipid extraction

E. huxleyi CCMP 2090 was obtained from the National Centre for the Culture of Marine Phytoplankton (CCMP, Maine, USA). *E. huxleyi* CCMP 2090 cells were grown in artificial seawater Marine Art SF-1 (Osaka-Yakken, Osaka, Japan), enriched with Erd-Schreiber's seawater containing 10 nM sodium selenite instead of soil extracts contained as an original component (Danbara and Shiraiwa, 1999). The temperature, light intensity and air flow rate were maintained at 20°C, 100 μmol quanta m^{-2} s^{-1} (continuously illuminated by fluorescent lamps) and aerated with ambient air at 75-80 mL min^{-1}/XX L culture, respectively. For the experiment a triplicate of *E. huxleyi* cells

were grown in flat oblong 1.5 L culture bottles and grown till stationary phase.

For lipidome analysis, sampling was done during mid logarithmic phase and stationary phases. Fifty mL of culture was used to collect cell pellet. The pellet was freeze dried in 2 mL centrifuge tubes and store at -80˚C until lipid extraction. Lipid extraction was performed according to Popendorf et al. (2013). According to this, total lipids were extracted from each filter using a modified Bligh and Dyer method (Blight and Dyer. 1959) by adding 0.5:0.5:0.45 (v:v) MeOH:dichloromethane (CH_2Cl_2):phosphate buffered saline, PBS (137 mmol sodium chloride, 2.7 mmol potassium chloride, 11.9 mmol phosphate, pH = 7.4). Initially 0.5 mL of MeOH and a 20 µL (100 ng μL^{-1}) of the internal recovery standard (DNPPE) were added to the cell pellet and subjected to sonication to disrupt the cells. Later, 0.5 mL and 0.45 mL of CH_2Cl_2 and PBS were added, respectively. The solvent mixture was vortexed thoroughly for solvent fractionation. The samples were then centrifuged for 5 min at 2,300 × *g*. The lower organic phase was transferred into pre combusted 2 mL glass vials by filtering them through a hand driven 0.20 µM organic filters (Millex-FG). The lipid fraction was further dried under N_2 gas, resuspended into 200 µL of 9:1 CH_2Cl_2:MeOH. This final lipid sample was transferred to a 350 µL inserts in a 2-mL screw cap borosilicate glass vials (Agilent) and used for analysis. The transfer of organic solvents at all steps was done using combusted glass pipettes. Highly concentrated lipid crude extracts were also prepared similarly and used for optimizing MS-methods to detect *E. huxleyi* specific lipids.

Lipid detection using QTRAP ESI-MS

According to the principle of ESI/MS, a strong electric field in the capillary

induces formation of positive or negatively ionized lipids. Later by using tandem mass spectrometry, detection of both molecular and fragment ions was done (Brugger et al. 1997; Han and Gross 2003, 2005). Here the first mass analyzer (quadrapole, Q1) is used for selection of molecular ion of interest, collision cell (Q2) is used for fragmentation and second mass analyzer (Q3) is used for selection of fragments of interest (Fig. 19). Based on the characteristic mass of detected product ion, it is considered as precursor ion scan (PIS) or neutral loss scan (NL). Q3 scan is referred as PIS when the mass of head group/FA of interest is equal to the mass detect at Q3 (product ion mass). On the contrary, if the difference between mass of molecular ion (at Q1) and product ion is equal to mass of head group/FA of interest, it is considered as NL (Fig. 19). An MPIS scan is also employed either in positive or negative ion mode. This is needed when several PIS scans for detecting a certain type lipid class is needed (TAG and DAG in positive mode; Fatty acids in negative mode).

Lipid ISTs samples at 100 ng concentration were introduced by syringe (direct injection) into the ESI source on a triple quadrupole MS/MS (QTRAP 5500, ABSciex, Framingham, MA, USA). But, by combining data from both methods provides details on lipid type and its fatty acyl composition as shown in Fig. 20. This approach will provide highly reliable qualitative lipidome data.

The lipid fragmentation was initially done using an automated optimization method by operating in both positive and negative ionization method. This resulted in certain electric potentials that are needed for lipid class selection (at Q1) and fragmentation (at Q2, Q3). These voltages for de-clustering potential (DP), collision energy (CE) and collision exit potential (CXP) were recorded, respectively. These

voltages were further fine-tuned by manual optimization. Here each of these parameters was manually changed until a large signal for lipid specific fragments was obtained. Except DP, CE and CXP, other parameters of the mass spectrometer were not changed. This was because of their minimal influence on lipid fragmentation. Thus, the entrance potentials at 10 V and -10 V, electron spray capillary voltage at 5500 V and -4500 V, source temperature (TEM) at 0˚C and ion source gases GS1 at 15 and 14, GS2 at 0, curtain gas (CUR) at 20 (arbitrary units) were fixed in both positive and negative ionization methods respectively.

In addition, special lipid classes like BLs (DGCC, DGTS and BLL), PDPT and GSLs (hGSL and sGSL) were detected in *E. huxleyi* according to Maltisky et al. (2016) and Fulton et al. (2014). However, except for DGTS, internal standards for these lipid classes were commercially not available. Therefore, to setup detection methods for these lipids, fragmentation-optimization was done using a highly concentrated crude lipid extract of *E. huxleyi* CCMP 2090. During this process, based on the lipidome data from Malitsky et al. (2016), mass of the most abundant lipid species of above lipid classes was taken as reference.

All the optimized parameters for fragmentation (DP, CE and CXP) and detection (NL and PIS) were used to make a method for lipidome analysis. Triplicates of *E. huxleyi* CCMP 2090 lipid samples were analyzed using this method. Data collection was done using Analyst 1.5 software, and data analysis was done using updated LipidView 1.1 (Zhou *et al.*, 2011) software (ABSciex, Framingham, MA, USA).

Lipid class separation using normal phase HPLC system

Here, I attempted to separate lipids based on their head group polarity. For this separation of lipids prior to ESI-MS analysis, I used a normal phase HPLC column (150 x 2.1 mm 5 µm BETASIL Diol). Chromatographic separation was performed on an Agilent 1200 HPLC system (Agilent Technologies) consisting of a binary pump, auto sampler (with sample tray cooled to 4°C). The HPLC eluents and gradient flow was adopted from Popendorf et al. 2013 and further optimized in several attempts to achieve lipid class separation. Eluants were as follows (by volume): eluent A= 600:400:1.0:0.4 *n*-hexane:isopropanol:formic acid: 25 % aqueous NH_4; eluent B= 880:120:1.0:0.4 isopropanol:water:formic acid: 25 % aqueous NH_4. For the separation of lipids, the gradient was performed as follows: 100 % A to 31.5 % B in 8 min, then increased to 51.5% B in 7 min, followed by increase in B up to 71.5% within 12 min, hold for 5 min, then recover to 100% A in 3 min and hold for 15 min. The flow rate was started from 0 to 0.1 mL min^{-1} within 8 min and hold at 0.1 mL min^{-1} till the end. The gradient flow and solvent parameters were provided in Table. 7 and Fig. 21. In case of *E. huxleyi* lipidome, PDPT lipids were observed to elute beyond 50 min, so the last step of gradient elution was extended to 60 min.

Computational methods for qualitative analysis and quantification

PERL programming based algorithms were developed for qualitative data analysis. The results from the PERL scripts confirm the presence of each lipid species and also calculate unknown fatty acyl groups. This flow chart for data processing is shown in Fig. 22. The PERL algorithm (Lipid-compare.pl) identifies a common lipid species in both ionization methods based on the head group name. The head group name (ex. PC 38:3) generally consist name of lipid class [PC], total carbons [38] and total

double bonds [3]. The lipid class in negative mode consists a fatty acyl group also (ex. PC 38:3 (-FA 24:3)). Once common lipid species are detected, other unknown FA of the lipid are calculated. For this, the fatty acyl group detected in negative mode is considered as FA1 (ex: FA 24:3) and the algorithm calculates other fatty acid of the lipid as FA2 (ex. FA 14:0). This is a simple subtraction where carbons and double bonds of FA1 are subtracted from total carbons and double bonds of head group to obtain FA2. Positive mode lipid species whose negative counterpart is not detected are not considered in this analysis. Similarly, the negative mode data without positive counter parts was also not considered in this analysis. The list of lipids with both FA1 and FA2 is considered as the reliable qualitative data, lipid-list-1. The intensities (from positive mode) of these lipids in lipid-list-1 were used for quantitative analysis. For quantification, the lipid MS-intensities were normalized by DNPPE and concentration was calculated. This list consist lipid class name, total number of carbons, double bonds, fatty acyl groups and their quantities (mol%). Once this is done, final qualitative check is done by selecting the lipids that are detected in more than half of the replicates in either of the methods. The final lipid-list-2 is generated, this list have most reliable and highly confirmed lipid species in lipidome.

For visualization of lipidome data, Biolayout 3D V 3.0 software was used (Theocharidis et al. 2009). This needs a specific input file format with interacting partners (ex. A interact with B is given as A-B) and parameters to represent them (like shape, size etc.). The interacting partners were considered as nodes (shown as a sphere, square etc.). The interaction is nothing but a relationship between two components. In networks this relation is represented as an edge (a line connecting two components).

The interaction list of all components is known as edge list. For all detected lipid species, we considered the lipid class, their FAs and carbon number as components (nodes) which are interacting with each other. Using this information I generated an edge list of E. huxleyi CCMP 2090 lipidome. The size of the node was based on the carbon number and total number of interactions with other nodes. As several lipid species are detected, making such input file for BioLayout is difficult. For this I developed PERL scripts (Pre-BioLayout.pl, BioLayout-Lipid-FA.pl and BioLayout-CC.pl). The output of these scripts is directly used as input for BioLayout tool which immediately shows a network of lipidome. Working protocols of all PERL scripts were included in Appendix for PERL scripts. Original PERL programs were attach at the end of this thesis.

Results

Lipid class ionization and fragmentation

Lipid class specific fragments were successfully obtained and parameters like DP, CE and CXP were optimized (Table. 5, 6). These were classified into neutral loss (NL) scans or precursor ion scans (PIS). This classification is based on the m/z of molecular ion detected at Q1 and product ion detected at quadrapole Q3 as shown in Fig. 19. A multiple precursor ion scan (MPIS) was used to detect fatty acyl groups of polar lipids in negative ion mode as $[M-H]^-$ ion. Head groups in positive ionization method were analyzed as $[M+H]^+$ and $[M+NH_4]^+$. While neutral lipids like TAG and DAGs were only detected in positive ion mode. The NL and PIS scans of TAG and DAG were obtained from ammonium adducts of FA group $[FA+NH_4]^+$ and FA-glycerol $[FA+C_3H_6O+NH_4]^+$ groups respectively (Table 5, 6). Among 13 lipid classes, 12 of them (phosphatidicacid (PA), phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylglycerol (PG), phosphatidylserine (PS), triacylglycerol (TAG), diacylglycerol (DAG), monogalactosyldiglyceride (MGDG), digalactosylglyceride (DGDG) and sulfoquinovosyl diacylglycerols (SQDG)) were detected in positive ionization method as shown in Table. 5. The fragments of these lipid classes were PA-115 m/z, PC-184 m/z, PE-141 m/z, PG-189 m/z, PS-185 m/z, MGDG-359 m/z and DGDG-197 m/z respectively. While phosphatidylionositol (PI – 241 m/z) and fatty acids (FA) of polar lipid classes were detected in negative ionization method (Table. 5). The mass spectra of 12 internal standards, detected in positive ionization mode are shown in Fig. 23. *E. huxleyi* specific lipids were also detected in positive mode through head group fragmentation and formation of $[M+H]^+$ molecular ions. The head group

fragments of these lipids are as follows, PDPT (201.2 m/z), DGCC (104.1 m/z), DGTS (236 m/z), BLL (190.1 m/z), hGSL (180.4 m/z) and sGSL (268.07 m/z). The details of *E. huxleyi* specific lipid fragmentation based on reference lipids were provided in Table. 6.

Neutral lipid, glycolipid and phospholipid class separation

Using the normal phase HPLC chromatography (Fig. 21), I successfully separated all major lipid classes within 50 min (60 min for *E. huxleyi*) in a single run. As shown in Fig. 24, neutral lipids like TAG and DAG were separated from the 13-lipid IST mixture within 9 min. Glycolipids like MGDG, DGDG, SGDG were separated within 23 min. The unique internal standard DNPPE which is used for quantification of crude extracts was detected at 23 min along with PG. Other phospholipids like PE, PI, PA, PS and PC were separated in a sequential order from 25 min to 40 min as show in Fig. 24. Further separation of *E. huxleyi* specific lipids from crude extract was also achieved. Among which, hGSL, sGSL, DGTS, BLL, DGCC and PDPT were separated nicely at 13, 26, 32, 37, 38 and 46 min respectively. Fig. 25 represents the extracted ion chromatogram of all major lipid classes detected in *E. huxleyi* and their retention times.

Cell growth, inorganic phosphate and chlorophyll

E. huxleyi CCMP 2090 cells grown at 20°C in +N batch culture were normal. The cell growth followed a general trend from logarithmic to stationary phase (Fig. 26a). The inorganic phosphate concentration (Pi) in the medium (the initial concentration: 30 µM) was nearly depleted to 1.4 µM under the by day 4 (Fig. 26b). The chlorophyll content appeared to be increasing in all replicates (Fig 26. C) during logarithmic to stationary phase.

Lipidome of *E. huxleyi* CCMP 2090

Lipdome analysis of *E. huxleyi* CCMP 2090 resulted in detection of nearly 600 lipid species in each positive and negative ion mode. PERL algorithm based selection of common lipids from positive and negative ionization resulted in identification of 365 lipid species (Table. 8). This data is lipid-list-1. From the quantification of these lipid species, no significant difference was observed in the accumulation of lipids during logarithmic and stationary phases (Fig. 27). Due to database technical issues, some lipid classes were not properly annotated. Phospholipids (PC and PC/SM) were detected as the major lipid class comprising of 73.5% of total lipidome in logarithmic phase and decreased to 65% during stationary phase. Whereas, glycolipids like MGDG were increased from 12% to 17%, DGDG and SQDG were nearly 2% to 3% during logarithmic phase to stationary phase. On the other hand, *E. huxleyi* specific lipids like sGSL and hGSL were not detected by the database, as no information was available in the database. Only manual observation indicated the presence of these lipids. I also attempted to add the lipid information for PDPT, DGTS, DGCC and BLL to the LipidView database. After addition, the composition of each of these lipid classes was detected to be less than 2% in the total lipidome. Further, I detected TAGs that are composed of FA 18:1 in the lipidome of *E. huxleyi*. However, the compositions of TAGs was very low (2% mol of total lipidome) under both conditions.

From the lipid-list-1, very long chain fatty acids (VLCFAs) like tetracosatetraenoic acid (24:4), tetracosaopentaenoic acid (24:5) and tetracosahexaenoic acid (24:6) in PDPT and DGTS like lipid species were also detected (Fig. 28). The amount of these lipids was also observed to be increased during stationary phase. In

addition, odd chain fatty acid FA 15:0 was also detected from this analysis (Fig. 28c). However, their content appears to be very less, as most of them are not detected in lipid-list-2. Nearly 80 lipid species were identified in lipid-list-2. These were assumed to be highly reliable lipids (Table. 9). Total lipidome processing resulted in detection of many fatty acids. Nearly 42 types of fatty acyl groups were identified as shown in Table. 10. Fatty acyl groups in lipid-list 1 and 2 are mainly composed of FA 14:0. From the qualitative analysis using PERL scripts network models for lipidome of *E. huxleyi* were successfully generated. Fig. 29 and Fig. 30 represent the network models based on the lipidome data from lipid-list-1.

Discussion

The main objective of this study was to establish methods for lipidomic profiling using MS. I identified the lipid classes based on the PIS, NL and MPIS techniques available in MS. These techniques were used for the comprehensive analysis of metabolites categorized by structural similarities of specific fragments (Han and Gross 2003 and 2005, Taguchi et al. 2005 and Ejsing et al. 2006). To improve the purity and detection sensitivity of the sample by ESI-MS, I also focused on establishing lipid separation technique using HPLC column.

Improved detection methods

I successfully optimized the fragmentation and detection parameters for major lipid classes using QTRAP ESI-MS machine. All the 13 lipid standards were successfully fragmented (Table. 4). 12 lipid classes were detected in positive ionization mode and only PI in negative mode as shown in Tables. 5, 6. Current method successfully achieved fragmentation of all most all lipid classes in a single mode. This could be an easy way for naïve users to perform lipidome analysis. To detect the acyl groups of lipids (especially polar lipids) it is necessary to use negative ionization mode (Ejsing et al. 2006). The intensities of detected lipid ISTs in positive mode were observed to be different even at same concentration (Fig. 21). This indicates that the sensitivity of fragmentation of each lipid class varies. Therefore it should be noted that, peak height does not mean actual concentration of the lipid.

Using highly concentrated *E. huxleyi* lipid extract was found to be helpful in optimizing parameters for lipids like GSLs, BLs and PDPT (Table. 6). Interest in

studying these specific lipid classes of marine Haptophyta is increasing recently (Armada et al. 2013, Van Mooy et al. 2009). Among which BLs play a key role under nutrient limitation (Van Mooy et al. 2009), sGSL and hGSL play a significant role in viral infection (Shemi et al. 2015; Rosenwasser et al. 2014; Malitsky et al. 2016; Vardi et al. 2009) and PDPT are major constituents in lipid body formation (Fulton et al. 2014; Harwood and Jones 1989; Hunter at al. 2015). Separation and detection of these lipid classes using QTRAP ESI/MS has not been previously reported. Therefore, current report could be helpful for future research on these lipids.

Lipid separation was achieved by modifying the method of Popendorf et al. (2013) using a normal phase BETASIL DIOL column (Table. 7 and Fig. 21). This resulted in separation of neutral lipids (nonpolar lipids), phospholipids, glycolipids and *E. huxleyi*/Haptophyte specific lipids in a single run (Fig. 24, 25). Previously available methods can separate various membrane lipid (polar lipids) classes within 30-45 min range (Hunter et al. 2015; Strut et al. 2004; Popendorf et al. 2013). However, in many cases fractionation of crude extracts into polar and non polar lipids is needed to detect all lipid classes. While, this improved method can separate almost all lipid classes including neutral lipids without any prior fractionation. Though this method consumes slightly more time for lipid separation (50-60 min), it could be advantageous over others.

Changes in lipidome of *E. huxleyi* and scope for new findings

From the lipidome analysis of *E. huxleyi* CCMP 2090, phospholipids, especially PC lipid class were observed to be dominant. The lipid composition remained similar throughout the experiment (Fig. 27). This indicates no significant influence of

growth stage of the cell on its lipidome. Higher content of PC was also observed in some coccolith forming strains (Pond and Harris 1996). Decrease in PC (8.5% mol) and an increase in MGDG (4% mol) was observed as the cell underwent changes from logarithmic to stationary phase. These slight changes might be due to nutrient concentration. During the stationary phase (day 6 to 12 in Fig. 26a), phosphate depletion can be observed in the medium (Fig. 26b). On the other hand, nitrogen limitation was not observed here as the chlorophyll content appears to be increasing during stationary phase (Fig. 26c). In a separate experiment on N-limitation, I observed that N-limitation occured after 12 days in a batch culture. Therefore the nutrient limitation is surely due to phosphate limitation. Because of this slight decrease in phospholipids might have occurred and chloroplast membrane lipids might be still accumulating. Studies by Malitsky et al. (2016) show that the lipidome of *E. huxleyi* CCMP 2090 is mainly composed of MGDG during logarithmic phase. This is quite opposite to my current finding. Therefore, the differences might be due to differences in culture conditions or lipidome analysis platforms. Moreover, the techniques used to analyze lipidome and database reliability, plays a key role in this kind of analysis. Therefore, validation of differences in mass spectrometric platforms should be made in future. Also, as the sensitivity of fragmentation varies between lipid classes, using an internal standard more similar to phospholipid might be the reason for such results. In order to reduce such effects during quantification, it is suggestible to use standards for each lipid class. In the current method, I tried quantification using DNPPE only, based on previous studies. Further optimization can be done to improve quantification of lipidome.

Qualitative lipidomics and computational methods

Detection of a lipid class and all its constituents (head group and acyl groups) is a crucial step in lipidomics. Till date, it was difficult to compare and combine large amount of lipidome data. So, previous reports were able to provide the lipidome data either from negative or positive ion mode only (Hunter et al. 2015; Malitsky et al. 2016; Milne et al. 2006; Taguchi et al. 2005). Recent advances in computational techniques provide us the opportunity to handle biological data effectively (Yetkuri et al. 2007). Using PERL programming language, I was able to develop simple computational algorithms for this purpose (Fig. 22). By combining the data from both ionization methods, I successfully identified lipid head groups and their fatty acyl tails (Fig. 20, 22 and Table. 8, 9). Out of 600 lipid species, 365 lipid species were detected in both methods and 80 lipid species were highly identified in more than half of the replicates. These 80 lipids must be consistently produced by *E. huxleyi* despite changes in the growth phase.

Current lipidome results of *E. huxleyi* indicate presence of VLCFAs and odd chain fatty acid like FA 15:0 in this strain (Fig. 27). Though they appear in a very low concentration, it is clear that these fatty acids are produced by this strain. These were not previously reported in this strain due to lack of such sensitive equipment. VLCFAs like FA 24:4, FA 24:5 and FA 24:6 were detected mainly in PDPT and DGTS. These FAs were intermediates of mammalian type of DHA biosynthesis “Sprecher” pathway (Sprecher et. al 1992, 1995, 2000). DHA is an essential fatty acid useful as nutrient source in food and industry. It is the most abundant fatty acid in *E. huxleyi* (Bell and Pond 1996). Therefore the accumulation of DHA is speculated to be in multiple ways.

However, general GC-MS techniques did not indicate presence of any intermediates. Previously, the presence of intermediate fatty acids and coenzymes for DHA biosynthesis in *E. huxleyi* from the calculated fatty ESI-MS approach was speculated (Ohi and Shiraiwa 2015). Here, I improved the detection methods and were able to clearly identify the presence of these intermediate fatty acids. Further, from the network view of 365 lipid species, the relation between lipid class and acyl group compositions can be clearly understood (Fig. 29, 30). From Fig. 29, large size of lipid class nodes (spheres) indicate the large number of specific lipid classes. Larger FA 14:0 node indicates that it is present in most of the lipid classes. Such nodes are considered as hub nodes as they have many links with other nodes. In general, these hub nodes are considered as key/central regulators of any pathway. Therefore, one can understand that by targeting such central nodes we can alter any pathway (especially here lipid biosynthesis). Similarly from Fig. 30, we can identify which lipid classes share similar carbon content and how significant the carbon composition for each lipid class. Therefore, these techniques could be effectively useful for understanding lipid metabolism and fatty acid exchange/transport mechanisms across lipid membranes.

In summary, this study successfully optimized detection of almost all major lipid classes in positive ionization mode. On the other hand, the developed LC method provided better separation of all described lipid classes based on their head group polarity. Using the optimized methods, I was able to identify several lipid species in *E. huxleyi* CCMP 2090. The combination of positive and negative ion mode data using computational methods resulted in successful detection of lipids and their acyl components.

General discussion

In part I, carbon allocation mechanism into various photosynthetic products of *E. huxleyi* CCMP 2090 was examined. I mainly focused on production of unique lipids like alkenones (Fig. 8 - 10) and their relation to N-availability and carbon storage mechanism (Fig. 13, 14). Lipidome results from my study indicated that alkenone accumulation was triggered under –N condition. Nearly 4 fold accumulation in alkenone content was observed, while no significant increase in other lipid content (neutral lipids like TAGs and membrane lipids) was detected (Fig. 11, 12). This assures the role of alkenones as storage lipids instead of TAGs. This could be very useful for biofuel research (Wu et al. 1999; O'Neil et al. 2015, 2016). Further, same levels of total fatty acids content in polar lipid fraction under +N and –N suggest that alkenone biosynthesis might not be strongly dependent on membrane lipid biosynthesis. However, the role of FA biosynthesis in alkenone accumulation and TAG accumulation cannot be ruled out completely (Rotnani et al. 2006). A future sensitive lipidome analysis might be needed. Also, by inducing alkenone production using such –N conditions, I might be able to identify genome, proteome or metabolite precursors involved in alkenone biosynthesis. This might lead to development of advanced methods for industrial level alkenone production for biofuel synthesis. On the other hand, geophysical studies that use alkenones as paleothermometers could also benefit from the current work (Brassell 1986; Müller et al. 1998; Prahl and Wakeham 1987). My results showed that the $U^{K'}_{37}$ index and accumulation of different types of unsaturated alkenones were not affected by N-limitation (Fig. 12, 10b). Therefore, alkenone unsaturation can be used as a temperature calibration tool even for samples collected from low nitrogen environments.

The unsaturation responses to nutrient limitation are more similar to some of the noncoccolith strains Popp et al. (1998) and opposite to coccolith forming strains (Pan and Sun (2011). The reason for these differences could be genetic; however there is no information on the actual molecular mechanism.

Further, a dynamic carbon directing and redirection mechanism was also identified in *E. huxleyi* CCMP 2090 in my study (Fig. 14, 18). Accumulation of more carbon into alkenones under –N conditions shows that alkenones are the immediate C sinks. These results provide a clear evidence for the role of alkenones as storage components (Eltgroth et al. 2005; Epstein et al. 1998; Prahl et al. 2003). Marine phytoplankton like *E. huxleyi* is responsible for driving global carbon cycle through photosynthetic C fixation into various components (Field et al. 1998). On the other hand, this is also linked with *E. huxleyi* primary production which is regulated by nutrient availability through ocean upwelling processes (Behrenfeld et al. 2006). Despite depletion of nitrogen by diatom blooms in the oceans, *E. huxleyi* continues to bloom (Holligan et al., 1993). This is due to presence of N-scavenging mechanisms (Rokitta et al 2014) which are simultaneously activated with carbon metabolism. Current results provide significant information on the C allocation in non coccolith strains of *E. huxleyi* and their survival under N-limitation conditions. Most of the carbon accumulated in the form of protein is re-allocated into alkenones (27% Ca.). In natural conditions, during algal blooms cells are exposed to high light and nutrient depletion (Iglesias-Rodriguez et al. 2002). During this condition storage of carbon in the form of TAGs might be subjected to photo oxidation. Whereas in alkenones, their trans-double bond geometry provides a more photostable form of energy storage (Eltgroth et al. 2005). Therefore,

accumulation of alkenones can be speculated as an obvious phenomina to store energy for growth and survival even in natural N-limited conditions. In addition to C storage into alkenones, I also speculate that LMCs (like mannitol) are accumulated (Fig. 17). This could be a secondary storage pool besides alkenones, as they can accommodate nearly 35% Ca. in several strains of *E. huxleyi* (Obate et al. 2015; Tsuji et al. 2015). All my findings suggest that, alkenones play a key role under N-limitation. This also builds hopes for future research on alkenone based climatic studies and development of alternate energy systems.

In Part II, I focused on developing methods for lipidomics. Here, I succeeded in improving LC separation and ESI/MS based lipid profiling methods. The optimized detection of almost all major lipid classes in positive ionization mode is beneficial for quick lipid analysis (Table. 5). This provides a good alternative for many MS lipid detection methods which require frequent transition from positive ionization to negative ionization modes (Brugger et al. 1997; Ejsing et al. 2006; Taguchi et al. 2005). Successful optimization of novel lipid fragmentation and detection from concentrated crude extracts was achieved in this study (Table. 6). These simple strategies could be beneficial for detecting lipids that do not have commercial standards. From these, previously described *E. huxleyi* specific lipids like hGSL and sGSL were detected along with some PDPT and beatine lipids (Shemi et al. 2015; Rosenwasser et al. 2014). On the other hand, the developed LC method provided better separation of all described lipid classes based on their head group polarity (Table. 7, Fig. 21, 24). This is an improvement of normal phase chromatographic technique and useful for most kinds of biological samples. However, due to limited database sources for Haptophyte specific

lipids and no MS-data processing tools several challenges are yet to be addressed. My work is the first attempt in my research group to establish advanced and sensitive mass spectrometry analysis methods. Despite hurdles in setting up from scratch, the developed lipid profiling methods show promising sensitivity than available GC/MS methods. Using the optimized methods, I am able to identify nearly 600 lipid species *E. huxleyi* CCMP 2090. Further, the developed computational algorithms generated qualitative data of 80-365 most reliable lipid species (Table. 8, 9). Unknown mechanisms like mammalian type DHA biosynthesis pathway in *E. huxleyi* were also speculated in this study (Fig. 28). Further, I also developed a platform to use network models to study lipid metabolism (Fig. 29, 30). From the improved lipid detection, separation and computational methods, I showed that a comprehensive lipid profiling can be made. This could therefore be helpful towards exploring alkenone and lipid biosynthesis mechanisms in future.

Tables and Figures

Table. 1. Carbon content of *E. huxleyi* cellular components from chemical structure

Cellular components	Composition	Carbon content
Neutral polysaccharides	Glucose	0.40 (β-glucan structure)
Acid polysaccharides	Galactouronic acid, mannose, Rhamminose, xylose, Ribose/arabinose	0.39 (Fichtinger-Schepman et al. 1981)
Proteins	Amino acids	0.53 (Laws , 1991)
Alkenones	C_{37}- C_{39},	0.84
Alkenes	C_{31} and C_{33}	0.86
Polar Lipids	All fatty acids detected	0.78

Table. 2. FTIR spectral band assignment with reference to Fig.15. Under lined wave numbers used for semi quantification of relate macromolecules

Label	Band assignment (wave number Cm^{-1})	Functional groups ≈ wavelength region	Macromolecule class
A	1732, <u>1734</u>, 1737	Symmetric strech($_v$) C=O ≈ 1745	Ester of lipids and fatty acids - LIPIDS
*****	1705.5	($_v$)C=O ≈ 1705-1706	Ketone (alkenone related)
B	<u>1655</u>, 1651	($_v$) C=O≈ 1655	PROTEINS (Amide I)
C	<u>1057</u>, 1060, 1163, 1165	($_v$) C-O-C ≈ 1200–980	CARBOHYDRATES
D	~<u>962.5</u>	Trans C=C ≈ 960-965	ALKENONES

Table. 3. Percentage of lipids measured from total lipid content

Time(d) / Lipids	4		5		6		7		8	
	%lipids	**± SD**	**%lipids**	**± SD**	**%lipids**	**± SD**	**%lipids**	**± SD**	**%lipids**	**± SD**
N-sufficient condition										
Alkenones	59.6	± 5.3	61.8	± 2.4	56.6	± 3.4	55.3	± 3.1	57.8	± 5.7
Alkenes	3.8	± 0.4	5.1	± 1.1	4.1	± 0.3	4.1	± 0.2	4.6	± 0.5
Fatty acids	36.9	± 5.6	33.1	± 2.0	39.3	± 3.7	40.6	± 3.2	37.5	± 6.2
N-limitation condition										
Alkenones	43.3	± 4.1	66.1	± 1.5	68.8	± 2.4	74.9	± 2.2	77.3	± 3.7
Alkenes	4.1	± 0.8	5.2	± 0.3	6.1	± 0.8	5.9	± 0.4	5.9	± 0.1
Fatty acids	52.6	± 5.2	28.7	± 1.8	25.1	± 2.3	19.1	± 2.0	16.8	± 3.8

Table .4. Lipid internal standards (ISTs) information

S.No	IST	Full Name	Solvent (stocks)
1	PA(17:0/17:0)	1,2-diheptadecanoyl-sn-glycero-3-phosphate	CHCl3/MeOH/H2O (65:35:8)
2	PC(17:0/17:0)	1,2-diheptadecanoyl-sn-glycero-3-phosphocholine	MeOH:CHCl3 (2:1)
3	PE(17:0/17:0)	1,2-diheptadecanoyl-sn-glycero-3-phosphoethanolamine	CHCl3/MeOH/H2O(4:1:0.1)
4	PG(17:0/17:0)	1,2-diheptadecanoyl-sn-glycero-3-phospho-(1'-rac-glycerol)	CHCl3/MeOH/H2O
5	PI(18:1(9Z)/18:1(9Z))	1,2-dioleoyl-sn-glycero-3-phospho-(1'-myo-inositol)	MeOH:CHCl3
6	PS(17:0/17:0)	1,2-diheptadecanoyl-sn-glycero-3-phospho-L-serine	CHCl3/MeOH/H2O
7	DGTS (d9)	1,2-dipalmitoyl-sn-glycero-3-O-4'-N,N,N-trimethyl(d9)]-homoserine	MeOH:CHCl3
8	DG d5-(17:0/0:0/17:0)	1,3(d5)-diheptadecanoyl-glycerol	MeOH:CHCl3
9	TGd5-(17:0/17:1(10Z)/17:0)	1,3(d5)-diheptadecanoyl-2-(10Z-heptadecenoyl)-glycerol	CHCl3
10	MGDG	Monogalactosyldiacylglycerol	CHCl3/MeOH/H2O
11	DGDG	Digalactosyldiacylglycerol	CHCl3/MeOH/H2O
12	SQDG(18:3/16:0)	Sulfoquinovosyldiacylglycerol	CHCl3
13	16:0 DNP PE	1,2-dihexadecanoyl-sn-glycero-3-phosphoethanolamine-N-(2,4-dinitrophenyl)	CHCl3

Table. 5. Lipid ISTs ESI/MS fragmentation parameters

S.No	Lipid class	molecular ion	Scan Type	Q3-m/z	kind of fragment	DP	CE	CXP
Positive ionization method								
1	PA	[M+NH4]+	NL	115	PA[PA+NH4]	126	23	36
2	PC	[M+H]+	PIS	184	LPC/PC+H	286	41	10
3	PE	[M+H]+	NL	141	PE[PE/LysoPE]	71	30	36
4	PG	[M+NH4]+	NL	189	PG [phosphorylglycerol+NH4]	136	28	36
5	PS	[M+H]+	NL	185	PS(headgroup)	116	21	36
6	DAG	[M+NH4]+	PIS	332	HEAD-dg-d5	101	35	32
7	DNP-PE	[M+NH4]+	NL	324	DNP-FA16	86	35	33
8	DGTS	[M+H]+	PIS	245	DGTS-D9 (head group)	46	63	14
	DGTS	[M+Na]+	NL	96	N+(CH3)3	101	46	39
9	DGDG	[M+NH4]+	NL	359	Diglycosyl	141	39	36
10	MGDG	[M+NH4]+	NL	197	Monoglycosyl	151	36	33
11	SQDG	[M+NH4]+	NL	261	SQDG[SQDG+NH4]	106	25	34
12	TAG	[M+NH4]+	NL	287	TAG	111	36	35
Negative ionization method								
13	PI	$[M-H]^-$	PIS	241	Inositol phosphate-H2O	-210	-52	-15
14	FA of lipid	$[M-H]^-$	MPIS	m/z of FA	FA-$[H]^-$	-100	-50	-15

Table. 6. *E. huxleyi* specific lipids and their fragmentation parameters.

S.No	m/z of lipids	Scan Type	PIS/NL	Fragment type	Formula	Lipid class	DP	EP	CE	CXP	Reference
1	710.6	PIS	236	DGTS [headgroup + H^+]	C10H22O5N	DGTS	150	10	55	22	1
2	856.4	PIS	190.1	BLL [headgroup + H^+]	C7H12O5	BLL	55	10	40	10	1,2
3	872	PIS	104.1	DGCC [headgroup + H^+]	C5H14ON	DGCC	226	10	120	12.5	1,2
4	849.5	PIS	201.2	PDPT [headgroup + H^+]	C5H14O4PS	PDPT	200	10	42	10.8	1,2
5	870.4	NL	268.07	sGSL [headgroup + H^+]	C9H14O8+H2O	sGSL	200	10	33.2	38.19	1,2
6	806	NL	180.4	hGSL [headgroup + H^+]	C6H12O6	hGSL	200	10	20	35	3
7	-	PIS	264.27	Ceramide backbone	d18:1	Cer	140	10	32	15	4,5,6
9	-	PIS	369.3	sterol head	C27H45	Chol	210	10	25	10	6

References: 1.Maltisky et al., 2016, 2. Fulton et al., 2014,3. Vardi et al., 2012, 4. Rebecca L.Shanner, 2009, 5. Hsu and Turk, 2000, 6. Hunter et al., 2015; Lipid abbreviation are as follows: phosphatidyldimethylpropanethiol(PDPT),Diacylglyceryl-3-O-carboxyhydroxymethylcholine (DGCC), Diacylglyceryl trimethylhomoserine (DGTS), betaine like lipids (BLL), host specific glyco sphingo lipids (hGSL) and salicylic glycol sphingo lipid (sGSL), Cerramide (cer), Cholesterol (Chol)

Table. 7. HPLC gradient elution parameters. Table is provided with reference to Fig. 20. A and B represent elution buffers used in chromatography. A= 600:400:1.0:0.4 n-hexane:isopropanol:formic acid: 25 % aqueous NH_4; eluent B= 880:120:1.0:0.4 isopropanol:water:formic acid: 25 % aqueous NH_4.

S.No	Time (min)	Flow rate (µl min^{-1})	%A	%B
0	0	0	100	0
1	8	100	68.5	31.5
2	15	100	48.5	51.5
3	27	100	28.5	71.5
4	32	100	28.5	71.5
5	35	100	100	0
6	50.2	100	100	0

Table. 8. Lipidome data of *E. huxleyi* CCMP 2090. Table representing lipid class and fatty acyl groups (FA1/FA2) detected by combination of negative and positive ionization method

S.No	Lipid	Fatty acyl groups	mass (m/z)	Log-mol%	Sta-mol%
1	BLL(24:0)	(FA 14:0/FA 10:0)	628	0.00	0.00
2	BLL(36:3)	(FA 18:1/FA 18:2)	791	0.07	0.01
3	BLL(36:6)	(FA 18:1/FA 18:5)	786.5	0.33	1.15
4	BLL(38:2)	(FA 14:1/FA 24:1)	820.5	0.00	0.00
5	BLL(38:3)	(FA 14:0/FA 24:3)	819	0.00	0.00
6	BLL(38:5)	(FA 18:1/FA 20:4)	816.5	0.02	0.07
7	BLL(38:6)	(FA 18:1/FA 20:5)	814.5	0.02	0.19
8	BLL(40:0)	(FA 14:0/FA 26:0)	853	0.00	0.00
9	BLL(40:2)	(FA 14:0/FA 26:2)	848.5	0.00	0.00
10	BLL(40:3)	(FA 18:1/FA 22:2)	848.5	0.10	0.12
11	BLL(40:5)	(FA 18:1/FA 22:4)	845	0.11	0.00
12	BLL(40:6)	(FA 18:1/FA 22:5)	843	0.02	0.00
13	BLL(42:4)	(FA 18:1/FA 24:3)	874.5	0.00	0.01
14	BLL(42:7)	(FA 18:5/FA 24:2)	866.5	0.00	0.00
15	BLL(44:3)	(FA 18:1/FA 26:2)	904.5	0.01	0.07
16	BLL(44:6)	(FA 18:1/FA 26:5)	898.5	0.25	0.19
17	BLL(48:6)	(FA 22:6/FA 26:0)	952.5	0.00	0.00
18	BLL(48:8)	(FA 22:6/FA 26:2)	948.5	0.00	0.00
19	DAG(26:0)	(FA 14:0/FA 12:0)	483.5	0.00	0.00
20	DAG(28:0)	(FA 16:0/FA 12:0)	511.5	0.00	0.00
21	DAG(30:0)	(FA 16:0/FA 14:0)	539	0.03	0.00
22	DAG(30:4)	(FA 14:0/FA 16:4)	531	0.00	0.00
23	DAG(32:1)	(FA 16:0/FA 16:1)	565	0.00	0.00
24	DAG(32:1)	(FA 18:1/FA 14:0)	584	0.01	0.01
25	DAG(32:2)	(FA 18:1/FA 14:1)	582.5	0.05	0.14
26	DAG(32:4)	(FA 16:0/FA 16:4)	559	0.00	0.00
27	DAG(34:2)	(FA 20:2/FA 14:0)	591.5	0.00	0.00
28	DAG(36:4)	(FA 22:4/FA 14:0)	615.5	0.00	0.00
29	DAG(36:5)	(FA 22:5/FA 14:0)	614	0.00	0.00
30	DAG(36:6)	(FA 22:6/FA 14:0)	612	0.03	0.00

31	DAG(36:9)	(FA 22:6/FA 14:3)	605.5	0.00	0.00
32	DAG(40:0)	(FA 16:0/FA 24:0)	680	0.00	0.00
33	DAG(40:1)	(FA 14:0/FA 26:1)	677.5	0.00	0.00
34	DAG(40:4)	(FA 14:0/FA 26:4)	672	0.00	0.00
35	DAG(40:5)	(FA 14:0/FA 26:5)	670	0.00	0.00
36	DAG(42:4)	(FA 16:0/FA 26:4)	700	0.00	0.00
37	DAG(46:4)	(FA 20:0/FA 26:4)	775	0.00	0.00
38	DGCC(30:0)	(FA 14:0/FA 16:0)	625.5	0.00	0.00
39	DGCC(36:2)	(FA 14:0/FA 22:2)	705.5	0.02	0.00
40	DGCC(40:1)	(FA 20:0/FA 20:1)	765.5	0.00	0.14
41	DGCC(42:1)	(FA 20:0/FA 22:1)	793.5	0.01	0.00
42	DGCC(44:1)	(FA 18:1/FA 26:0)	844	0.15	0.00
43	DGDG(30:0)	(FA 14:0/FA 16:0)	863.5	0.00	0.00
44	DGDG(32:0)	(FA 16:0/FA 16:0)	951	0.00	0.00
45	DGDG(32:1)	(FA 14:0/FA 18:1)	949.5	0.04	0.00
46	DGDG(32:1)	(FA 18:1/FA 14:0)	949	0.00	0.35
47	DGDG(32:4)	(FA 14:0/FA 18:4)	944.5	0.01	0.00
48	DGDG(32:5)	(FA 14:1/FA 18:4)	882	0.01	0.00
49	DGDG(34:0)	(FA 14:0/FA 20:0)	980	0.00	0.00
50	DGDG(34:1)	(FA 14:0/FA 20:1)	978.5	0.01	0.00
51	DGDG(34:2)	(FA 14:0/FA 20:2)	975	0.01	0.00
52	DGDG(34:3)	(FA 14:0/FA 20:3)	913.5	0.00	0.00
53	DGDG(34:3)	(FA 16:0/FA 18:3)	932	0.17	0.00
54	DGDG(34:5)	(FA 14:0/FA 20:5)	969	0.01	0.01
55	DGDG(34:6)	(FA 14:0/FA 20:6)	926	0.00	0.00
56	DGDG(36:1)	(FA 14:0/FA 22:1)	945.5	0.01	0.01
57	DGDG(36:3)	(FA 14:0/FA 22:3)	941	0.00	0.00
58	DGDG(36:7)	(FA 14:1/FA 22:6)	933.5	0.10	0.00
59	DGDG(36:8)	(FA 18:3/FA 18:5)	991	0.00	0.00
60	DGDG(38:0)	(FA 14:0/FA 24:0)	975.5	0.01	0.00
61	DGDG(38:2)	(FA 14:1/FA 24:1)	971	0.00	0.02
62	DGDG(38:3)	(FA 14:0/FA 24:3)	988.5	0.06	0.00
63	DGDG(38:4)	(FA 14:0/FA 24:4)	967	0.00	0.00
64	DGDG(38:4)	(FA 16:0/FA 22:4)	968	0.00	0.00
65	DGDG(38:5)	(FA 14:0/FA 24:5)	966	0.01	0.00
66	DGDG(40:11)	(FA 18:5/FA 22:6)	981	0.00	0.00

67	DGDG(40:2)	(FA 14:0/FA 26:2)	999.5	0.00	0.00
68	DGDG(40:3)	(FA 14:0/FA 26:3)	998	0.00	0.00
69	DGDG(40:4)	(FA 18:0/FA 22:4)	995.5	0.00	0.00
70	DGDG(40:5)	(FA 16:0/FA 24:5)	993	0.00	0.00
71	DGDG(40:6)	(FA 16:0/FA 24:6)	992	0.00	0.00
72	DGDG(40:7)	(FA 18:5/FA 22:2)	989.5	0.08	0.00
73	DGDG(40:8)	(FA 18:3/FA 22:5)	987.5	0.00	0.00
74	DGTS(26:1)	(FA 14:0/FA 12:1)	724	0.28	0.00
75	DGTS(28:0)	(FA 14:0/FA 14:0)	732	0.96	0.00
76	DGTS(28:1)	(FA 18:1/FA 10:0)	730.5	0.07	0.01
77	DGTS(38:1)	(FA 14:0/FA 24:1)	868.5	0.15	0.01
78	DGTS(38:3)	(FA 14:0/FA 24:3)	865	0.00	0.04
79	DGTS(38:5)	(FA 14:0/FA 24:5)	860.5	0.00	1.51
80	DGTS(40:0)	(FA 14:0/FA 26:0)	899	0.00	0.01
81	DGTS(40:1)	(FA 14:0/FA 26:1)	896.5	0.01	0.01
82	DGTS(40:1)	(FA 20:0/FA 20:1)	916	0.00	0.00
83	DGTS(40:3)	(FA 14:0/FA 26:3)	892.5	0.00	0.00
84	DGTS(40:5)	(FA 14:0/FA 26:5)	888.5	0.11	0.00
85	DGTS(40:6)	(FA 16:3/FA 24:3)	886.5	0.00	0.00
86	DGTS(42:1)	(FA 16:0/FA 26:1)	925	0.00	0.00
87	DGTS(42:4)	(FA 16:0/FA 26:4)	919	0.00	0.00
88	DGTS(42:5)	(FA 16:0/FA 26:5)	916.5	0.00	0.02
89	DGTS(42:6)	(FA 18:1/FA 24:5)	915	0.00	0.00
90	DGTS(44:11)	(FA 22:6/FA 22:5)	933	0.00	0.00
91	DGTS(44:4)	(FA 18:2/FA 26:2)	946.5	0.00	0.00
92	DGTS(44:7)	(FA 18:1/FA 26:6)	960	0.09	0.00
93	DGTS(44:7)	(FA 18:5/FA 26:2)	940.5	0.00	0.00
94	DGTS(44:9)	(FA 18:3/FA 26:6)	936.5	0.01	0.00
95	DGTS(46:10)	(FA 22:6/FA 24:4)	963	0.00	0.00
96	DGTS(46:5)	(FA 22:1/FA 24:4)	972.5	0.01	0.05
97	DGTS(46:8)	(FA 22:6/FA 24:2)	966.5	0.00	0.00
98	DGTS(46:9)	(FA 20:3/FA 26:6)	965	0.00	0.00
99	DGTS(48:10)	(FA 22:6/FA 26:4)	990.5	0.00	0.00
100	DGTS(48:11)	(FA 22:6/FA 26:5)	988.5	0.00	0.00
101	LPI(30:0)	(FA 15:0/FA 15:0)	767	0.12	0.00
102	LPI(30:1)	(FA 15:0/FA 15:1)	765.5	0.06	0.00

103	MGDG(24:0)	(FA 14:0/FA 10:0)	636	0.02	0.00
104	MGDG(26:0)	(FA 14:0/FA 12:0)	664	0.08	0.00
105	MGDG(28:0)	(FA 14:0/FA 14:0)	692	0.78	0.01
106	MGDG(32:1)	(FA 14:0/FA 18:1)	787.5	0.10	0.01
107	MGDG(34:1)	(FA 14:0/FA 20:1)	816	0.53	0.21
108	MGDG(36:0)	(FA 14:0/FA 22:0)	785.5	0.12	0.19
109	MGDG(36:1)	(FA 14:0/FA 22:1)	844	0.21	0.66
110	MGDG(36:8)	(FA 18:4/FA 18:4)	788	0.07	0.03
111	MGDG(38:1)	(FA 14:0/FA 24:1)	811.5	0.10	0.03
112	MGDG(38:2)	(FA 14:0/FA 24:2)	870.5	0.92	0.01
113	MGDG(38:3)	(FA 14:0/FA 24:3)	867	0.01	0.01
114	MGDG(38:6)	(FA 16:0/FA 22:6)	820	0.21	0.01
115	MGDG(40:1)	(FA 14:1/FA 26:0)	840	0.03	0.01
116	MGDG(40:1)	(FA 16:0/FA 24:1)	900.5	0.00	0.00
117	MGDG(40:2)	(FA 14:1/FA 26:1)	838	0.02	0.06
118	MGDG(40:2)	(FA 16:0/FA 24:2)	898	0.00	0.02
119	MGDG(40:3)	(FA 14:0/FA 26:3)	895	0.00	0.21
120	MGDG(40:6)	(FA 14:0/FA 26:6)	890.5	0.01	0.00
121	MGDG(44:0)	(FA 18:0/FA 26:0)	958	0.00	0.00
122	MGDG(46:6)	(FA 22:6/FA 24:0)	974.5	0.00	0.00
123	PA(30:0)	(FA 16:0/FA 14:0)	619.5	0.01	0.00
124	PA(30:4)	(FA 14:0/FA 16:4)	611	0.01	0.00
125	PA(32:1)	(FA 14:0/FA 18:1)	646	0.00	0.00
126	PA(34:1)	(FA 14:0/FA 20:1)	674	0.00	0.00
127	PA(34:2)	(FA 14:0/FA 20:2)	671.5	0.00	0.00
128	PA(34:5)	(FA 14:0/FA 20:5)	666	0.00	0.00
129	PA(36:0)	(FA 14:0/FA 22:0)	703	0.00	0.00
130	PA(36:1)	(FA 14:0/FA 22:1)	701	0.00	0.00
131	PA(36:1)	(FA 18:0/FA 18:1)	701.5	0.00	0.00
132	PA(36:2)	(FA 14:0/FA 22:2)	699.5	0.00	0.00
133	PA(36:6)	(FA 14:0/FA 22:6)	692	0.00	0.00
134	PA(38:1)	(FA 14:0/FA 24:1)	730	0.07	0.00
135	PA(38:2)	(FA 14:0/FA 24:2)	727	0.01	0.00
136	PA(40:9)	(FA 24:4/FA 16:5)	741.5	0.00	0.00
137	PA(46:1)	(FA 22:0/FA 24:1)	841.5	0.00	0.00
138	PC(28:3)	(FA 14:0/FA 14:3)	672.5	0.03	0.00

139	PC(30:0)	(FA 14:0/FA 16:0)	704	0.00	0.00
140	PC(30:1)	(FA 16:0/FA 14:1)	705	0.00	0.00
141	PC(30:3)	(FA 14:0/FA 16:3)	698.5	0.00	0.00
142	PC(30:4)	(FA 14:0/FA 16:4)	697	0.00	0.00
143	PC(32:1)	(FA 14:0/FA 18:1)	730.5	0.01	0.00
144	PC(32:2)	(FA 14:0/FA 18:2)	728.5	0.03	0.00
145	PC(32:3)	(FA 14:0/FA 18:3)	726.5	0.00	0.00
146	PC(34:0)	(FA 14:0/FA 20:0)	760	0.00	0.00
147	PC(34:1)	(FA 14:0/FA 20:1)	758.5	0.01	0.00
148	PC(34:2)	(FA 14:0/FA 20:2)	756.5	0.02	0.00
149	PC(34:3)	(FA 14:0/FA 20:3)	754	0.00	0.00
150	PC(34:3)	(FA 18:1/FA 16:2)	755	0.01	0.00
151	PC(34:4)	(FA 16:0/FA 18:4)	752.5	0.00	0.00
152	PC(34:5)	(FA 14:1/FA 20:4)	750.5	0.02	0.00
153	PC(36:2)	(FA 16:0/FA 20:2)	784	0.03	0.00
154	PC(36:3)	(FA 14:0/FA 22:3)	783	0.01	0.00
155	PC(36:4)	(FA 14:0/FA 22:4)	780.5	0.00	0.00
156	PC(36:5)	(FA 18:0/FA 18:5)	778.5	0.00	0.00
157	PC(36:6)	(FA 22:6/FA 14:0)	776.5	0.01	0.00
158	PC(38:1)	(FA 14:1/FA 24:0)	814	0.01	0.00
159	PC(38:2)	(FA 14:0/FA 24:2)	812.5	0.05	0.01
160	PC(38:3)	(FA 14:1/FA 24:2)	810.5	0.07	0.00
161	PC(38:4)	(FA 14:1/FA 24:3)	810	0.01	0.00
162	PC(38:4)	(FA 14:1/FA 24:3)	808	0.00	0.02
163	PC(38:4)	(FA 18:1/FA 20:3)	809	0.06	0.01
164	PC(38:7)	(FA 14:1/FA 24:6)	803	0.00	0.04
165	PC(40:1)	(FA 16:0/FA 24:1)	842.5	0.01	0.00
166	PC(40:1)	(FA 18:1/FA 22:0)	843	0.00	0.00
167	PC(40:2)	(FA 14:0/FA 26:2)	841	0.03	0.01
168	PC(40:5)	(FA 14:0/FA 26:5)	834.5	0.17	0.00
169	PC(42:0)	(FA 16:0/FA 26:0)	873	0.00	0.00
170	PC(42:0)	(FA 18:0/FA 24:0)	872.5	0.00	0.00
171	PC(42:2)	(FA 16:0/FA 26:2)	869	0.01	0.00
172	PC(42:7)	(FA 22:6/FA 20:1)	861	0.00	0.04
173	PC(42:9)	(FA 20:4/FA 22:5)	855	0.00	0.00
174	PC(44:8)	(FA 22:6/FA 22:2)	887	0.08	0.00

175	PC(46:11)	(FA 22:6/FA 24:5)	906.5	0.00	0.02
176	PC(46:11)	(FA 22:6/FA 24:5)	929	0.00	0.00
177	PC(46:3)	(FA 20:1/FA 26:2)	922.5	0.00	0.00
178	PC(46:4)	(FA 20:2/FA 26:2)	921	0.00	0.00
179	PC(46:8)	(FA 22:6/FA 24:2)	912.5	0.00	0.00
180	PC(46:8)	(FA 22:6/FA 24:2)	915	0.00	0.00
181	PC(48:11)	(FA 22:6/FA 26:5)	934.5	0.00	0.00
182	PC(48:7)	(FA 22:6/FA 26:1)	942.5	0.01	0.01
183	PC(48:9)	(FA 22:6/FA 26:3)	939	0.01	0.00
184	PDPT(28:0)	(FA 16:0/FA 12:0)	695.5	0.00	0.00
185	PDPT(30:0)	(FA 14:0/FA 16:0)	724	0.00	0.00
186	PDPT(30:0)	(FA 16:0/FA 14:0)	723.5	0.00	0.01
187	PDPT(30:1)	(FA 14:0/FA 16:1)	722	0.00	0.00
188	PDPT(30:1)	(FA 15:0/FA 15:1)	721.5	0.00	0.00
189	PDPT(30:2)	(FA 14:0/FA 16:2)	719.5	0.00	0.06
190	PDPT(30:3)	(FA 14:0/FA 16:3)	717	0.14	0.00
191	PDPT(30:4)	(FA 14:0/FA 16:4)	716	0.62	0.00
192	PDPT(32:2)	(FA 14:0/FA 18:2)	748	0.00	0.00
193	PDPT(32:3)	(FA 18:0/FA 14:3)	746	0.00	0.02
194	PDPT(32:4)	(FA 14:0/FA 18:4)	744	0.37	0.00
195	PDPT(32:5)	(FA 14:0/FA 18:5)	742	0.01	0.00
196	PDPT(34:1)	(FA 14:0/FA 20:1)	777.5	0.00	0.01
197	PDPT(34:2)	(FA 14:0/FA 20:2)	776	0.00	0.00
198	PDPT(34:3)	(FA 14:0/FA 20:3)	773.5	0.00	0.00
199	PDPT(34:4)	(FA 14:0/FA 20:4)	771.5	0.09	0.01
200	PDPT(34:5)	(FA 14:0/FA 20:5)	770	0.03	0.00
201	PDPT(34:6)	(FA 14:0/FA 20:6)	767.5	0.00	0.00
202	PDPT(36:1)	(FA 16:0/FA 20:1)	806	0.00	0.00
203	PDPT(36:2)	(FA 14:0/FA 22:2)	804	0.01	0.00
204	PDPT(36:2)	(FA 14:0/FA 22:2)	803.5	0.01	0.00
205	PDPT(36:3)	(FA 20:2/FA 16:1)	802	0.00	0.00
206	PDPT(36:4)	(FA 14:0/FA 22:4)	800	0.01	0.00
207	PDPT(36:4)	(FA 16:0/FA 20:4)	799.5	0.00	0.00
208	PDPT(36:5)	(FA 16:0/FA 20:5)	797.5	0.01	0.00
209	PDPT(36:6)	(FA 14:0/FA 22:6)	797.5	0.00	0.00
210	PDPT(38:2)	(FA 16:0/FA 22:2)	831.5	0.00	0.00

211	PDPT(38:3)	(FA 14:0/FA 24:3)	830	0.04	0.00
212	PDPT(38:3)	(FA 18:0/FA 20:3)	829.5	0.03	0.00
213	PDPT(38:4)	(FA 16:0/FA 22:4)	828	0.04	0.00
214	PDPT(38:5)	(FA 14:0/FA 24:5)	826	0.00	0.00
215	PDPT(38:6)	(FA 16:0/FA 22:6)	824	0.00	0.00
216	PDPT(40:0)	(FA 14:0/FA 26:0)	864	0.00	0.00
217	PDPT(40:1)	(FA 14:0/FA 26:1)	862	0.02	0.00
218	PDPT(40:2)	(FA 14:0/FA 26:2)	860	0.00	0.00
219	PDPT(40:3)	(FA 14:0/FA 26:3)	858.5	0.00	0.00
220	PDPT(40:3)	(FA 22:0/FA 18:3)	857.5	0.00	0.00
221	PDPT(40:4)	(FA 16:0/FA 24:4)	855.5	0.05	0.00
222	PDPT(40:5)	(FA 14:0/FA 26:5)	854	0.01	0.01
223	PDPT(40:5)	(FA 18:5/FA 22:0)	853.5	0.00	0.00
224	PDPT(40:6)	(FA 14:0/FA 26:6)	852	0.02	0.00
225	PDPT(42:11)	(FA 22:6/FA 20:5)	869.5	0.00	0.00
226	PDPT(42:2)	(FA 16:0/FA 26:2)	888	0.00	0.00
227	PDPT(42:6)	(FA 22:6/FA 20:0)	880	0.00	0.00
228	PDPT(42:8)	(FA 22:6/FA 20:2)	876	0.03	0.00
229	PDPT(44:1)	(FA 18:1/FA 26:0)	917.5	0.00	0.01
230	PDPT(44:10)	(FA 22:6/FA 22:4)	899.5	0.00	0.00
231	PDPT(44:12)	(FA 24:6/FA 20:6)	895.5	0.00	0.00
232	PDPT(44:3)	(FA 18:3/FA 26:0)	914	0.00	0.00
233	PDPT(44:5)	(FA 18:3/FA 26:2)	910.5	0.01	0.00
234	PDPT(46:6)	(FA 22:6/FA 24:0)	936	0.01	0.00
235	PDPT(46:8)	(FA 22:6/FA 24:2)	931.5	0.12	0.00
236	PDPT(48:11)	(FA 22:6/FA 26:5)	954	0.00	0.00
237	PDPT(48:12)	(FA 26:6/FA 22:6)	951.5	0.00	0.00
238	PDPT(48:2)	(FA 24:1/FA 24:1)	972	0.00	0.00
239	PDPT(48:6)	(FA 22:6/FA 26:0)	963.5	0.17	0.00
240	PDPT(48:6)	(FA 22:6/FA 26:0)	964	0.00	0.00
241	PDPT(48:7)	(FA 22:6/FA 26:1)	961.5	0.03	0.01
242	PDPT(48:8)	(FA 22:6/FA 26:2)	960	0.00	0.00
243	PE(26:0)	(FA 14:0/FA 12:0)	606.5	0.01	0.00
244	PE(28:0)	(FA 14:0/FA 14:0)	634.5	0.00	0.00
245	PE(30:0)	(FA 14:0/FA 16:0)	662	0.00	0.00
246	PE(30:0)	(FA 16:0/FA 14:0)	663	0.00	0.00

247	PE(30:2)	(FA 14:0/FA 16:2)	659	0.00	0.00
248	PE(30:2)	(FA 16:0/FA 14:2)	658	0.00	0.00
249	PE(30:3)	(FA 14:0/FA 16:3)	656	0.08	0.00
250	PE(32:0)	(FA 14:0/FA 18:0)	690.5	0.00	0.00
251	PE(32:2)	(FA 14:0/FA 18:2)	686.5	0.00	0.00
252	PE(32:3)	(FA 14:0/FA 18:3)	684.5	0.01	0.00
253	PE(32:4)	(FA 14:0/FA 18:4)	682.5	0.00	0.00
254	PE(32:5)	(FA 14:0/FA 18:5)	681	0.00	0.00
255	PE(34:0)	(FA 16:0/FA 18:0)	718.5	0.02	0.00
256	PE(34:2)	(FA 14:0/FA 20:2)	714	0.00	0.00
257	PE(34:3)	(FA 16:0/FA 18:3)	712	0.00	0.00
258	PE(34:4)	(FA 14:0/FA 20:4)	711	0.00	0.00
259	PE(34:4)	(FA 16:0/FA 18:4)	710	0.00	0.00
260	PE(34:5)	(FA 14:1/FA 20:4)	708.5	0.00	0.00
261	PE(36:3)	(FA 18:1/FA 18:2)	741	1.01	0.13
262	PE(36:4)	(FA 16:0/FA 20:4)	738.5	0.00	0.00
263	PE(36:5)	(FA 14:1/FA 22:4)	737	0.75	0.11
264	PE(36:5)	(FA 16:3/FA 20:2)	736.5	0.00	0.00
265	PE(38:1)	(FA 14:0/FA 24:1)	772.5	0.12	0.00
266	PE(38:3)	(FA 18:1/FA 20:2)	768.5	0.07	0.00
267	PE(38:4)	(FA 16:1/FA 22:3)	766	0.01	0.00
268	PE(38:5)	(FA 18:3/FA 20:2)	764.5	0.08	0.01
269	PE(38:6)	(FA 14:0/FA 24:6)	762.5	0.16	0.02
270	PE(40:1)	(FA 14:1/FA 26:0)	801	0.00	0.00
271	PE(40:6)	(FA 14:0/FA 26:6)	790.5	0.02	0.01
272	PE(40:8)	(FA 22:6/FA 18:2)	786.5	0.05	0.00
273	PE(42:2)	(FA 16:0/FA 26:2)	826.5	0.01	0.00
274	PE(42:3)	(FA 16:1/FA 26:2)	824.5	0.00	0.00
275	PE(42:4)	(FA 20:1/FA 22:3)	822.5	0.00	0.00
276	PE(44:7)	(FA 22:2/FA 22:5)	844.5	0.00	0.00
277	PE(46:5)	(FA 24:1/FA 22:4)	877	0.00	0.00
278	PE(48:10)	(FA 22:6/FA 26:4)	894	0.00	0.00
279	PE(48:3)	(FA 22:3/FA 26:0)	908.5	0.00	0.00
280	PE(48:5)	(FA 22:5/FA 26:0)	904.5	0.00	0.00
281	PE(48:6)	(FA 22:3/FA 26:3)	902.5	0.00	0.00
282	PG(30:0)	(FA 15:0/FA 15:0)	693.5	0.00	0.00

283	PG(30:2)	(FA 14:0/FA 16:2)	689.5	0.00	0.00
284	PG(30:3)	(FA 14:0/FA 16:3)	688	0.04	0.00
285	PG(30:5)	(FA 14:0/FA 16:5)	683.5	0.02	0.00
286	PG(32:2)	(FA 16:0/FA 16:2)	718	0.00	0.00
287	PG(32:4)	(FA 16:0/FA 16:4)	713.5	0.00	0.00
288	PG(34:5)	(FA 14:0/FA 20:5)	740	0.00	0.00
289	PG(36:4)	(FA 16:0/FA 20:4)	769	0.59	0.03
290	PG(38:5)	(FA 14:1/FA 24:4)	795	0.00	0.00
291	PG(38:6)	(FA 16:0/FA 22:6)	793.5	0.00	0.00
292	PG(38:7)	(FA 18:2/FA 20:5)	792	0.00	0.00
293	PG(42:0)	(FA 16:0/FA 26:0)	861.5	0.00	0.00
294	PG(42:3)	(FA 16:0/FA 26:3)	858	0.12	2.46
295	PG(44:12)	(FA 22:6/FA 22:6)	868	0.00	0.02
296	PG(48:8)	(FA 22:6/FA 26:2)	929	0.01	0.00
297	PI(30:0)	(FA 15:0/FA 15:0)	781.5	0.00	0.00
298	PI(36:0)	(FA 14:0/FA 22:0)	865.5	0.00	0.00
299	PI(36:7)	(FA 14:1/FA 22:6)	851	0.01	0.00
300	PI(38:1)	(FA 20:1/FA 18:0)	892	0.00	0.00
301	PI(38:2)	(FA 14:0/FA 24:2)	889.5	0.01	0.00
302	PI(38:3)	(FA 14:0/FA 24:3)	887	0.01	0.00
303	PI(38:4)	(FA 14:0/FA 24:4)	885.5	0.01	0.00
304	PI(38:5)	(FA 14:0/FA 24:5)	884	0.02	0.00
305	PI(38:5)	(FA 14:0/FA 24:5)	883	0.00	0.01
306	PI(38:6)	(FA 14:0/FA 24:6)	881.5	0.07	0.00
307	PI(40:1)	(FA 14:0/FA 26:1)	920	0.00	0.00
308	PI(40:2)	(FA 14:0/FA 26:2)	918	0.01	0.00
309	PI(40:5)	(FA 14:0/FA 26:5)	912	0.01	0.00
310	PI(40:5)	(FA 16:0/FA 24:5)	911	0.00	0.00
311	PI(40:6)	(FA 14:0/FA 26:6)	909.5	0.00	0.00
312	PI(42:0)	(FA 16:0/FA 26:0)	950	0.00	0.00
313	PI(42:1)	(FA 16:0/FA 26:1)	947.5	0.01	0.00
314	PI(42:6)	(FA 16:0/FA 26:6)	938	0.01	0.00
315	PI(42:6)	(FA 22:6/FA 20:0)	937	0.00	0.01
316	PI(44:0)	(FA 18:0/FA 26:0)	977.5	0.04	0.00
317	PI(44:10)	(FA 18:4/FA 26:6)	957	0.03	0.00
318	PI(44:11)	(FA 22:6/FA 22:5)	955	0.00	0.00

319	PI(46:10)	(FA 22:5/FA 24:5)	986	0.00	0.02
320	PI(46:10)	(FA 22:6/FA 24:4)	985	0.20	0.00
321	PI(46:11)	(FA 22:6/FA 24:5)	983.5	0.01	0.00
322	PI(46:6)	(FA 24:1/FA 22:5)	994	0.00	0.00
323	PI(46:9)	(FA 22:5/FA 24:4)	988	0.00	0.00
324	PS(26:0)	(FA 14:0/FA 12:0)	650	0.00	0.01
325	PS(28:0)	(FA 14:0/FA 14:0)	680	0.00	0.00
326	PS(30:0)	(FA 14:0/FA 16:0)	706.5	0.00	0.00
327	PS(32:0)	(FA 16:0/FA 16:0)	734.5	0.00	0.00
328	PS(36:10)	(FA 22:6/FA 14:4)	770.5	0.12	0.00
329	PS(40:0)	(FA 14:0/FA 26:0)	846	0.00	0.00
330	SQDG(26:0)	(FA 14:0/FA 12:0)	709.5	0.01	0.00
331	SQDG(26:1)	(FA 14:1/FA 12:0)	708	0.00	0.01
332	SQDG(30:1)	(FA 14:0/FA 16:1)	763.5	0.00	0.01
333	SQDG(30:2)	(FA 14:0/FA 16:2)	762	0.01	0.00
334	SQDG(32:2)	(FA 14:0/FA 18:2)	789.5	0.02	0.00
335	SQDG(34:2)	(FA 14:1/FA 20:1)	818	0.01	0.00
336	SQDG(34:2)	(FA 16:0/FA 18:2)	817.5	0.00	0.00
337	SQDG(34:2)	(FA 18:1/FA 16:1)	836.5	0.03	0.00
338	SQDG(34:3)	(FA 14:1/FA 20:2)	834	0.02	0.00
339	SQDG(36:0)	(FA 14:0/FA 22:0)	849	0.00	0.03
340	SQDG(36:0)	(FA 16:0/FA 20:0)	850	0.14	0.41
341	SQDG(36:2)	(FA 18:1/FA 18:1)	864	0.27	0.00
342	SQDG(38:0)	(FA 14:0/FA 24:0)	878	0.00	0.00
343	SQDG(38:1)	(FA 14:0/FA 24:1)	875.5	0.00	0.00
344	SQDG(38:2)	(FA 14:0/FA 24:2)	873.5	0.00	0.00
345	SQDG(38:5)	(FA 14:0/FA 24:5)	867.5	0.24	0.00
346	SQDG(40:0)	(FA 14:0/FA 26:0)	905.5	0.00	0.00
347	SQDG(40:1)	(FA 14:0/FA 26:1)	903	0.00	0.00
348	SQDG(40:2)	(FA 14:0/FA 26:2)	901.5	0.00	0.00
349	SQDG(40:4)	(FA 14:0/FA 26:4)	897.5	0.00	0.00
350	SQDG(40:5)	(FA 20:0/FA 20:5)	914	0.00	0.00
351	SQDG(42:2)	(FA 16:1/FA 26:1)	930	0.00	0.00
352	SQDG(42:3)	(FA 20:3/FA 22:0)	928	0.00	0.00
353	SQDG(44:1)	(FA 24:1/FA 20:0)	959.5	0.03	0.00
354	SQDG(44:3)	(FA 24:1/FA 20:2)	955.5	0.01	0.00

355	SQDG(44:4)	(FA 18:1/FA 26:3)	953.5	0.00	0.02
356	TAG(48:4)	(FA 18:1/FA 14:0/FA 16:3)	821.5	0.03	0.00
357	TAG(50:4)	(FA 18:1/FA 14:0/FA 18:3)	849.5	0.00	0.00
358	TAG(50:5)	(FA 18:1/FA 14:0/FA 18:4)	847.5	0.06	2.07
359	TAG(52:7)	(FA 18:1/FA 14:2/FA 16:1)	866.5	1.75	0.00
360	TAG(54:10)	(FA 18:1/FA 14:3/FA 22:6)	888.5	0.00	0.01
361	TAG(58:1)	(FA 18:1/FA 16:0/FA 24:0)	962.5	0.03	0.01
362	TAG(58:10)	(FA 18:1/FA 16:4/FA 24:5)	945	0.00	0.00
363	TAG(58:11)	(FA 18:1/FA 16:4/FA 24:6)	943	0.00	0.00
364	TAG(58:8)	(FA 18:1/FA 16:3/FA 24:4)	948.5	0.07	0.00
365	TAG(58:9)	(FA 18:1/FA 16:4/FA 24:4)	946.5	0.07	0.00

Table. 9. Lipidome of *E. huxleyi* CCMP 2090.Table representing most common lipid species detected (in half or more than half of the replicates) in both stationary and logarithmic phases

S.No	Lipid	FA1	FA2	Mol-wt
1	DAG(32:1)	FA 16:0	FA 16:1	565
2	DAG(36:4)	FA 22:4	FA 14:0	615.5
3	BLL(38:3)	FA 14:0	FA 24:3	819
4	DGCC(36:2)	FA 14:0	FA 22:2	705.5
5	DGTS(40:1)	FA 14:0	FA 26:1	896.5
6	DGTS(42:5)	FA 16:0	FA 26:5	916.5
7	DGTS(44:7)	FA 18:5	FA 26:2	940.5
8	DGDG(30:0)	FA 14:0	FA 16:0	863.5
9	DGDG(32:1)	FA 18:1	FA 14:0	949
10	DGDG(32:4)	FA 14:0	FA 18:4	944.5
11	DGDG(32:5)	FA 14:1	FA 18:4	882
12	DGDG(34:0)	FA 14:0	FA 20:0	980
13	DGDG(34:1)	FA 14:0	FA 20:1	978.5
14	DGDG(34:2)	FA 14:0	FA 20:2	975
15	DGDG(34:3)	FA 14:0	FA 20:3	913.5
17	DGDG(34:5)	FA 14:0	FA 20:5	969
18	DGDG(34:6)	FA 14:0	FA 20:6	907.5
19	DGDG(36:1)	FA 14:0	FA 22:1	945.5
22	DGDG(36:3)	FA 14:0	FA 22:3	942
23	DGDG(36:7)	FA 14:1	FA 22:6	933.5
24	DGDG(36:8)	FA 18:3	FA 18:5	991
25	DGDG(38:0)	FA 14:0	FA 24:0	975.5
26	DGDG(38:3)	FA 14:0	FA 24:3	969.5
27	DGDG(38:4)	FA 14:0	FA 24:4	967
28	DGDG(38:4)	FA 16:0	FA 22:4	968
29	DGDG(38:5)	FA 14:0	FA 24:5	966
30	DGDG(40:12)	FA 22:6	FA 18:6	979.5
31	MGDG(26:0)	FA 14:0	FA 12:0	705
32	MGDG(28:0)	FA 14:0	FA 14:0	733

33	MGDG(32:1)	FA 14:0	FA 18:1	787.5
34	MGDG(34:1)	FA 14:0	FA 20:1	816
35	MGDG(38:1)	FA 14:0	FA 24:1	811.5
37	MGDG(38:2)	FA 14:0	FA 24:2	870.5
38	PC(26:1)	FA 14:0	FA 12:1	646.5
39	PC(28:4)	FA 14:0	FA 14:4	668
40	PC(30:0)	FA 14:0	FA 16:0	704
41	PC(30:1)	FA 16:0	FA 14:1	702.5
42	PC(30:3)	FA 14:0	FA 16:3	698.5
43	PC(30:4)	FA 14:0	FA 16:4	697
44	PC(30:5)	FA 14:0	FA 16:5	694
45	PC(32:6)	FA 14:0	FA 18:6	720.5
46	PC(34:1)	FA 14:0	FA 20:1	758.5
47	PC(34:2)	FA 14:0	FA 20:2	756.5
48	PC(34:3)	FA 18:1	FA 16:2	755
49	PC(34:3)	FA 14:0	FA 20:3	754
50	PC(34:4)	FA 16:0	FA 18:4	752.5
51	PC(34:5)	FA 14:1	FA 20:4	750.5
52	PC(36:5)	FA 18:0	FA 18:5	778.5
53	PC(38:1)	FA 14:1	FA 24:0	814
55	PC(38:4)	FA 14:1	FA 24:3	808
56	PC(38:4)	FA 18:1	FA 20:3	809
57	PC(38:5)	FA 18:0	FA 20:5	806.5
58	PC(38:7)	FA 14:1	FA 24:6	803
59	PC(40:5)	FA 14:0	FA 26:5	834.5
60	PC(40:6)	FA 14:0	FA 26:6	832.5
61	PC(42:7)	FA 22:6	FA 20:1	859
62	PC(42:9)	FA 20:4	FA 22:5	855
63	PC(44:8)	FA 22:6	FA 22:2	884.5
64	PC(46:8)	FA 22:6	FA 24:2	912.5
65	PDPT(34:2)	FA 14:0	FA 20:2	776
66	PDPT(36:6)	FA 14:0	FA 22:6	796
67	PDPT(40:2)	FA 14:0	FA 26:2	860
68	PDPT(40:6)	FA 14:0	FA 26:6	852
69	PDPT(48:11)	FA 22:6	FA 26:5	954
70	PDPT(48:2)	FA 24:1	FA 24:1	972

71	PDPT(48:8)	FA 22:6	FA 26:2	960
72	PE(38:6)	FA 14:0	FA 24:6	762.5
73	PS(28:0)	FA 16:0	FA 12:0	678.5
74	PS(28:0)	FA 14:0	FA 14:0	679
75	SQDG(30:1)	FA 14:0	FA 16:1	763.5
76	SQDG(34:2)	FA 16:0	FA 18:2	817.5
77	SQDG(34:2)	FA 14:1	FA 20:1	818
78	SQDG(38:1)	FA 14:0	FA 24:1	875.5
79	SQDG(38:5)	FA 14:0	FA 24:5	867.5

Table. 10. Total fatty acyl groups detected directly by negative ionization mode and calculation of counter parts in lipidome of *E. huxleyi* CCMP 2090. Fatty acids with " * "(asterick) are those which are present in mammalian type DHA biosynthesis pathway as shown in Fig. 26b

S.No	Fatty acid	S.No	Fatty acid
1	FA 12:0	22	FA 20:3*
2	FA 12:1	23	FA 20:4*
3	FA 14:0	24	FA 20:5*
4	FA 14:1	25	FA 20:6
5	FA 14:4	26	FA 22:1
6	FA 16:0	27	FA 22:2
7	FA 16:1	28	FA 22:3
8	FA 16:2	29	FA 22:4*
9	FA 16:3	30	FA 22:5*
10	FA 16:4	31	FA 22:6*
11	FA 16:5	32	FA 24:0
12	FA 18:0	33	FA 24:1
13	FA 18:1*	34	FA 24:2
14	FA 18:2*	35	FA 24:3
15	FA 18:3*	36	FA 24:4
16	FA 18:4	37	FA 24:5*
17	FA 18:5	38	FA 24:6*
18	FA 18:6	39	FA 26:1
19	FA 20:0	40	FA 26:2
20	FA 20:1	41	FA 26:5
21	FA 20:2	42	FA 26:6

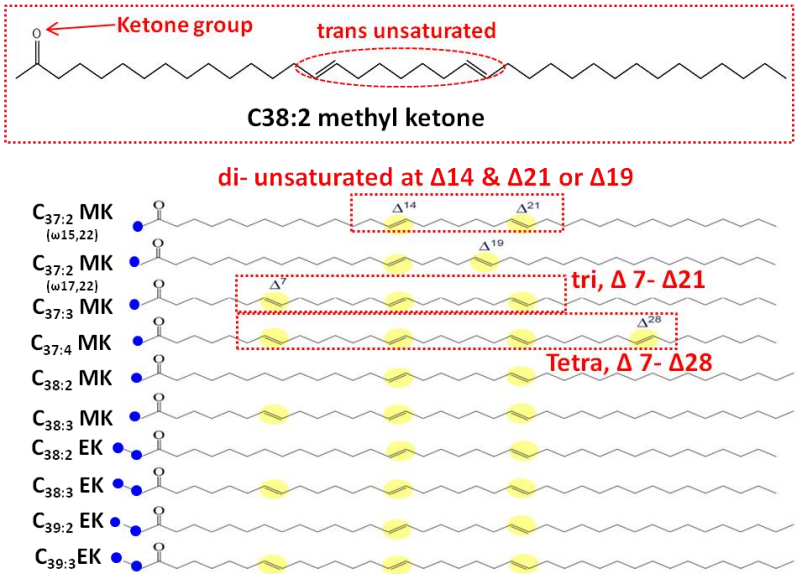

Fig. 1. Molecular structures of alkenone species. Image credits: Laboratory of plant physiology and metabolism. Purple spots indicate methyl and ethyl groups. Yellow spots trans indicated unsaturation.

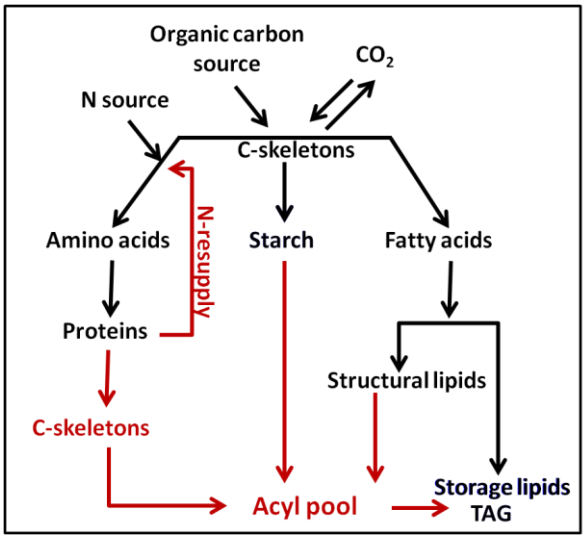

Fig. 2. General C-allocation mechanism under normal N-sufficient (+N) and N-limitation (–N) conditions. Arrows indicate direction of carbon flow in +N (black arrows) and –N (red arrows). Model credits Gonclaves et al. 2016.

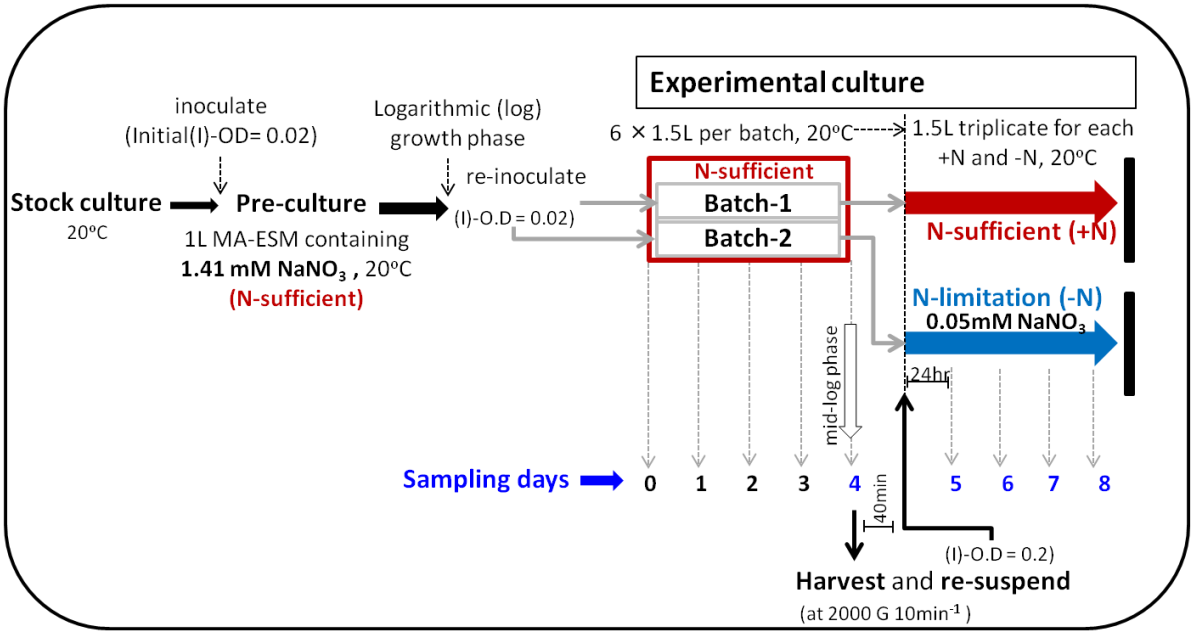

Fig. 3. Culture conditions and experimental procedure to study carbon allocation mechanism under N-sufficent (+N) and N-limitation (-N) conditions.

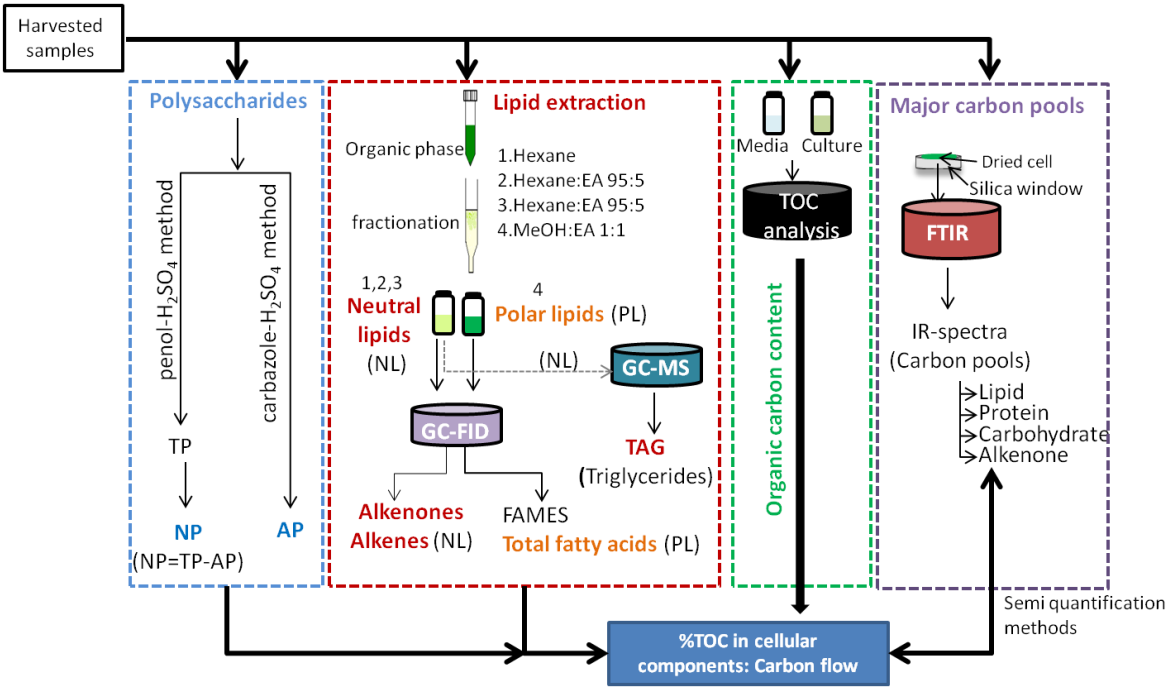

Fig.4. An overview of sample analysis procedure. Samples from day 4 to 8 are used to anlyze polysaccharides, lipids, total organic carbon and macromolecular carbon absorption spectra(FTIR). Here NP is neutral polysaccharides, AP is acid polysaccharides, FAMES is fatty acid methyl esters, GC-FID/MS is gas chromatography- flame ionization detection/ Mass spectrometry, TOC is total organic carbon and FTIR is fourrier transformed infrared spectroscopy.

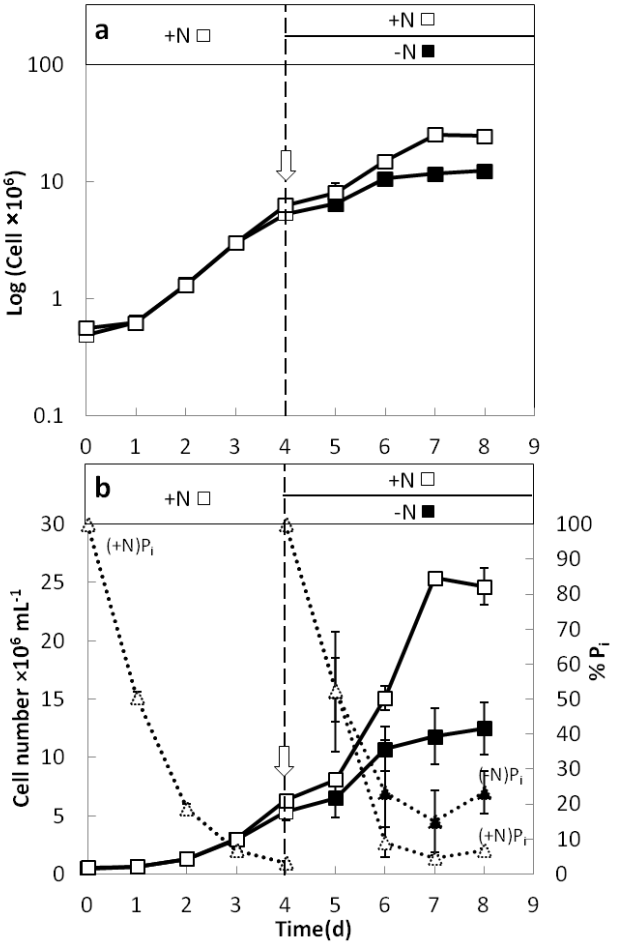

Fig. 5. Changes in growth and inorganic phosphate (Pi) concentration in the medium during the culture of the alkenone-producing haptophyte *E. huxleyi* CCMP2090 under nitrogen- sufficient (+N) and -limiting (–N) conditions. The cells were grown under +N conditions for first 4 days (stage-I), and then cells harvested by centrifugation were transferred to either +N or –N medium (stage-II), as described in Materials and Methods. **a,** growth curve with logarithmic scale on the y-axis. **b,** growth curve with linear scale on the y-axis and Pi concentration in the medium expressed as a relative value (% Pi). Open and closed symbols, data from experiments performed under +N and –N conditions, respectively. The initial Pi concentration (100% Pi) was 30 μM. Open arrows, the time when cells were transferred to either +N or –N conditions. Error bars indicate standard deviations from triplicate experiments (n=3).

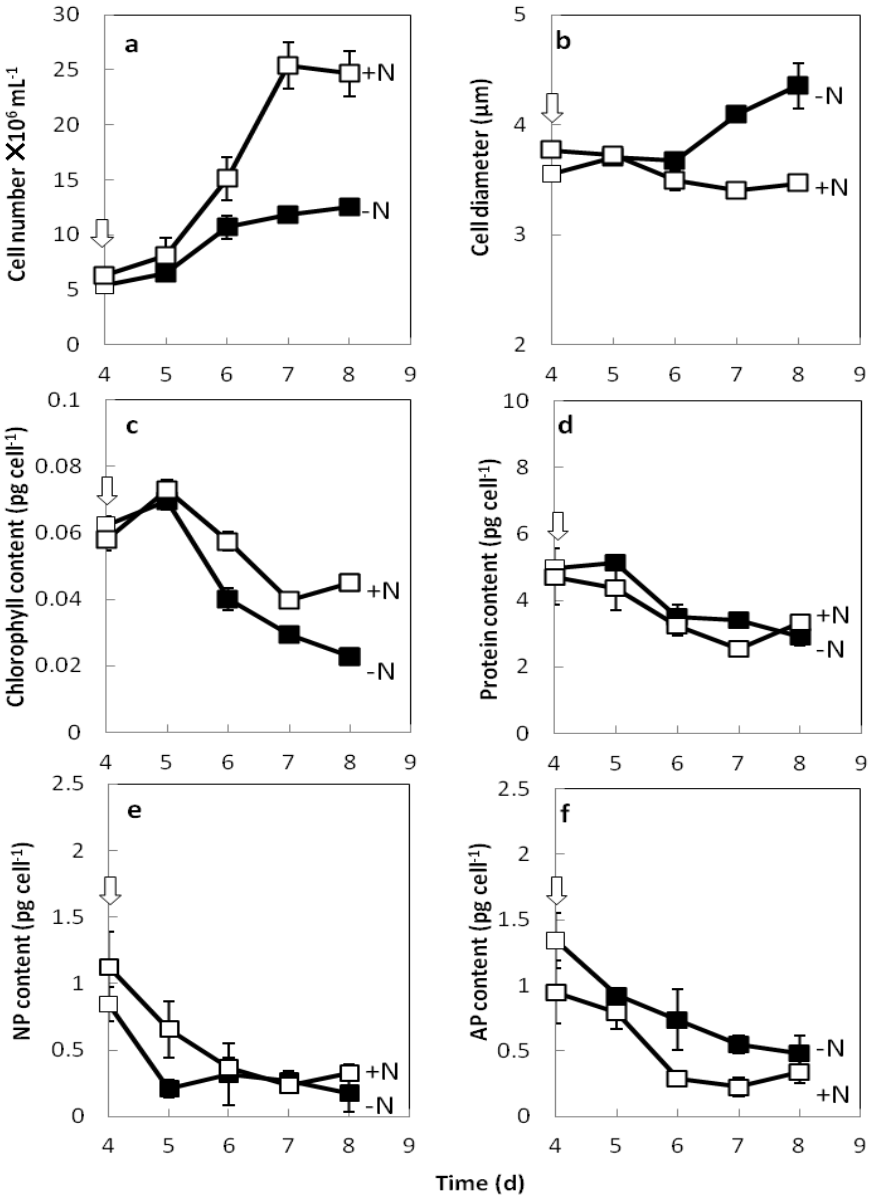

Fig. 6. Changes in the parameters of cell growth and the contents of cellular components (pg cell^{-1}) after the transfer of 4-d-grown +N cells to either +N (open circles) or –N conditions (closed circles) in *E. huxleyi* CCMP2090. **a,** cell number. **b,** cell size. **c-f,** the contents (pg cell^{-1}) of cellular components such as chlorophylls **(c)**, proteins **(d)**, neutral polysaccharides (NP) **(e)** and acid polysaccharides (AP) **(f)**. Refer Fig. 5 for symbols, error bars and the whole experiments during 8 days.

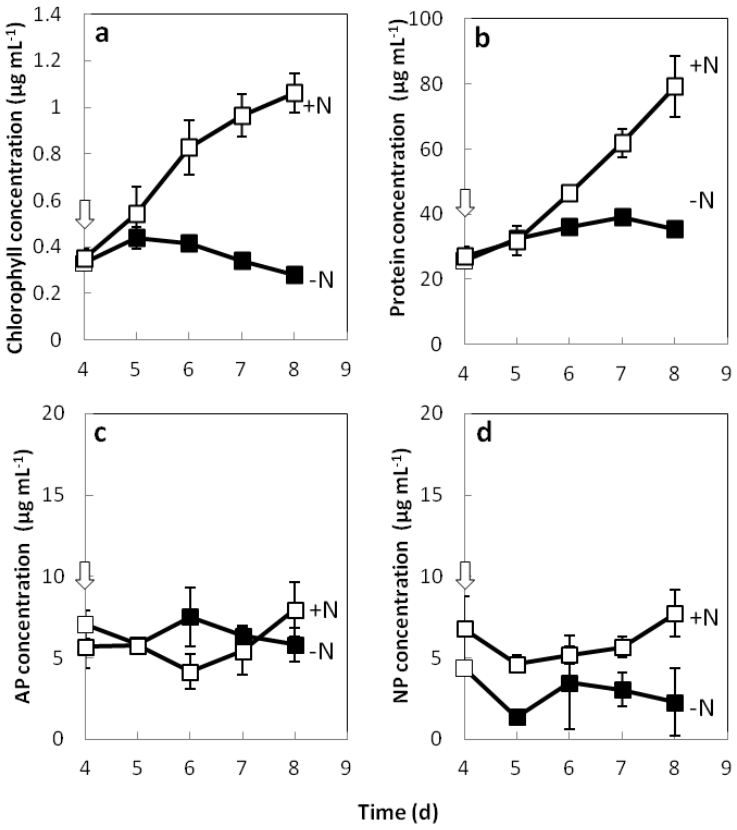

Fig. 7. Changes in the amounts of cellular components (μg (mL culture)$^{-1}$) after the transfer of 4-d-grown +N cells to either +N (open circles) or –N conditions (closed circles) in *E. huxleyi* CCMP2090. **a,** chlorophylls. **b,** proteins. **c,** neutral polysaccharides (NP). **d,** acid polysaccharides (AP). Refer Fig. 5 for symbols, error bars and the whole experiments during 8 days.

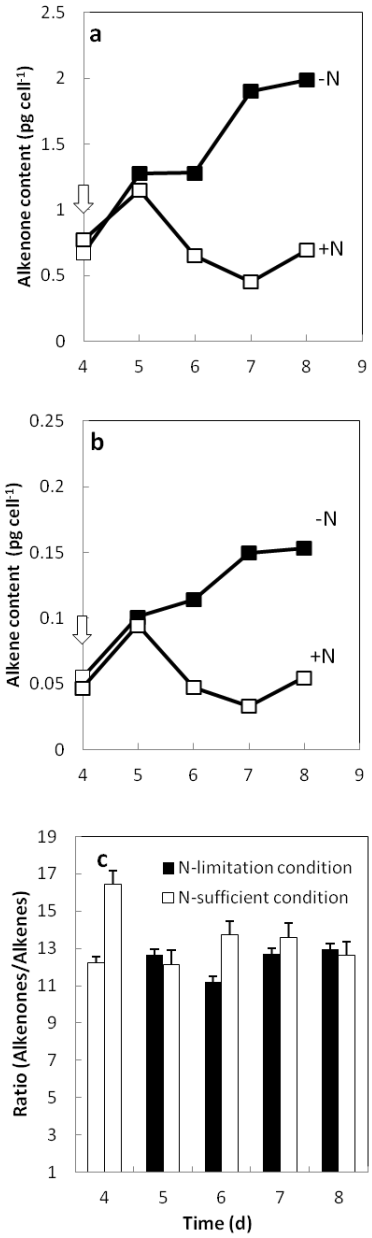

Fig. 8. Changes in neutral lipid contents after the transfer of 4-d-grown +N cells to either +N (open circles) or –N conditions (closed circles) in *E. huxleyi* CCMP2090. **a** and **b,** the contents of alkenones and alkenes, respectively. **c,** the relative accumulation patterns between alkenones and alkenes. Refer Fig. 5 for symbols, error bars and the whole experiments during 8 days.

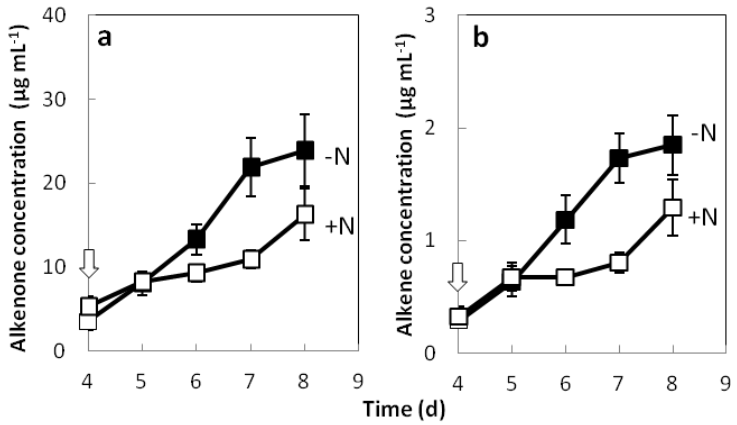

Fig. 9. Changes in the amounts of **a,** alkenones and **b**, alkenes (µg (mL culture)$^{-1}$) after the transfer of 4-d-grown +N cells to either +N (open circles) or –N conditions (closed circles) in *E. huxleyi* CCMP2090. Refer Fig. 5 for symbols, error bars and the whole experiments during 8 days.

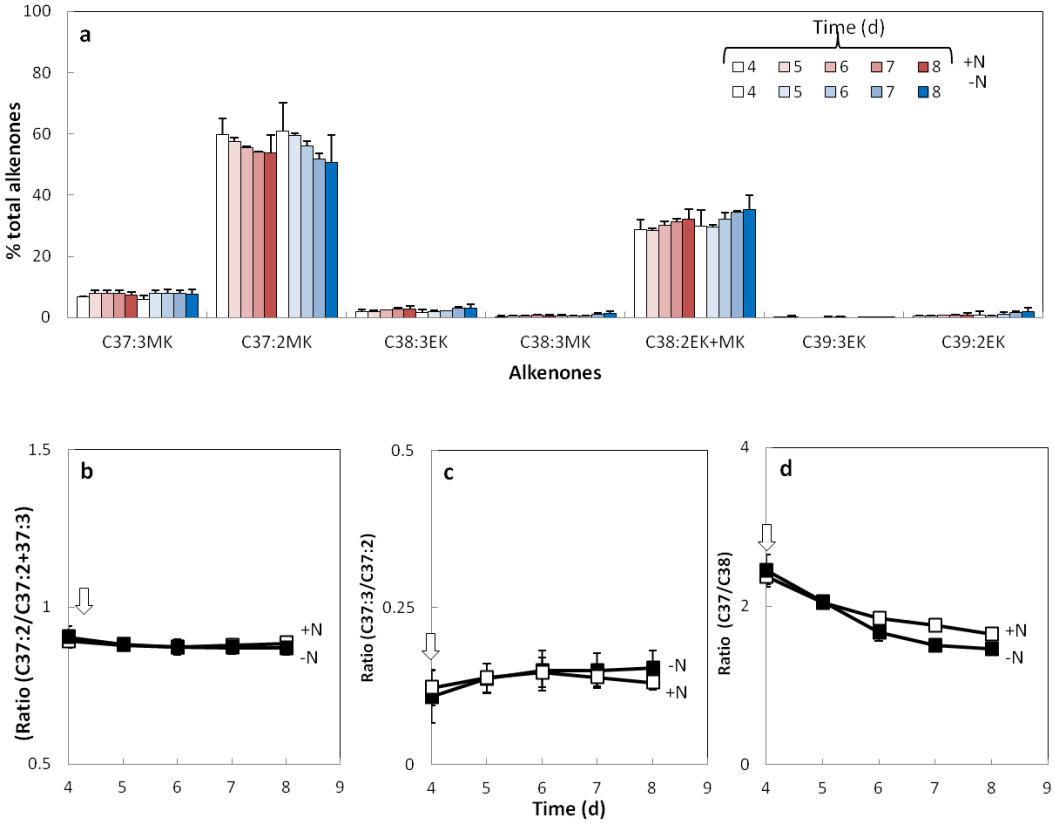

Fig. 10. Changes in alkenone components with unsaturation degree and carbon chain length during culture of *E. huxleyi* under +N (open squares) and –N conditions (closed squares). **a,** the compositions of various alkenone molecules expressed as % of total alkenone contents. **b,** the alkenone unsaturation index ($U^{K'}_{37}$) as a parameter of unsaturation degree. **c,** the alkenone unsaturation degree expressed as the ratio of $C_{37:3}/C_{37:2}$ alkenones. **d,** the carbon chain length expressed as the ratio of C_{37}/C_{38} alkenones. Red and blue bars indicate +N and –N conditions, respectively, in Fig. 10a. Refer Fig. 5 for symbols, error bars and the whole experiments during 8 days.

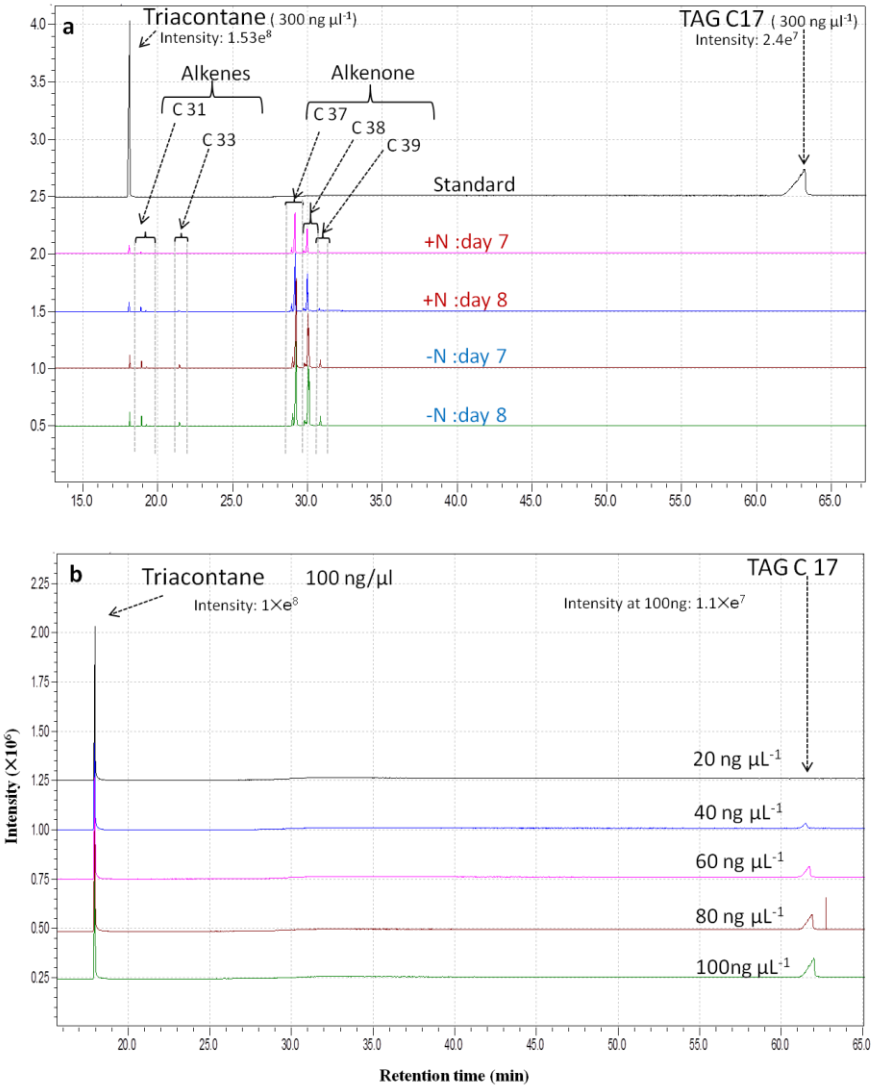

Fig. 11. Analysis for presence of TAGs in neutral lipid fractions. **a,** GC-MS analysis profiles of neutral lipid fractions from cells on various days. **b,** Calibration of TAG detection by GC-MS which was performed to confirm the sensitivity of GC-MS to TAG. Internal standards were a series of various concentrations of TAG C17 ranging from 20 ng μL^{-1} to 100 ng μL^{-1} and triacontane at 100 ng μL^{-1}.

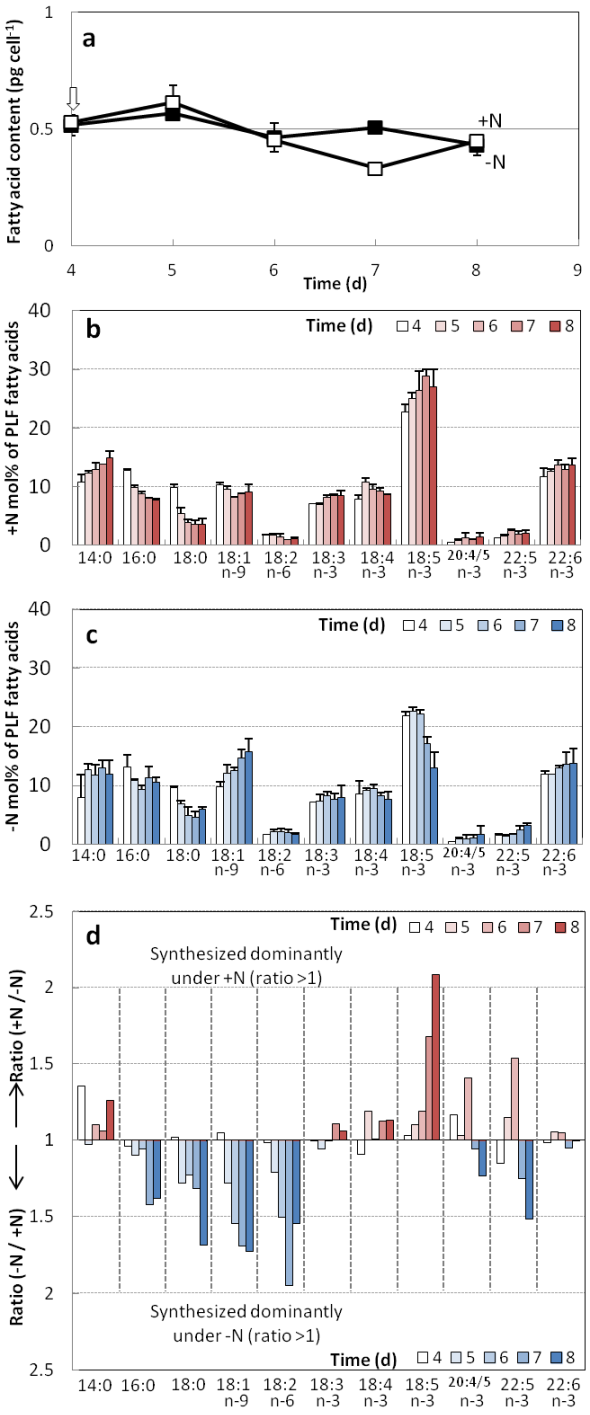

Fig. 12. Changes in FA contents after the transfer of 4-d-grown +N cells to either +N (open circles) or –N conditions (closed circles) in *E. huxleyi* CCMP2090. **a,** TFA contents in the cells. **b** and **c,** FA profiles of polar lipid fraction (PLF) in +N and –N cells, respectively. **d,** the ratios of mol% of FAs between +N cells and –N cells. Upper

(red) and lower (blue) columns, FAs which are dominant in +N cells and –N cells, respectively. Refer Fig. 5 for symbols, error bars and the whole experiments during 8 days.

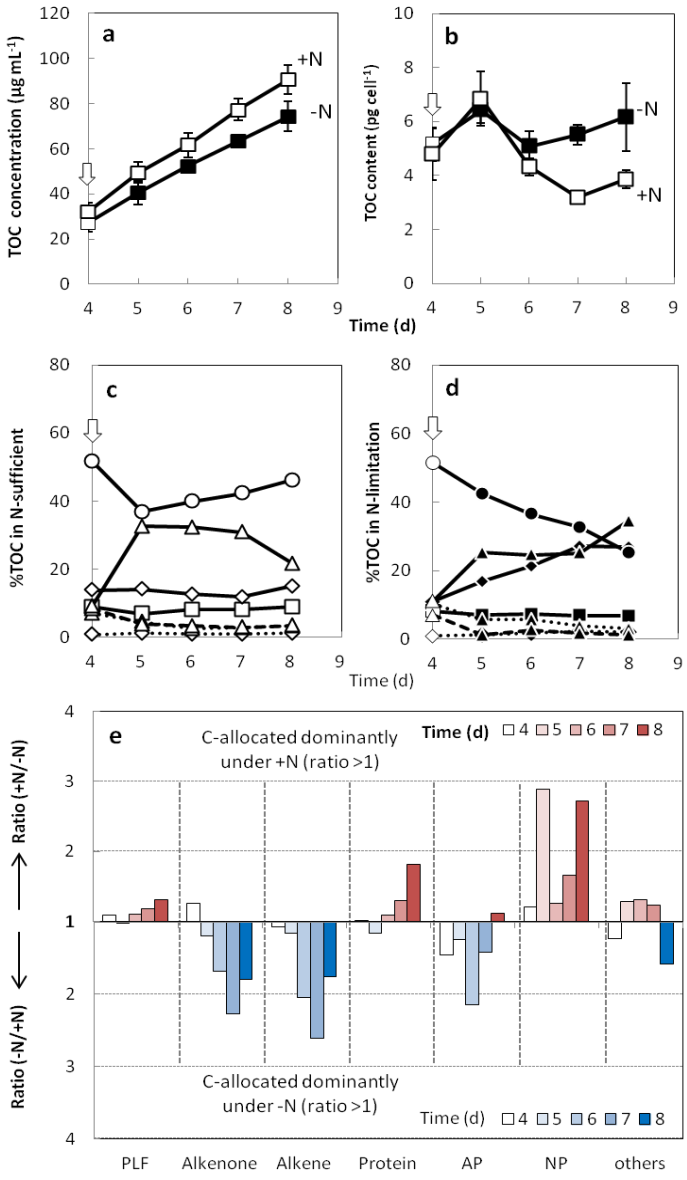

Fig. 13. Changes in the carbon allocation into various cellular components after the transfer of 4-d-grown +N cells to either +N (open symbols) or –N conditions (closed symbols) in *E. huxleyi* CCMP2090. **a** and **b,** TOC in +N and –N cells expressed per mL culture and cell, respectively. **c** and **d,** Carbon allocation into various components (%) in +N and –N cells, respectively. Symbols; cell components such as proteins (circles), alkenones (diamonds), FAs as PLF (squares), acid polysaccharides (triangles with dashed line), neutral polysaccharides (diamonds with dotted line) and others including unidentified low molecular compounds (triangles). **e,** the ratios of carbons allocated into various components between +N cells and –N cells. Upper (red) and lower (blue)

columns, the ratio of C-allocation in various compounds in which are dominant in +N cells and –N cells, respectively. Refer Fig. 1 for open arrow, error bars and the whole experiments during 8 days.

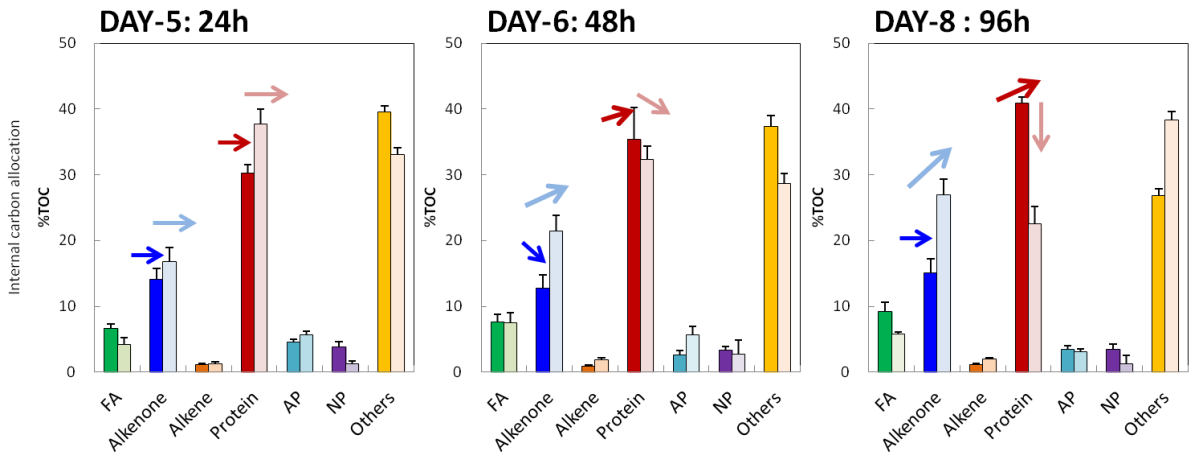

Fig. 14. Model of carbon distribution into various molecular pools during **day 5**, **day 6** and **day 8** of *E. huxleyi* CCMP 2090 grown under +N (dark colored bars) and –N (light colored bars). Each color is specific for each metabolite. Arrows represent pattern of carbon flow under +N and –N condition. Colored arrows represent direction of carbon flow for that specific component relative to the adjacent colored bars.

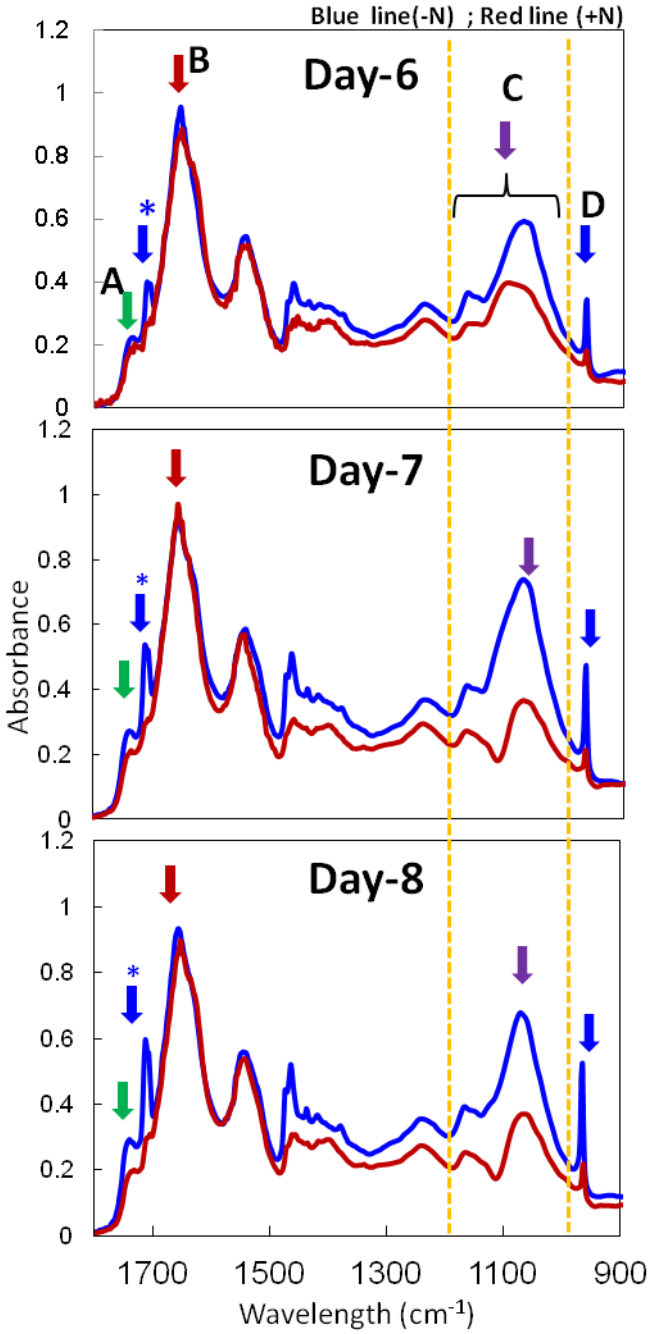

Fig. 15. Carbon bond excitation absorption spectra obtained using Fourier transformed infra red spectroscopy (FTIR). Data showing FTIR spectra of samples from **day 6, 7, 8** of N-sufficient (+N) and N-limitation (–N) conditions. Spectra is average of triplicate data.

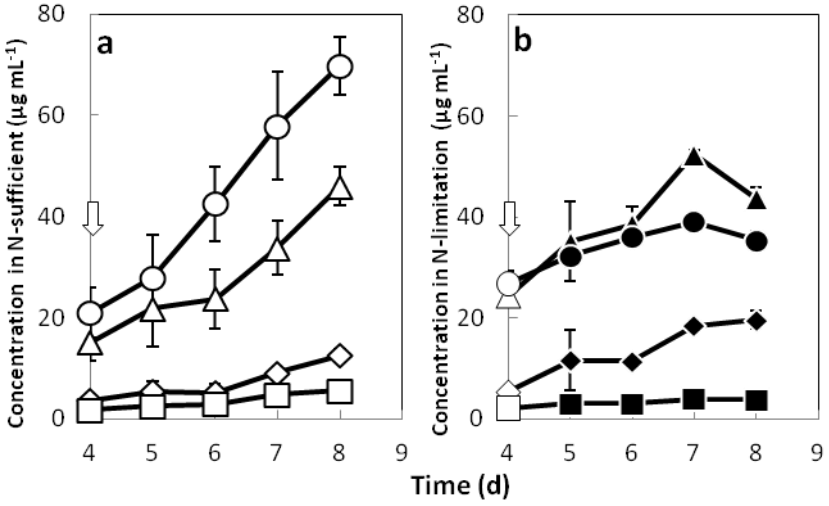

Fig. 16. Changes in major macro molecules from FTIR semi-quantification data. **a,** N-sufficient (open symbols, +N) and **b,** N-limitation (closed symbols, -N) conditions. Here Circles, triangles, diamonds and squares indicate Protein, other unknown compounds including carbohydrates, Alkenone, and Fatty acid (PLF), respectively. Refer Fig. 5 for open arrow, error bars and the whole experiments during 8 days.

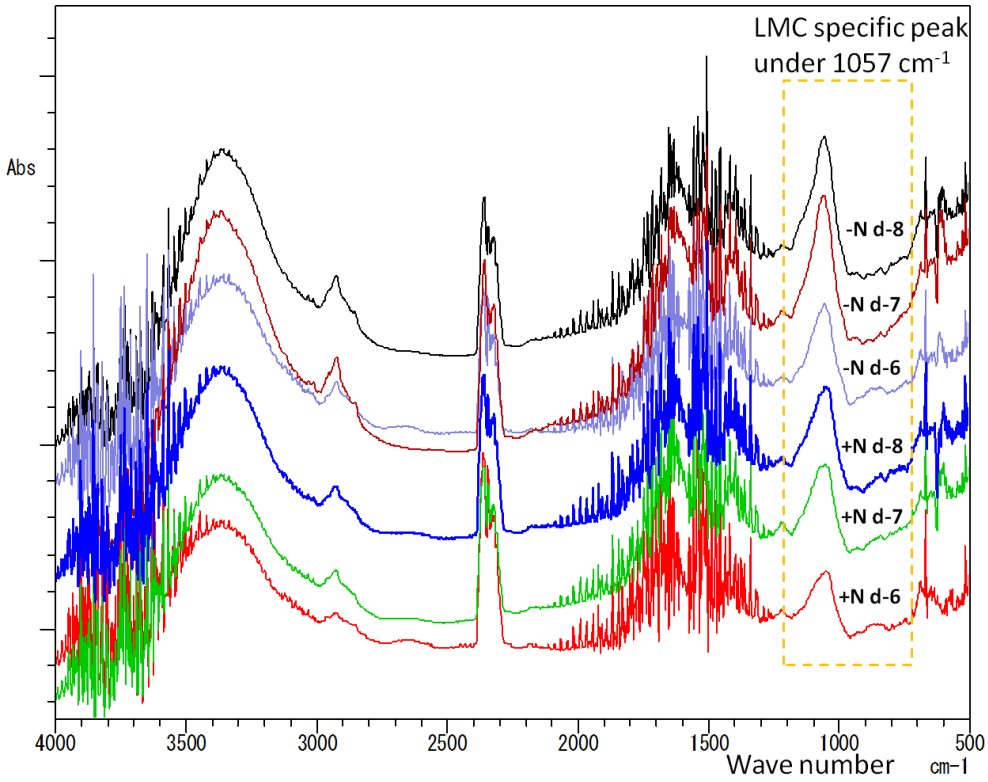

Fig. 17. FTIR spectra of MeOH-Dw fraction of samples from day(d) 6, 7, 8 of N-sufficient (+N) and N-limitation (–N) conditions. Peak under wave number 1057cm^{-1} represent absorption of infra red spectra by C-C, C-O-C bond vibrations of LMCs like mannitol and other sugars.

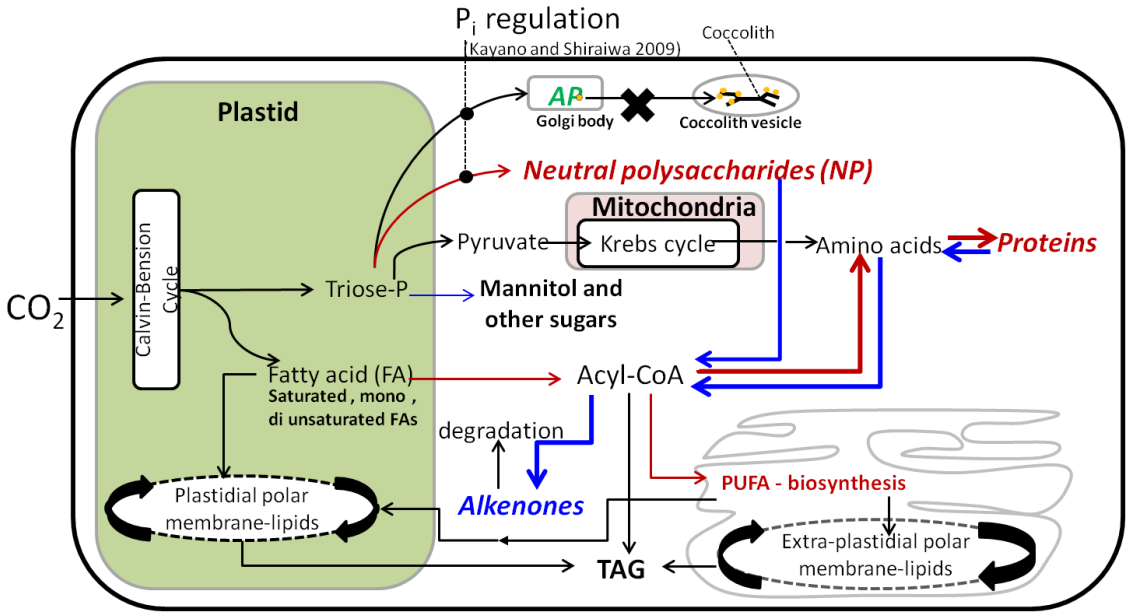

Fig. 18. A model representing the direction of C-allocation into various cell components in *E. huxleyi* CCMP 2090 under +N and –N conditions. The direction of C flow is represented by arrows. Metabolic processes and metabolites in black, red and blue represent reactions which are neutral, suppressive (stimulatory) and stimulatory (suppressive) under –N (+N) conditions, respectively. The thickness of arrow represents strength of the reaction. Metabolites in italics represent components quantified experimentally in this study. Metabolites in green represent no regulation by N-availability. Cross mark represents inhibitory. Symbols: AP, acid polysaccharides; PUFA, polyunsaturated FAs.

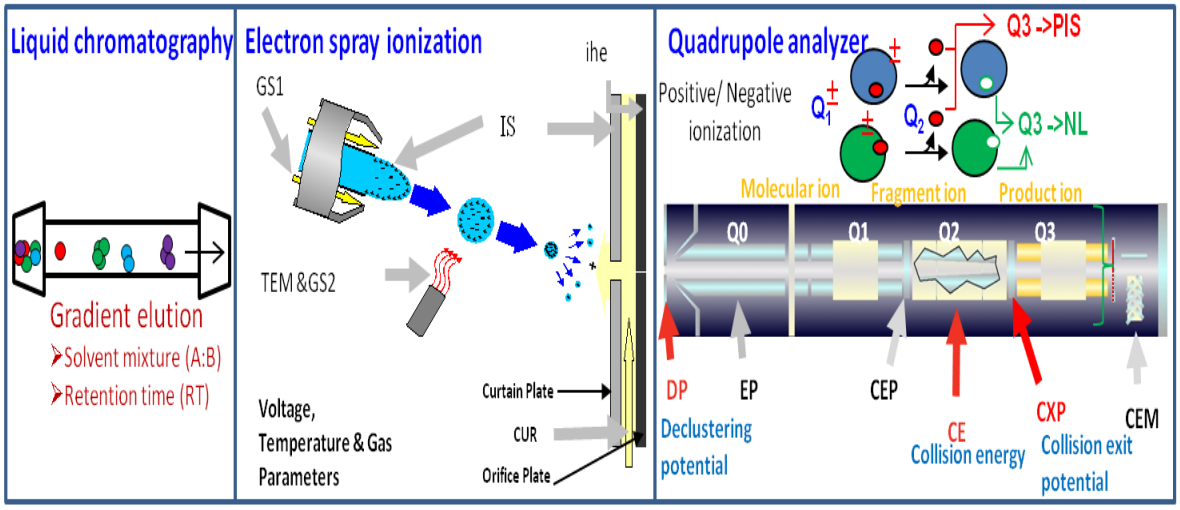

Fig. 19. Schematic cartoon of liquid chromatography-electronspray ionization mass spectrometer (LC-ESI/MS) setting for lipidomic analysis. First indicate indicates lipid separation (colored dots) using a HPLC column based on gradient elution methods. Second box indicate formation fine droplets from liquid sample (principle of ESI) at the capillary ending of sample injector. Third box indicate the quadrapole (Q) chamber (tandem masspectrometer, QTRAP) through which the ionized molecular ion passes and undergoes fragmentation at various points. Q1 is quadrapole 1 where molecular ions are selected; Q2 is collision chamber where fragmentation of molecular ion occurs and at Q3 selection of specific lipid fragments occurs. Here DP, EP, CE, CEP, CXP and CEM are various electric potentials. DP (declustering potential), CE (collision energy) and CXP (collision exit potential) are optimized for method establishment according to ABSciex QTRAP 5500 procedure.

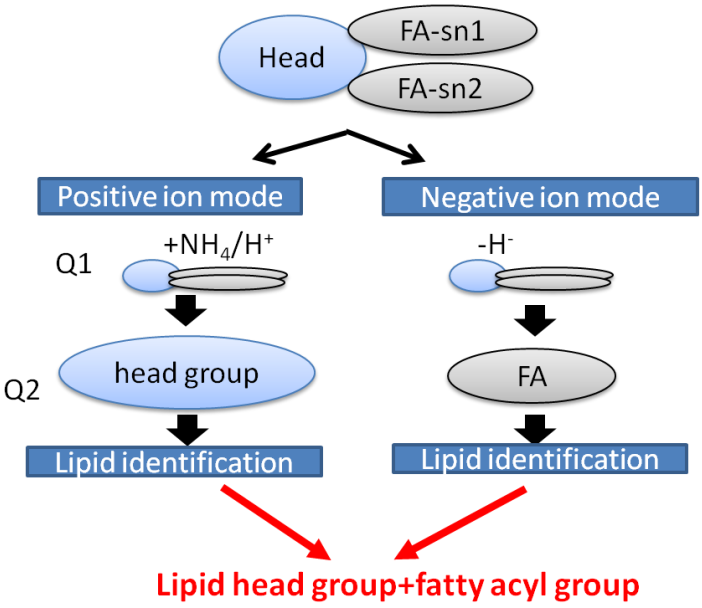

Fig. 20. Flow for lipid detection and qualitative analysis based on head group and fatty acyl fragmentation in positive and negative ion mode. $+NH_4$, H+ and H- are adductions formed during ionization of a lipid molecule. Q1 and Q2 are quadruples at which molecular ion and fragment ions and their masses (m/z) are detected.

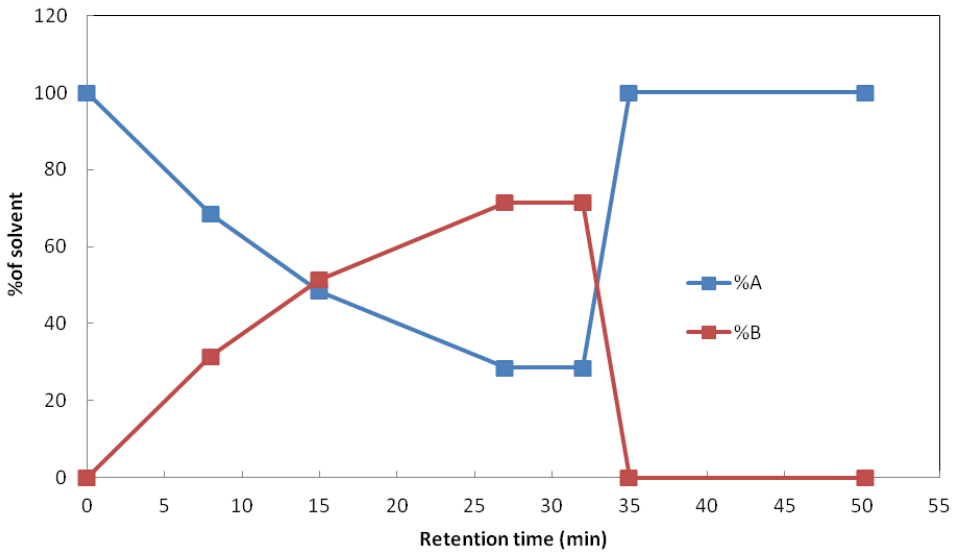

Fig. 21. High performance liquid chromatography (HPLC) gradient elution graph. A and B represent HPLC elution buffers. The graph represents combination of elution buffers at different time points to enable lipid class separation. A= 600:400:1.0:0.4 n-hexane:isopropanol:formic acid: 25 % aqueous NH_4; eluent B= 880:120:1.0:0.4 isopropanol:water:formic acid: 25 % aqueous NH_4.

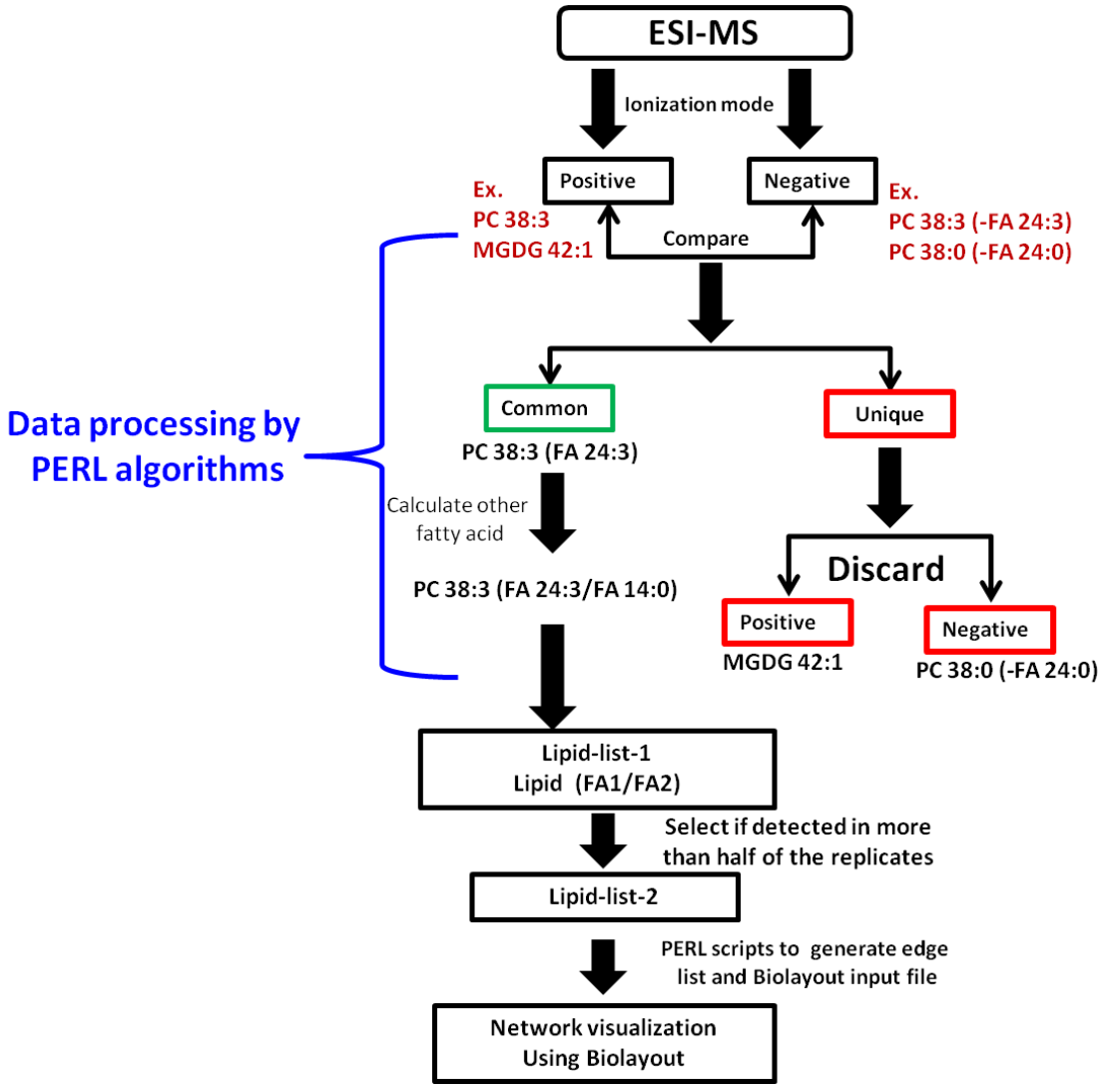

Fig. 22. Flow chart of lipidome data processing using computational methods.

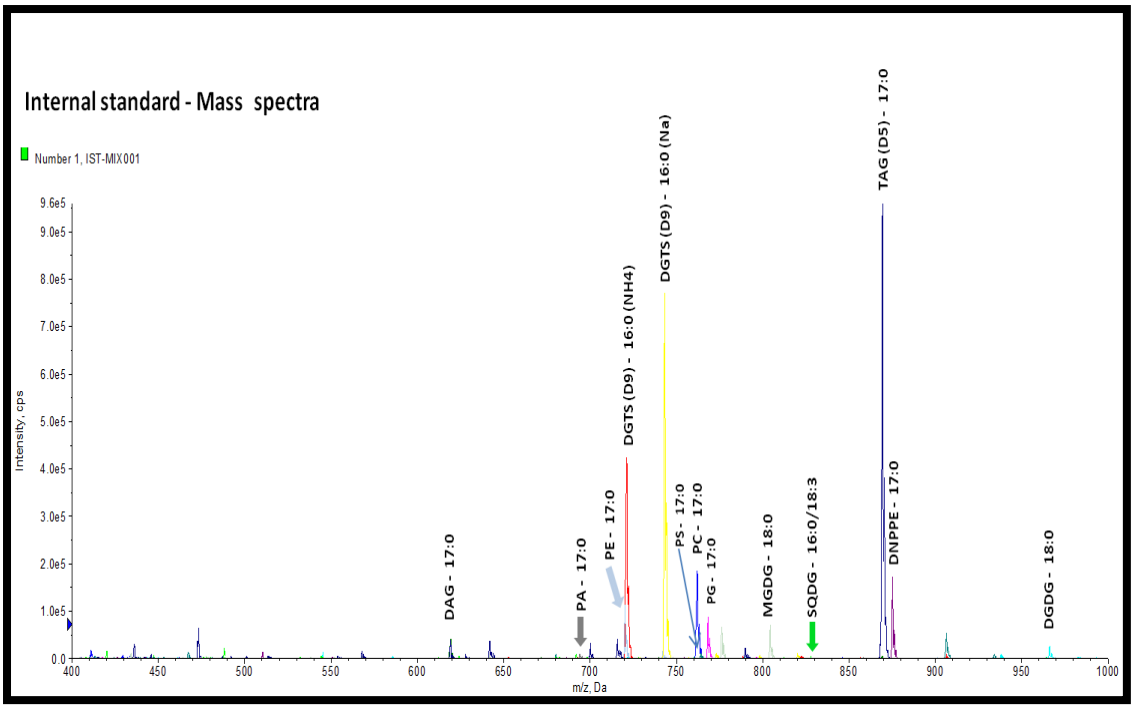

Fig. 23. Mass spectra of 13-IST mixture, showing 12 ISTs detected in positive ionization mode. Refer table 4 and table 6 for lipid abbreviations.

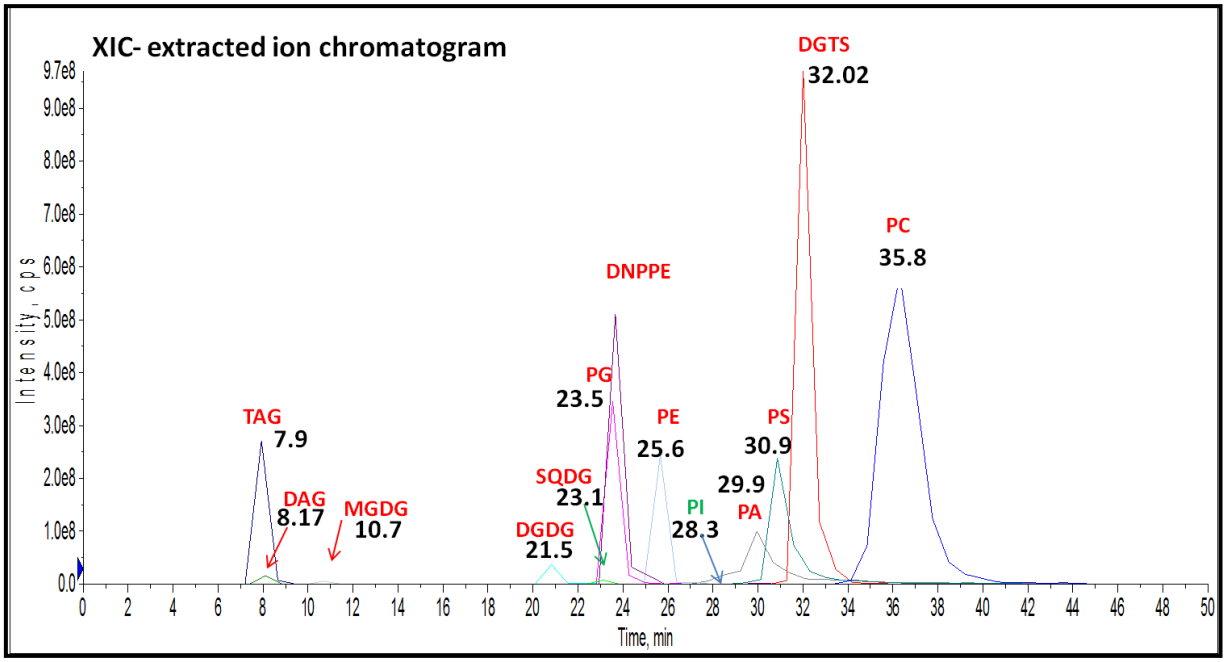

Fig. 24. Extracted ion spectra (XIC) of 13-IST mixture in positive ionization mode. Each IST is at 100ng concentration. Spectra represent retention time at which each lipid class specific IST is detected. PI is detected in negative ion mode and its retention time is mapped on this spectra. Differences in peak heights indicate difference in sensitivity of detection for each lipid class. Refer table 4 and table 6 for lipid abbreviations.

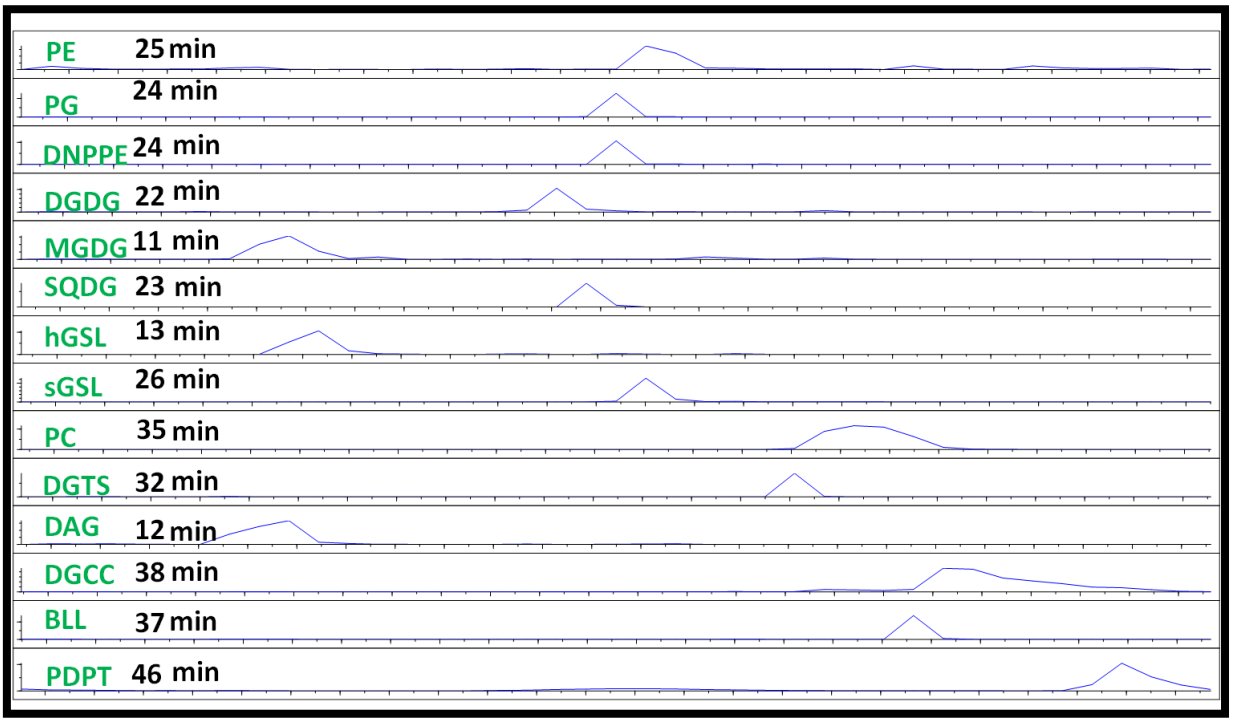

Fig. 25. Extracted ion chromatogram (XIC) of *E. huxleyi* CCMP 2090 lipidome. XIC of each lipid class show in separate panes to clearly represent the separation various lipid classes present in sample. Refer table 4 and table 6 for lipid abbreviations.

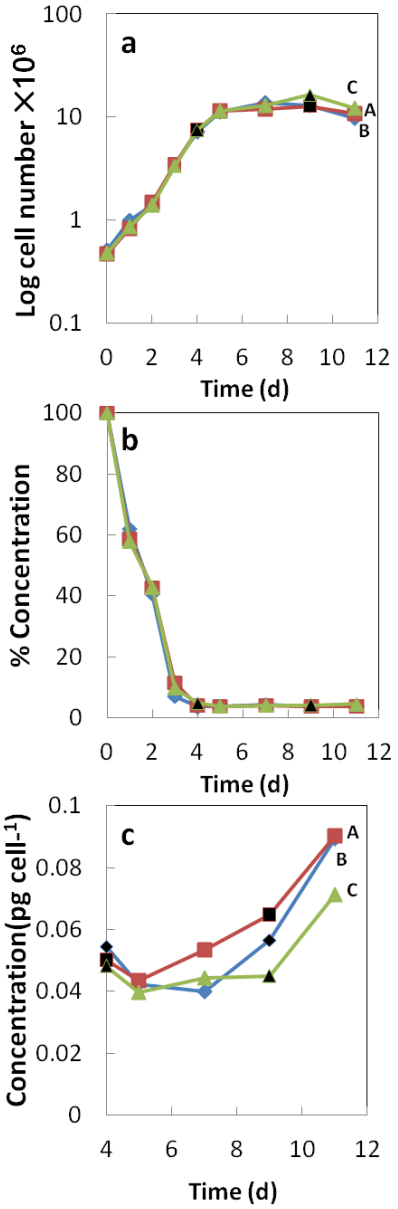

Fig. 26. Changes in growth and cellular components of *E. huxleyi* CCMP 2090 during logarithmic and stationary phase. **a,** Cell growth. **b,** Inorganic phosphate concentration. **c,** Chlorophyll concentration. Squares, diamonds and triangles indicate replicate experiments A, B and C. Dark filled symbols indicate point at which sampling was donne for lipidome analysis.

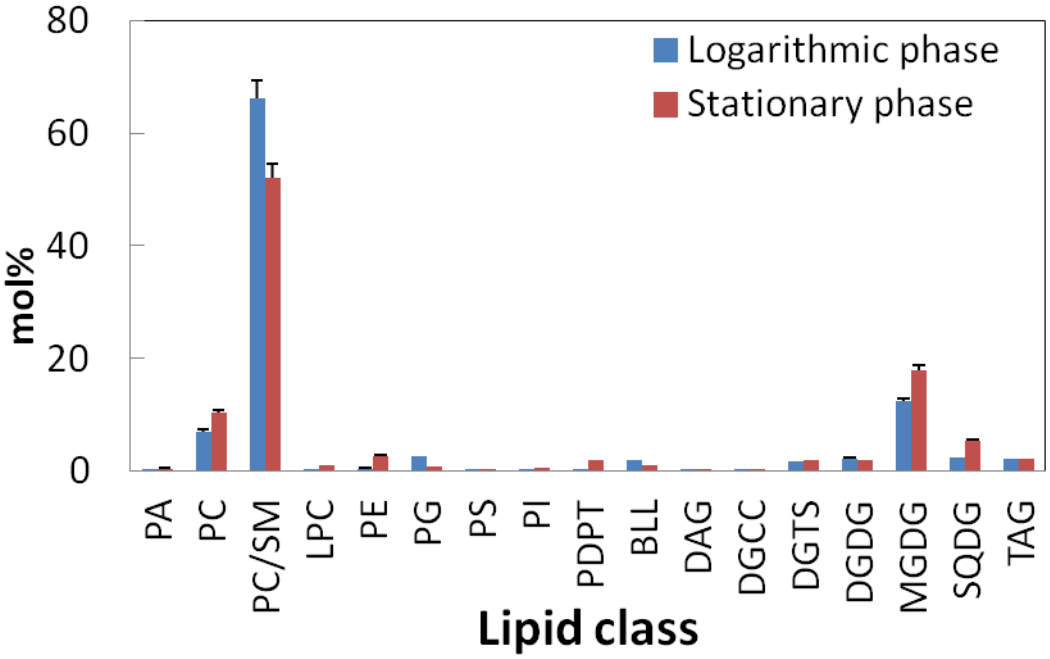

Fig. 27. Lipidome composition of *E. huxleyi* CCMP 2090 during logarithmic and stationary phases. Errora bars indicate standard deviation of triplicate experiments. PC/SM or PC represent phosphatidyl choline (PC) and sphingomylein (SM). Refer table 4 and table 6 for lipid class abbreviations.

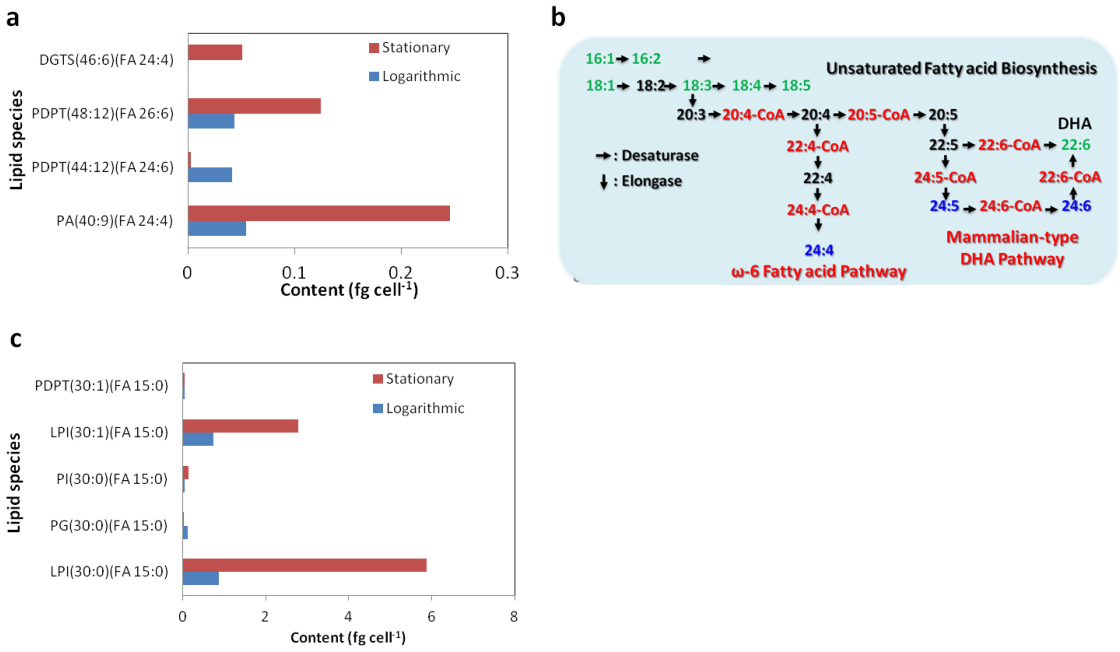

Fig. 28. Additional fatty acid pathways detected from lipidome of *E. huxleyi* CCMP 2090 during log and stationary growth phase. **a,** very long chain fatty acids detected in various lipid species. **b,** possible mammalian type DHA biosynthesis pathway in *E. huxleyi*. **c,** odd chain fatty acid 15:0 detected in various lipid species.

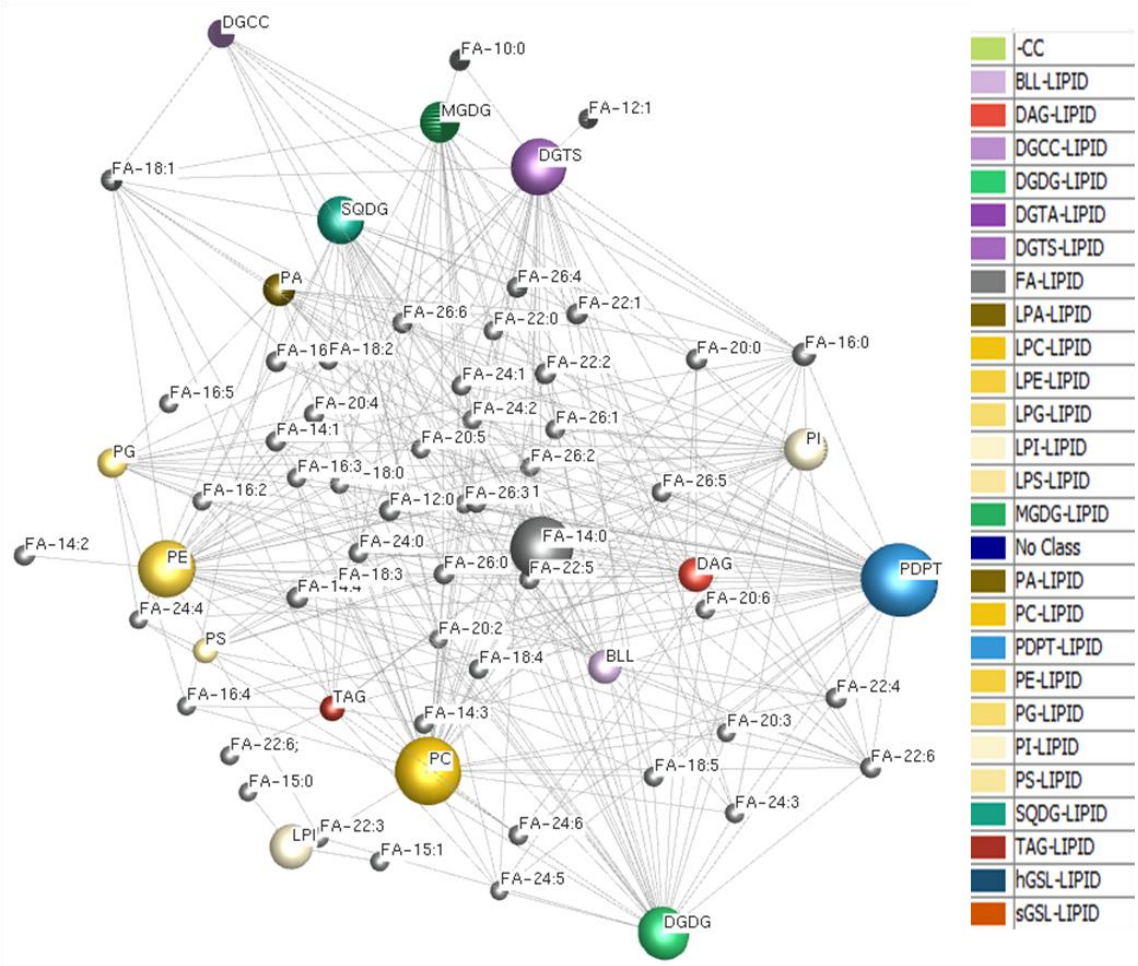

Fig. 29. Network view of lipids from *E. huxleyi* CCMP 2090. Data represent the 365 lipid species that were sorted using computational methods. Image generated using Biolayout 3D V 3.0 software. Here the colored spheres (nodes) represent different lipid classes. Grey spheres represent fatty acyl groups (FA). The lines (edges) connecting a lipid class and a FA group indicate lipid species have that FA. FA with more edges indicates its distribution in different lipid classes. This distribution is also represented by increased size of nodes (nodes with more edges have larger size).

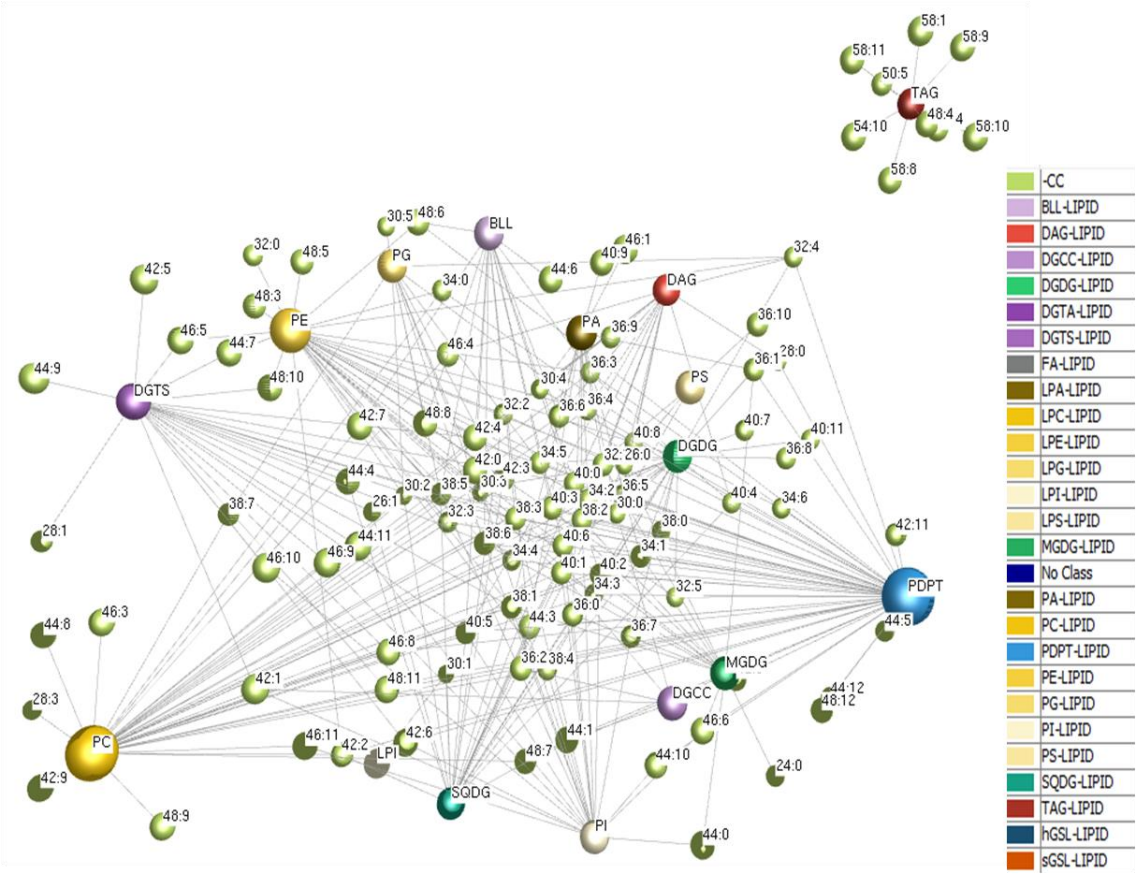

Fig. 30. Network view of lipids from *E. huxleyi* CCMP 2090. Data represent the 365 lipid species that were sorted using computational methods. Image generated using Biolayout 3D V 3.0 software. Here the colored spheres (nodes) represent different lipid classes and light green spheres (CC) represent total carbon number and double bonds associated to each lipid class. The lines (edges) connecting a lipid class and the carbons (CC) indicate detected lipid species with that carbon composition. CC with more edges indicates different lipid classes with similar CC composition. This distribution is also represented by increased size of lipid class nodes (nodes with more edges have larger size).

Appendix for PERL scripts

Lipid-compare.pl

- ✧ This program compares two lipidome data files (output tables from lipid view software
- ✧ Files can be positive (POS) or negative (NEG), this coded can compare POS-POS or NEG-NEG or NEG-POS data.
- ✧ This code generates 3 output files, 1. Unique lipids in each data file; 2. common lipids between two data files
- ✧ The FA list file should have fatty acid names in this way "10:0" just carbon-number: double-bonds (no need to say FA)
- ✧ From the COMMON data output-file select lipid names column for EXCEL processing "ex. PC 26:1 (FA 14:0/12:1)”
- ✧ In EXCEL make lipid with one FA like " PC 26:1 (FA 14:0) and PC 26:1 (FA 14:0) in each line. Save in text file, for Pre-BioLayout.pl

Pre-BioLayout.pl

- ✧ This program is pre processing script for Biolayout-input-file program
- ✧ The input file for this program consist lipid names in this way " PC 26:1 (FA 14:0) " in each line
- ✧ The input fatty acid list file should be consist fatty acid names in this way " FA(10:0) " in each line

- ✧ The output data files LC-INP and LF-INP are input files for two kinds of Biolayout-input-file programs
- ✧ LC-INP consist lipid class name, total carbons and total double bonds. This is input for BioLayout-CC.pl
- ✧ LF-INP consist Lipid class name, Fatty acid. This is input for BioLayout-Lipid-FA
- ✧ Redundant data is automatically removed, manual check can be done again

BioLayout-Lipid-FA.pl

- ✧ This program generates input file for BioLayout tool
- ✧ Here EDGE list for network is created using the lipid class names and fatty acids
- ✧ In the network view spheres (known as NODES) represent lipid class name and also fatty acids
- ✧ Using this, I can see how many lipid classes share similar fatty acids in the lipidome. This also help to quickly review whole lipidome
- ✧ Size of spheres increase with 1. Number of lipid species for each lipid class and also 2. number of Fatty acids shared by each lipid class
- ✧ Rename the output file with ".layout" extension
- ✧ Once the network is generated in BioLayout software. You can give colors to nodes

BioLayout-CC.pl

- ✧ This program generates input file for BioLayout tool
- ✧ Here EDGE list for network is created using the lipid class names and carbon number
- ✧ In the network view spheres (known as NODES) represent lipid class name and also Carbon number (along with double bonds)
- ✧ Using this, I can see how many lipid classes share similar carbon number in the lipidome
- ✧ Size of spheres increase with 1. Number of lipid species for each lipid class and also 2. number of classes that share same carbon number #
- ✧ Rename the output file with ".layout" extension
- ✧ Once the network is generated in BioLayout software. You can give colors to nodes

All PERL codes are attached at the end of this thesis.

References

Ahmed Z, Mayr M, Zeeshan S, Dandekar T, Mueller MJ, Fekete A (2015) Lipid-Pro: a computational lipid identification solution for untargeted lipidomics on data-independent acquisition tandem mass spectrometry platforms. Bioinformatics 31: 1150–1153

Alipanah L, Rohloff J, Winge P, Bones AM, Brembu T (2015) Whole-cell response to nitrogen deprivation in the diatom *Phaeodactylum tricornutum*. Journal of Experimental Botany 66:6281–6296

Allen JW, DiRusso CC, Black PN (2015) Triacylglycerol synthesis during nitrogen stress involves the prokaryotic lipid synthesis pathway and acyl chain remodeling in the microalgae *Coccomyxa subellipsoidea*. Algal Research 10:110–120

Bai X, Song H, Lavoie M, Zhu K, Su Y, Ye H, Chen S, Fu Z, Qian H (2016) Proteomic analyses bring new insights into the effect of a dark stress on lipid biosynthesis in *Phaeodactylum tricornutum*. Scientific Reports 6:25494

Ball SG, Dirick L, Decq A, Martiat JC, Matagne R (1990) Physiology of starch storage in the monocellular alga *Chlamydomonas reinhardtii*. Plant Science 66:1–9

Bell MV, Pond D (1996) Lipid composition during growth of motile and coccolith forms of *Emiliania huxleyi.* Phytochemistry 2:465–471

Benavente-Valdésa JR, Aguilara C, Contreras-Esquivela JC, Méndez-Zavalab A, Montañezb J (2016) Strategies to enhance the production of photosynthetic pigments and lipids in chlorophycae species. Biotechnology Reports 10:117–125

Benthien A, Zondervan I, Engel A, Hefter J, Terbruggen A, Riebesell U (2007) Carbon isotopic fractionation during a mesocosm bloom experiment dominated by *Emiliania huxleyi*: Effects of CO_2 concentration and primary production. Geochimica Cosmochimca Acta 71:1528–1541

Bendall DS, Howe CJ, Nisbet EG, Nisbet RER (2008) Introduction. Photosynthetic and atmospheric evolution. Philosophical Transactions of the Royal Society B: Biological Sciences 363: 2625–2628

Behrenfeld MJ, O'Malley RT, Siegel DA, McClain CR, Sarmiento JL, Feldman GC, Milligan AJ, Falkowski PG, Letelier RM, Boss ES (2006) Climate-driven trends in contemporary ocean productivity. Nature 444: 752–755

Bitter T, Muir HM (1962) A modified uronic acid carbazole reaction. Anal Biochem 4:330–334

Blankenship RE, Hartman H (1998) The origin and evolution of oxygenic photosynthesis. Trends in Biochemical Scences 23: 94–97

Bligh EG, Dyer WJ (1959) A rapid method of total lipid extraction and purification. Canadian Journal of Biochemistry and Physiology 37: 911–917

Brassell SC, Eglinton G, Marlowe IT, Pflaumann U, Sarnthein M (1986) Molecular Stratigraphy: A new tool for climatic assessment. Nature 320:129–133

Brassell SC (1993) Applications of biomarkers for delineating marine paleoclimatic fluctuations during the Pleistocene. In. Engel MH, Macko SA (eds) Organic Geochemistry, Chapter 34, Plenum Press, New York, pp 699:738

Brown MR, Dunstan GA, Norwood SJ, Miller KA (1996) Effects of harvest stage and light on the biochemical composition of the diatom *Thalassiosira pseudonana*. Journal of Phycology 32:64–73

Brugger B, Erben G, Sandhoff R,Wieland FT, Lehmann WD (1997) Quantitative analysis of biological membrane lipids at the low picomole level by nano-electrospray ionization tandem mass spectrometry. Proceedings of the National Academy of Sciences of the United States of America 94: 2339–2344

Cajka T, Fiehn O (2014) Comprehensive analysis of lipids in biological systems by liquid chromatography-mass spectrometry. Trends in Analytical Chemistry 61:192–206

Choi GG, Kim BH, Ahn CY, Oh HM (2011) Effect of nitrogen limitation on oleic acid biosynthesis in *Botryococcus braunii*. Journal of Applied Phycology 23:1031–1037

Christer S. Ejsing, Eva Duchoslav, Julio Sampaio, Kai Simons, Ron Bonner, Christoph Thiele, Kim Ekroos, Andrej Shevchenko (2006) Automated Identification and Quantification of Glycerophospholipid Molecular Species by Multiple Precursor Ion Scanning. Analytical Chemistry 78: 6202–6214

Conte MH, Volkman JK, Eglinton G (1994) Lipid biomarkers of the Haptophyta. In. Green JC, Leadbeater BSC (eds) The Haptophyte Algae, Clarendon Press, pp. 351–77

Conte MH, Thompson A, Eglinton G (1995) Lipid biomarker diversity in the coccolithophorid *Emiliania huxleyi* (Prymnesiophyceae) and related species *Gephyrocapsa oceanica*. Journal of Phycology 31: 272–282.

Conte MH, Thompson A, Lesley D, Harris RP (1998) Genetic and physiological influences on the alkenone/alkenoate versus growth temperature relationship in *Emiliania huxleyi* and *Gephyrocapsa oceanica*. Geochimica Cosmochimca Acta 62:51–68

Cranwell PA (1985) Long-chain unsaturated ketones in recent lacustrine sediments. Geochimica Cosmochimca Acta 49:1545–1551

Danbara A, Shiraiwa Y (1999) The requirement of selenium for the growth of marine coccolithophorids, *Emiliania huxleyi*, *Gephyrocapsa oceanica* and *Helladosphaera* sp. (Prymnesiophyceae). Plant Cell Physiology 40:762–766

Dassow VP, Ogata H, Probert I, Wincker P, Da Silva C, Audic S, Claverie JM, De Vargas C (2009) Transcriptome analysis of functional differentiation between haploid and diploid cells of *Emiliania huxleyi*, a globally significant photosynthetic calcifying cell. Genome Biology 10: R114

Dunstan GA, Volkman JK, Barrett SM, Garland CD, (1993) Changes in the lipidcomposition and maximization of the polyunsaturated fatty-acid content of 3 microalgae grown in mass-culture. Journal of Appl ied Phycology 5: 71–83

Eltgroth ML, Watwood RL, Wolfe GV (2005) Production and cellular localization of neutral long-chain lipids in the haptophyte algae *Isochrysis galbana* and *Emiliania huxleyi*. Journal of Phycology 41:1000–1009

Elton C.G, Ann C.W, Matias K and Bala R (2016) Metabolic regulation of triacylglycerol accumulation in the green algae: identification of potential targets for engineering to improve oil yield, Plant Biotechnology Journal 14:1649–1660

Epstein BL, D'Hondt SD, Quinn JG, Zhang J, Hargraves PE (1998) An effect of dissolved nutrient concentrations on based temperature estimates alkenone. Paleooceanography 13:122–126

Epstein BL, D'Hondt S, Hargraves PE (2001) The possible metabolic role of C_{37} alkenones in *Emiliania huxleyi*. Organic Geochemistry 32: 867–875

Fan J, Andre C, Xu C (2011) A chloroplast pathway for the de novo biosynthesis of triacylglycerol in *Chlamydomonas reinhardtii*. FEBS Letters 585:1985–1991

Falkowski PG (1994) The role of phytoplankton photosynthesis in global biogeochemical cycles. Photosynthesis Research 39:235–258

Fenn JB, Mann M, Meng CK,Wong SF,Whitehouse CM (1989) Electrospray ionization for mass spectrometry of large biomolecules. Science 246: 64–71

Fenn JB, Mann M, Meng CK,Wong SF,Whitehouse CM (1990) Electrospray ionization-principles and practice. Mass Spectrometry Reviews 9:37–70

Fernández E, Balch WM, Marañón E, Holligan PM (1994) High rates of lipid biosynthesis in cultured, mesocosm and coastal populations of the coccolithophore *Emiliania huxleyi*. Marine Ecology Progress Series 114:13–22

Fernández E, Marañon E, Balch WM (1996a) Intracellular carbon partitioning in the coccolithophorid *Emiliania huxleyi*. Journal of Marine Systems 9:57–66

Fernández E, Maranon E, Harbour DS, Kristiansen S, Heimdal BR (1996b) Patterns of carbon and nitrogen uptake during blooms of *Emiliania huxleyi* in two Norwegian fjords. Journal of Plankton Research 18:2349–2366

Fernández E, Galvan A (2007) Inorganic nitrogen assimilation in *Chlamydomonas*. Journal of Experimental Botony 58:2279–2287

Fichtinger-Schepman AMJ, Kamerling JP, Versluis C, Vliegenthart JFG (1981) Structural studies of the methylated, acidic polysaccharide associated with coccoliths of *Emiliania huxleyi* (Lohmann) Kamptner. Carbohydrate Research 93:105–123

Field CB, Behrenfeld MJ, Randerson JT, Falkowski P (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. Science 281: 237–240

Flynn KJ, Davidson K, Cunningham A (1993a) Relations between carbon and nitrogen during growth of *Nannochloropsis oculata* (Droop) Hibberd under continuous Illumination. New Phytologist 125:717–722

Flynn KJ, Zapata M, Garrido JL, Öpik H, Hipkin CR (1993b) Changes in carbon and nitrogen physiology during ammonium and nitrate nutrition and nitrogen starvation in *Isochrysis galbanaZ*. European Journal of Phycology 28: 47–52

Freeman KH, Hayes JM (1992) Fractionation of carbon isotopes by phytoplankton and estimates of ancient CO_2 levels. Global Biogeochemical Cycles 6:185–198

Fulton JM, Fredricks HF, Bidle KD, Vardi A, Kendrick BJ, DiTullio GR et al. (2014) Novel molecular determinants of viral susceptibility and resistance in the lipidome of *Emiliania huxleyi*. Environmental Microbiology 16:1137–1149.

Garnier M, Bougaran G, Pavlovic M, Berard JB, Carrier G, A Charrier A, Le GF, Lukomska E, , Schreiber N, Cadoret JP, Rogniaux H, Saint-Jean B (2016) Use of a lipid rich strain reveals mechanisms of nitrogen limitation and carbon partitioning

in the haptophyte *Tisochrysis lutea*. Algal Research 2:229–248

Goodson C, Roth R, Wang ZT, Goodenough U (2011) Structural correlates of cytoplasmic and chloroplast lipid body synthesis in *Chlamydomonas reinhardtii* and stimulation of lipid body production with acetate boost. Eukaryotic Cell 10:1592–1606

Han X, Gross RW (2003) Global analyses of cellular lipidomes directly from crude extracts of biological samples by ESI mass spectrometry: a bridge to lipidomics. Journal of Lipid Research 44: 1071–1079

Han X, Gross RW (2005) Shotgun lipidomics: Electrospray ionization mass spectrometric analysis and quantitation of cellular lipidomes directly from crude extracts of biological samples. Mass Spectrometry Reviews 24: 367–412

Han X, Yang K, Gross RW (2011) Multi-dimensional mass spectrometry-based shotgun lipidomics and novel strategies for lipidomic analyses. Mass Spectrometry Reviews 31:134–78

Harwood JL, Jones AL (1989) Lipid metabolism in algae. Advances in Botanical Research 16: 1–53

Hartler J, Trötzmüller M, Chitraju C, Spener F, Köfeler HC, Thallinger GG (2011) Lipid Data Analyzer: unattended identification and quantitation of lipids in LC-MS data. *Bioinformatics* 27: 572-577

Herzog R, Schuhmann K, Schwudke D, Sampaio JL, Bornstein SR, Schroeder M, Shevchenko A (2012) LipidXplorer: a software for consensual cross-platform

lipidomics. PlosONE 7:e29851

Hodge JE, Hofreiter BT (1962) Determination of reducing sugars and carbohydrates. In. Whistler RL, Wolfram ML (eds) Methods in carbohydrate chemistry, vol. 1, pp 380–394

Holligan PM, Fernandez E, Aiken J, Balch WM, Boyd P, Burkill PH, et al. (1993) A biogeochemical study of the coccolithophore, *Emiliania huxleyi*, in the North-Atlantic. Global Biogeochemical Cycles 7: 879–900

Houdan A, Probert I, Van Lenning K, Lefebvre S (2005) Comparison of photosynthetic responses in diploid and haploid life-cycle phases of *Emiliania huxleyi* (Prymnesiophyceae). Marine Ecology Progress Series 292: 139–146

Hummel Jan, Segu Shruthi, Li Yan, Irgang Susann, Jueppner Jessica, Giavalisco Patrick, (2011) Ultra Performance Liquid Chromatography and High Resolution Mass Spectrometry for the Analysis of Plant Lipids. Frontiers in Plant Science 2:54

Hunter JE, Frada MJ, Fredricks HF, Vardi A, Van Mooy BAS (2015) Targetted and Untargetted lipidomics of *Emiliania huxleyi* viral infection and life cycle phases highlights molecular biomarkers of infection, susceptibility, and ploidy. Frontiers in Marine Science 2:81

Iglesias-Rodriguez MD, Brown CW, Doney SC, Kleypas J, Kolber D, Kolber Z, et al. (2002) Representing key phytoplankton functional groups in ocean carbon cycle models: coccolithophorids. Global Biogeochemical Cycles 16: 47-1–47-20

Ikeda K, Mutoh M, Teraoka N, Nakanishi H, Wakabayashi K, Taguchi R (2011)

Increase of oxidant-related triglycerides and phosphatidylcholines in serum and small intestinal mucosa during development of intestinal polyp formation in Min mice. Cancer Science 102:79

Isabel Armada, Ismael Hachero-Cruzado, Narciso Mazuelos, José Luis Ríos, Manuel Manchado, José Pedro Cañavate, (2013) Differences in betaine lipids and fatty acids between VLP and VLP isolates (Haptophyta). Phytochemistry 95: 224–233

Iwai M, Ikeda K, Shimojima M, Ohta H (2014) Enhancement of extraplastidic oil synthesis in *Chlamydomonas reinhardtii* using a type-2 diacylglycerol acyltransferase with a phosphorus starvation-inducible promoter. Plant Biotechnology Journal 12:808–819

Jasper JP, Hayes JM (1990) A carbon isotope record of CO_2 levels during the late Quaternary. Nature 347:462–464

Jeffrey SW (1972) Preparation and some properties of crystalline chlorophyll c1 and c2 from marine algae. Biochemica et Biophysica Acta 279:15–33

Jia J, Han D, Gerken HG, Li Y, Sommerfeld M, Hu Q, Xu J (2015) Molecular mechanisms for photosynthetic carbon partitioning into storage neutral lipids in *Nannochloropsis oceanica* under nitrogen-depletion conditions. Algal Research 7:66–77

Johnson X, Alric J (2013) Central carbon metabolism and electron transport in *Chlamydomonas reinhardtii*: metabolic constraints for carbon partitioning between oil and starch. Eukaryotic Cell 12: 776–793

Kaffes A, Thoms S, Trimborn S, Rost B, Langer G, Richter KU, Angela K, Norici A, Giordano M (2010) Carbon and nitrogen fluxes in the marine coccolithophore *Emiliania huxleyi* grown under different nitrate concentrations. Journal of Experimental Marine Biology and Ecology 393:1–8

Kayano K, Shiraiwa Y (2009) Physiological regulation of coccolith polysaccharide production by phosphate availability in the coccolithophorid *Emiliania huxleyi*. Plant Cell Physiology 50:1522–1531

Kimberly J Popendorf, Helen F. Fredricks, Benjamin A. S. Van Mooy (2013) Molecular Ion-Independent Quantification of Polar Glycerolipid Classes in Marine Plankton Using Triple Quadrupole MS. Lipids 48:185–195

Kind T, Liu KH, Lee DY, DeFelice B, Meissen JK, Fiehn O (2013) LipidBlast *in silico* tandem mass spectrometry database for lipid identification. Nature Methods 10:755–758

Klein U (1987) Intracellular Carbon Partitioning in *Chlamydomonas reinhardtii,* Plant Physiology 85:892–897

Knittelfelder OL, Weberhofer BP, Eichmann TO, Kohlwein SD, Rechberger GN (2014) A versatile ultra-high performance LC-MS method for lipid profiling. Journal of Chromatography B, Analytical Technologies in the Biomedical and Life Sciences 952:119–128.

Kotajima T, Shiraiwa Y, Suzuki I (2014) Functional screening of a novel Δ15 fatty acid desaturase from the coccolithophorid Emiliania huxleyi. Biochimica et Biophysica Acta 1842:1451–1458

Lacour T, Sciandra A, Talec A, Mayzaud P (2012) Neutral lipid and carbohydrate productivities as a response to Nitrogen status in *Isochrysis sp.* (T-ISO; Haptophyceae): Starvation versus limitation. Journal of Phycology 48:647–656

Laws EA (1991) Photosynthetic quotients, new production and net community production in the open ocean. Deep-Sea Research 38:143–167

Li Y, Han D, Sommerfeld M, Hu Q (2011) Photosynthetic carbon partitioning and lipid production in the oleaginous microalga *Pseudochlorococcum sp.* (Chlorophyceae) under nitrogen-limited conditions. Bioresource Technology 102:123–129

Loebl M, Cockshutt AM, Campbell D, Finkel ZV (2010) Physiological basis for high resistance to photoinhibition under nitrogen depletion in E*miliania huxleyi*. Limnology and Oceanography 55: 2150–2160

Lloyd NDH, Canvin DT, Culver DA (1977) Photosynthesis and Photorespiration in Algae. Plant Physiology 59:936-940

Malitsky S, Ziv C, Rosenwasser S, Zheng S, Schatz D, Porat Z, Ben-Dor S, Aharoni A, Vardi A (2016) Viral infection of the marine alga *Emiliania huxleyi* triggers lipidome remodeling and induces the production of highly saturated triacylglycerol. New Phytologist 210:88–96

Marlowe IT, Brassell SC, Eglinton G, Green JC (1984) Long chain unsaturated ketones and esters in living algae and marine sediments. Org Geochemistry 6:135–141

McKew BA, Stephane CL, Eric PA, Gergana M, Christine AR, Metodi V, Richard J (2013) Plasticity in the proteome of *Emiliania huxleyi* CCMP 1516 to extremes of

light is highly targeted. New Phytologist 200:61–73

Meng D, Zhang Q, Gao X, Wu S, Lin G (2014) LipidMiner: a software for automated identification and quantification of lipids from multiple liquid chromatography-mass spectrometry data files. Rapid Communications in Mass Spectrometry : Rapid Communications in Mass Spectrometry 28:981–985

Michel G, Tonon T, Scornet D, Mark C.J and Kloareg B (2010) Central and storage carbon metabolism of the brown alga *Ectocarpus siliculosus*: insights into the origin and evolution of storage carbohydrates in Eukaryotes. New Phytologist 188: 67–81

Msanne J, Xu D, Konda AR, Casas-Mollano JA, Awada T, Cahoon EB, Cerutti H (2012) Metabolic and gene expression changes triggered by nitrogen deprivation in the photoautotrophically grown microalgae *Chlamydomonas reinhardtii* and *Coccomyxa sp. C*-169. Phytochemistry 75:50–59

Müller PJ, Kirst G, Ruhland G, von Storch I, Rosell-Mele A (1998) Calibration of the alkenone paleotemperature index U37K' based on core-tops from the eastern South Atlantic and the global ocean (60°N–60°S). Geochimica et Cosmochimica Acta 62:1757–1772

Murphy J, Riley JP (1962) A modified single solution method for the determination of phosphate in natural waters. Analytica Chimica Acta 27:31–36

Nakamura H, Sawada K, Araie H, Suzuki I, Shiraiwa Y (2014) Long chain alkenes, alkenones and alkenoates produced by the haptophyte alga *Chrysotila lamellosa* CCMP1307 isolated from a salt marsh. Org Geochemistry 66:90–97

Negi S, Barry AN, Friedland N, Sudasinghe N, Subramanian S, Shayani P, Omar FH, Barry D, Tanner S, Richard S (2016) Impact of nitrogen limitation on biomass, photosynthesis, and lipid accumulation in *Chlorella sorokiniana*. Journal of Applied Phycology 28:803–812

Nunes-Nesi A, Fernie AR, Stitt M (2010) Metabolic and signaling aspects underpinning the regulation of plant carbon nitrogen interactions. Molecular Plant 3:973–996

Obata T, Schoenefeld S, Krahnert I, Bergmann S, Scheffel A, Fernie AR (2013) Gas-chromatography mass-spectrometry (GC-MS) based metabolite profiling reveals mannitol as a major storage carbohydrate in the coccolithophorid alga *Emiliania huxleyi*. Metabolites 3:168–184

Ohi N, Shiraiwa Y (2015) Lipidomics in marine microalgae using electronspray ionization tandem mass spectrometry. Radioisotopes 64:255–264

Ono M, Sawada K, Shiraiwa Y, Kubota M (2012) Changes in alkenone and alkenoate distributions during acclimatization to salinity change in *Isochrysis galbana*: Implication for alkenone-based paleosalinity and paleothermometry. Geo Chemical Journal 46:235–247

O'Neil GW, Williams JR, Wilson-Peltier J, Knothe G, Reddy CM (2016) Experimental Protocol for Biodiesel Production with Isolation of Alkenones as Coproducts from Commercial *Isochrysis* Algal Biomass. Journal of Visualized Experiments doi:10.3791/54189

Pagani M, Freeman KH, Ohkouchi N, Caldeira K (2002) Comparison of water column [CO_2aq] with sedimentary alkenone-based estimates: a test of the alkenone-CO_2

proxy. Paleoceanography 4:21-1–21-12

Pan H, Sun MY (2011) Variations of alkenone based paleotemperature index ($U^{K`}_{37}$) during *Emiliania huxleyi* cell growth, respiration (auto-metabolism) and microbial degradation. Organic Geochemistry 42:678–687

Pan H, Culp RA, Noakes JE, Sun MY (2014) Effects of growth stages, respiration, and microbial degradation of phytoplankton on cellular lipids and their compound-specific stable carbon isotopic compositions. Journal of Experimental Marine Biology and Ecology 461:7–19

Pan H, Culp RA, Sun MY (2017) Influence of physiological states of *Emiliania huxleyi* cells on their lipids and associated molecular isotopic compositions during microbial degradation. Journal of Experimental Marine Biology and Ecology 488:1–9

Parrish CC, Wells JS, Yang ZP, Dabinett P (1998) Growth and lipid composition of scallop juveniles, *Placopecten magellanicus*, fed the agellate *Isochrysis galbana* with varying lipid composition and the diatom *Chaetoceros muelleri*, Mar Biology 133: 461–471

Pond DW, Harris RP (1996) The lipid composition of the coccolithophore *Emiliania huxleyi* and its possible ecophysiological significance. Journal of the Marine Biological Association of the United Kingdom 76:579–594

Popp BN, Kenig F, Wakeham SG, Laws EA, Bidigare RR (1998) Does growth rate affect ketone unsaturation and intercellular carbon isotopic variability in *Emiliania huxleyi*? Paleoceanography 13: 35–41

Prahl FG, Wakeham SG (1987) Calibration of unsaturation patterns in long-chain ketone compositions for palaeotemperature assessment. Nature 330:367–369

Prahl FG, Wolfe GV, Sparrow MA (2003) Physiological impacts on alkenone paleothermometry. Paleoceanography 18:1025

Pluskal T, Castillo S, Villar-Briones A, Oresic M (2010) MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinformatics 11:395

Raven JA, Crawfurd K (2012) Environmental controls on coccolithophore calcification. Marine Ecology Progress Series 470:137–166

Read BA, Kegel J, Klute MJ, Kuo A, Lefebvre SC, Maumus F, et al. (2013) Pan genome of the phytoplankton *Emiliania* underpins its global distribution. Nature 499:209–213

Richard Y (2012) Do alternative energy sources displace fossil fuels? Nature Climate Change 2:441-443

Riegman R, Stolte W, Noordeloos AAM, Slezak D (2000) Nutrient uptake and alkaline phosphatase (EC 3:1:3:1) activity of *Emiliania huxleyi* (Prymnesiophyceae) during growth under N and P limitation in continuous cultures. J of Phycology 36:87–96

Reitan K.I, Rainuzzo J.R and Olsen Y (1994) Effect of nutrient limitation on fatty acid and lipid content of marine microalgae. Journal of Phycology 30:972–979

Riebesell U, Revill AT, Holdsworth DG, Volkman JK (2000) The effects of varying CO_2 concentration on lipid composition and carbon isotope fractionation

in *Emiliania huxleyi*. Geochimica Cosmochimica Acta 64:4179–4192

Rokitta SD, Peter VD, Björn R, Uwe J (2014) *Emiliania huxleyi* endures N-limitation with an efficient metabolic budgeting and effective ATP synthesis. BMC Genomics 15:1051

Rontani BB, Volkman JK (2004) Long-chain alkenones and related compounds in the benthic haptophyte *Chrysotila lamellosa* Anand HAP 17. Phytochemistry 65:117–126

Rontani JF, Prahl FG, Volkman JK (2006) Re-examination of the double bond positions in alkenones and derivatives: biosynthetic implications. Journal of Phycology 42:800–813

Rosenwasser S, Mausz M A, Schatz D, Sheyn U, Malitsky S, Aharoni A, et al. (2014) Rewiring host lipid metabolism by large viruses determines the fate of *Emiliania huxleyi*, a bloom-forming alga in the Ocean. Plant Cell 26:2689–2707.

Sawada K, Handa N, Shiraiwa Y, Danbara A, Montani S (1996) Long-chain alkenones and alkyl alkenoates in the coastal and pelagic sediments of the northwest North Pacific, with special reference to the reconstruction of *Emiliania huxleyi* and *Gephyrocapsa oceanica* ratios. Organic Geochemistry 24:751–764

Sawada K, Shiraiwa Y (2004) Alkenone and alkenoic acid compositions of the membrane fractions of *Emiliania huxleyi*. Phytochemistry 65:1299–1307

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Ideker T (2003) Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Research 13: 2498–2504

Shannon A Roche, Jeffrey D Leblond (2010) Betaine lipids in chlorarachniophytes. Physiological Research 58:298–305

Shemi A, Ben-Dor S, Vardi A (2015) Elucidating the composition and conservation of the autophagy pathway in photosynthetic eukaryotes. Autophagy 11:701–715.

Shemi A, Schatz D, Fredricks HF, Van Mooy BA, Porat Z, Vardi A (2016) Phosphorus starvation induces membrane remodeling and recycling in Emiliania huxleyi. New Phytologist 211:886–898

Shi Q, Araie H, Bakku RK, Fukao Y, Rakwal R, Suzuki I, Shiraiwa Y (2015) Proteomic analysis of lipid body from the alkenone-producing marine haptophyte alga Tisochrysis lutea, Proteomics 15:4145–4158

Simionato D, Block MA, La Rocca N, Jouhet J, Maréchal E, Finazzi G, Morosinotto T (2013) The response of *Nannochloropsis gaditana* to nitrogen starvation includes *de novo* biosynthesis of triacylglycerols, a decrease of chloroplast galactolipids, and reorganization of the photosynthetic apparatus. Eukaryotic Cell 12: 665–676

Singh A, Elvitigala T, Bhattacharyya-Pakrasi M, Aurora R, Ghosh B, Pakrasi H (2008) Integration of carbon and nitrogen metabolism with energy production is crucial to light acclimation in the cyanobacterium *Synechocystis*. Plant Physiology 148: 467–478

Sorrosa JM, Yamamoto M, Siraiwa Y (2003) Non thermal factors affecting production and unsaturation of alkenones in *Emiliania huxleyi* and *Gephyrocapsa oceanica*. In. Murata N, Yamada M, Nishida I, Okuyama H, Sekiya J, Wada H (eds) Adv Res

Plant Lipids, Kluwer Academic Publishers, Dordrecht, Netherlands, pp 133:136

Sprecher H (1992) A reevaluation of the pathway for the biosynthesis of 4,7,10,13,16,19-docosahexaenoic acid. Omega-3 News 7:1–3

Sprecher H, Luthria DL, Mohammed BS, Baykousheva SP (1995) Reevaluation of the pathways for the biosynthesis of polyunsaturated fatty acids. Journal of lipid research 36:2471–2477

Sprecher H (2000) Metabolism of highly unsaturated n−3 and n−6 fatty acids. Biochimica Biophysica Acta 1486:219–231

Stephen Milne, Pavlina Ivanova, Jeffrey Forrester, H. Alex Brown, (2006) Lipidomics: An analysis of cellular lipids by ESI-MS. Methods 39:92–103

Stahlman M, Ejsing CS, Tarasov K, Perman J, Borén J, Ekroos K (2009) High-throughput shotgun lipidomics by quadrupole time-of-flight mass spectrometry. Journal of Chromatography B 877: 2664–2672

Sturt HF, Summons RE, Smith K, Elvert M, Hinrichs K-U (2004) Intact polar membrane lipids in prokaryotes and sediments deciphered by high-performance liquid chromatography/electrospray ionization multistage mass spectrometry–new biomarkers for biogeochemistry and microbial ecology. Rapid Communications in Mass Spectrometry 18:617–628 7

Taguchi R, Houjou T, Nakanishi H, Yamazaki T, Ishida M, Imagawa M, Shimizu T. (2005) Focused lipidomics by tandem mass spectrometry. Journal of Chromatography B 823: 26–36

Talmy D, Blackford J, Hardman-Mountford NJ, Polimene L, Follows MJ, Geider RJ

(2014) Flexible C : N ratio enhances metabolism of large phytoplankton when resource supply is intermittent. Biogeosciences 11:4881–4895

Theissen KM, Zinniker DA, Moldowan JM, Dunbar RB, Rowe HD (2005) Pronounced occurrence of long-chain alkenones and dinosterol in a 25,000-year lipid molecular fossil record from Lake Titicaca, South America. Geochimica Cosmochimica Acta 69:623–636

Theocharidis A, van Dongen S, Enright AJ, Freeman TC, (2009) Network visualization and analysis of gene expression data using BioLayout Express3D, Nature Protocols 4:1535 – 1550

Theroux S, D'Andrea WJ, Toney J, Amaral-Zettler L, Huang YS (2010) Phylogenetic diversity and evolutionary relatedness of alkenone-producing haptophyte algae in lakes: Implications for continental paleotemperature reconstructions. Earth and Planetary Science Letters 300:311–320

Toney JL, Huang Y, Fritz SC, Baker PA, Grimm E, Nyren P (2010) Climatic and environmental controls on the occurrence and distributions of long chain alkenones in lakes of the interior United States. Geochimica Cosmochimica Acta 74:1563–1578

Tsuji Y, Yamazaki M, Suzuki I, Shiraiwa Y (2015) Quantitative analysis of carbon flow into photosynthetic products functioning as carbon storage in the marine coccolithophore, *Emiliania huxleyi*. Marine Biotechnology 17:428–440

Turpin DH (1991) Effects of inorganic N availability on algal photosynthesis and carbon metabolism. Journal of Phycology 27:14–20

Van Mooy BAS, Fredricks HF, Pedler BE, Dyhrman ST, Karl DM, Koblı´zek M, Lomas MW, Mincer TJ, Moore LR, Moutin T, Rappe´ MS, Webb EA (2009) Phytoplankton in the ocean substitute phospholipids in response to phosphorus scarcity. Nature 458:69–72

Vardi A, Van Mooy BA, Fredricks HF, Popendorf KJ, Ossolinski JE, Haramaty L, Bidle KD (2009) Viral glycosphingolipids induce lytic infection and cell death in marine phytoplankton. Science 326: 861–865

Varum KM, Kvam BJ, Myklestad S (1986) Structure of a food-reserve β-D-glucan produced by the haptophyte alga *Emiliania huxleyi* (Lohmann) Hay and Mohler. Carbohydrate Research 152:243–248

Volkman JK, Eglinton G, Corner EDS, Forsserg TEV (1980a) Long-chain alkenes and alkenones in the marine coccolithophorid *Emiliania huxleyi*. Phytochemistry 19:2619–2622

Volkman JK, Eglinton G, Corner EDS, Sargent JR (1980b) Novel unsaturated straight-chain C_{37}-C_{39} methyl and ethyl ketones in marine sediments and a coccolithophore *Emiliania huxleyi*. Physics and Chemistry of the Earth 12:219–227

Volkman JK, Barrett SM, Blackburn SI, Sikes EL (1995) Alkenones in Gephyrocapsa oceanica: implications for studies of paleoclimate. Geochimica Cosmochimica Acta 59:513–520

Watson AD (2006) Thematic review series: systems biology approaches to metabolic and cardiovascular disorders. Lipidomics: a global approach to lipid analysis in

biological systems. Journal of Lipid Research 47: 2101–2111.

Weger HG, Turpin DH (1989) Mitochondrial respiration can support NO_3^- and NO_2^- reduction during photosynthesis. Interactions between photosynthesis, respiration, and N assimilation in the N-limited green alga *Selenastrum minutum*. Plant Physiology 89:409–415

Welti R, Li W, Li M, Sang Y, Biesiada H, Zhou HE, Rajashekar CB, Williams TD, Wang X (2002) Profiling of membrane lipids in plant stress responses. Journal of Biological Chemistry 277: 31994–32002

Wenk MR (2005) The emerging field of lipidomics. Nature Reviews Drug Discovery 4: 594–610.

Xu L, Reddy CM, Farrington JW, Frysinger GS, Gaines RB, Johnson CG, Nelson RK, Eglinton TI (2001) Identification of a novel alkenone in Black Sea sediments. Organic Geochemistry 32: 633–645

Yamamoto M, Shiraiwa Y, Inouye I (2000) Physiological responses of lipids in Emiliania huxleyi and Gephyrocapsa oceanica (Haptophyceae) to growth status and their implications for alkenone paleothermometry. Organic Geochemistry 31:799-811

Yetukuri L, Katajamaa M, Medina-Gomez G, Seppanen-Laakso T, Vidal-Puig A, Oresic M (2007) Bioinformatics strategies for lipidomics analysis: characterization of obesity related hepatic steatosis. BMC Systems Biology 1: 12

Zhao J, An C, Longo WM, Dillon JT, Zhao Y, Shi C, Chen Y, Huang Y (2014) Occurrence of extended chain length C41 and C42 alkenones in hypersaline lakes. Organic Geochemistry 75: 48–53

Zhou Z, Marepally SR, Nune DS, Pallakollu P, Ragan G, Roth MR, Wang L, Lushington GH, Visvanathan M, Welti R (2011) LipidomeDB data calculation environment: online processing of direct-infusion mass spectral data for lipid profiles. Lipids 46: 879–884

Acknowledgements

I am glad to have this opportunity to thank all those who have filled valuable experiences in my life and extended their hands in completing this research work.

I would like to express my deepest heartfelt gratitude to Professor Dr. Iwane Suzuki, University of Tsukuba. It is his affectionate guidance and constant encouragement that motivated me to successfully carryout this work.

Mere words cannot express my gratitude for Professor Dr. Yoshihiro Shiraiwa, University of Tsukuba. His valuable suggestions and the scientific knowledge he provided to me during discussions were precious.

I am also grateful to Dr. Hiroya Araie, Kanto Gakuin University, for his constant motivation and guidance throughout my work. I would like to express my gratitude to him.

I also thank Dr. Hanawa, University of Tsukuba, for his support and training me on various techniques

Very special thanks to all my current and past lab members at University of Tsukuba. This is because of them a stimulating scientific environment is around me all the time. Thanks for great discussion and exchange of knowledge.

I extend my special thanks and gratitude to Dr. Randeep Rakwal, University of Tsukuba, for his personal care and support throughout my stay in Japan.

My acknowledgement will be incomplete if I don’t express my thanks to my parents and family who supported me all the time and encouraged me in every aspect of my life.

Perl scripts

1. **Lipid-compare.pl**

2. **Pre-BioLayout.pl**

3. **BioLayout-Lipid-FA.pl**

4. **BioLayout-CC.pl**

```
##########################################
#              Lipid-compare.pl          #
##########################################


use List::MoreUtils qw(uniq);
$list="list"; #input file name here. File consist list of files to be compared; names of two
files written in one line separated by a tab
open(LIST,$list);@Ls=<LIST>;$i=0; $len=@Ls;

$tf=~s/:/Y/g;my $FAT;
$falist="FA-list";open(FC,$falist);#File name. This file consist all fatty acids that are
expected to present in lipidome.
@FAS=<FC>;$LF=@FAS;open OUTFA,"+>","FATAble";
my @Fci;my @Fbi; my @Fcj; my @Fbj;
my @FAtable;$Fci[0][0]=0;$Fbi[0][0]=0;
$Fcj[0][0]=0;$Fbj[0][0]=0;$FAtable[0][0]=0;
for($fi=0;$fi<$LF;$fi++)
{ chomp $FAS[$fi];
  @FI= split /\:/,$FAS[$fi];#print "$LF; $FI[0]:$FI[1]\n";
  $Fci[0][$fi]=$FI[0];$Fbi[0][$fi]=$FI[1];
  $Fcj[$fi][0]=$FI[0];$Fbj[$fi][0]=$FI[1];
  $FAtable[$fi][0]="$FAS[0]";$FAtable[0][$fi]="$FAS[0]";
}
     for($fi2=1;$fi2<$LF;$fi2++)
     {
       for($fi3=1;$fi3<$LF;$fi3++)
        {
         $aa=$Fci[0][$fi2]+$Fcj[$fi3][0];$bb=$Fbi[0][$fi2]+$Fbj[$fi3][0];
         $FAtable[$fi2][$fi3]="$aa:$bb";#$FAtable[$fi2][$fi3]=~s/:/y/g;
             #print OUTFA "FA: $FAtable[$fi2][$fi3]\t";
        }#print OUTFA "\n";

     }



my @load;
foreach $line(@Ls)
{

 chomp $line;
 @load= split /\t+/,$line;#Loads names of files in array elements #better two file names per line
 $fnum=@load;@name1= split /\./,$load[0];@name2= split
 /\./,$load[1];$outflDIR="$name1[0]-$name2[0]";
 mkdir "$outflDIR";
 #open OUT1a,"+>","$name1[0]-common.txt";
 #open OUT1b,"+>","$name2[0]-common.txt";
 @Filenm= split /\-/,$outflDIR;
 my @ARY12;$x=0;my @Len;undef @FFAA1;undef @Flen;
     for ($i=0;$i<$fnum;$i++)
     {$F1="$load[$i]";my @F1array;
      chomp $F1;
      unless (open(F1,$F1)){print "Error-1:Cannot open file-1 $F1";};@F1=<F1>;$Len[$x]=@F1;$ij=0;
```

```
        foreach $fl1(@F1)
        {
                chomp $fl1;$fl11= rearrange($fl1);#print "ONE:$fl1\nTWO:$fl11\n";#can remove
                underscore in experiment details from raw file
                my @F1line; @F1line= split /\t+/,$fl11;$len=@F1line;#print "Length1:$len\n";
                $new=$F1line[0];$mass=$F1line[1];$pis=$F1line[2];
                #$compute=replace($new);
                #splitting each line of file 1 with SPACE/TAB as delimmiter and saving in an
                array
         push (@F1array, \@F1line);$FFAA1[$i][$ij]=$fl11;$Flen[$i][$ij]="$len";
                # Saving the array elements of a line in another array.. technically making
                multi dimensional array
         $ij++;
                }

         push (@ARY12, \@F1array);#stores in another array a[][][]
       $x++;
      }
for($j=0;$j<$fnum;$j++)
{
      #print "Loop1:$j";
      $LL=$Len[$j];

      for($j2=0;$j2<$fnum;$j2++)
  {
                $LL2=$Len[$j2]-1;
                if($j2 ne $j)
                    { undef @head1; undef @head2;
                         open OUT2,"+>","$outflDIR/$Filenm[$j]-UNIQUE.txt";
                         #output file consist lipids that are unique betwween files
                         compared (positive/negative). This data can be discarded

              if($j==0){open OUT1,"+>","$outflDIR/$Filenm[$j]-common.txt";}
              #output file with common lipids in files compared. This data is important.

              if($j==0){open OUT3,"+>","$outflDIR/$Filenm[$j]-common-UNN.txt";}
                         open OUT4,"+>","$outflDIR/$Filenm[$j]-UNIQUE-UNN.txt";
                         #UNN is unknown data. this have fatty acids which are not listed
                         in FA-list. This data can be discarded.


                         for ($k=0;$k<$LL;$k++)
                               { $C=0;$U=0;undef @CTOT; undef @UTOT;#print"$j,$j2, $k,
                               $ARY12[$j][$k][0]\n";
                                 $LP1=$ARY12[$j][$k][0];$mz1=$ARY12[$j][$k][1];$LP1=~
                                 s/\r|\n//g;$mz1=~ s/\r|\n//g;
                                 $pis1=$ARY12[$j][$k][2];$cnt1=$ARY12[$j][$k][3];#print
                                 "$ARY12[$j][$k][3]\n";
                                 $new=$LP1; $Lp1= replace($new);@head1= split
                                 /\s/,$Lp1;$am=$head1[0];$Q1=replace2($am);
                                 $HUNT=0;#print "Q1:$am\n";

                    for ($k2=0;$k2<=$LL2;$k2++)
```

```
 {
$LP2=$ARY12[$j2][$k2][0];$mz2=$ARY12[$j2][$k2][1];
                          $LP2=~ s/\r|\n//g;$mz2=~ s/\r|\n//g;

                          $pis2=$ARY12[$j2][$k2][2];$cnt2=$ARY12[$j2][$k2
                          ][3];
$new=$LP2;$Lp2= replace($new);@head2= split
/\s/,$Lp2;$am=$head2[0];$Q2=replace2($am);
#print "subject: $Q1 and $Q2 \n\n";
if ($Q2=~m/$Q1/)
  {$Real=0;#print "subject: $Lp1 and $Lp2 : $ARY12[$j][$k][3] and
  $ARY12[$j2][$k2][3] \n\n";
                                  if("Negative"=~m/$ARY12[$j][$k][3]/ ||
                                  $len<3)

                                  {$new1=$ARY12[$j2][$k2][0];$N1=0;}else{$N1=
                                  1;}
                                  if("Negative"=~m/$ARY12[$j2][$k2][3]/  ||
                                  $len<3)

                                  {$new2=$ARY12[$j][$k][0];$N2=0;}else{$N2=1;
                                  }
                                  #print
                                  "Neg:$ARY12[$j2][$k2][0]\n";#cannot
                                  transfer values of array elements to
                                  variables
                                  if($N1==$N2 && $N1==0)
                                    {$N=$N1;$new=$ARY12[$j][$k][0];}
                                  elsif($N1==$N2 &&
                                  $N1==1){$N=$N1;$new=$ARY12[$j][$k][0];}
                                  elsif($N1!=$N2)
                                     {

                                       if($N1==0){$N=$N1;$new=$ARY12[$j][$k][
                                       0];}

                                       if($N2==0){$N=$N2;$new=$ARY12[$j2][$k2
                                       ][0];}
                                     }#print "$N and $new\n";
   if($len<3)

                                 {if($FFAA1[$j][$k]=~m/FA/){$new=$ARY12[$j][$
                                 k][0]}

                                 elsif($FFAA1[$j2][$k2]=~m/FA/){$new=$ARY12[$
                                 j2][$k2][0]}}
    $lpp=replace($new,$N);$lpp=~ s/\r|\n//g;if($N eq
    0){$lipid_fa=findFA($lpp);}
                                  @tot= split /\t/,$FFAA1[$j][$k];
                                  @tote=split
                                  /[\s]/,$lpp;$tot[0]="$tote[0]\t($lipid_fa)"
                                  ;
                                  $ctot1= join("\t",@tot);#print"$ctot1\n";
```

```
                        @tot2= split /\t/,$FFAA1[$j2][$k2];splice @tot2,0,1;
                                        $ctot2 = join("\t",@tot2);
                        $ctot1=~ s/\r|\n//g;$ctot2=~ s/\r|\n//g;
                                        $CTOT[$C]= "$ctot1\t$ctot2";#print
                                        "$CTOT[$C]\n";
                        $value="$tot[0]\t$tot[1]";$Real=SearchFA($value,$N);
                                        #print"Value:$value: $Real\n\n";
                                        if($Real>1 && $j==0){print
                                        OUT1"$CTOT[$C]\n";}elsif($Real<=1){print
                                        OUT3"$CTOT[$C]\n";}
                                        $C++;#print OUT1"$CTOT[$C]\n";
                        #print OUT1b "$FFAA1[$j2][$k2]\n";
                                        $HUNT=1;goto nl1;
                        }
                }nl1:
                if($HUNT<1)

                {$N=0;$Real=0;

                                        if("Positive"=~m/$ARY12[$j][$k][3]/  ||
                                        $len<3){$N=1;$new=$ARY12[$j][$k][0];}
                                        if("Negative"=~m/$ARY12[$j][$k][3]/ ||
                                        $len<3){$N=0;$new=$ARY12[$j][$k][0];}
                                        #if($len<3){$new=$ARY12[$j][$k][0];$N=0;}
                                          $lpp=replace($new,$N);if($N eq
                                          0){$lipid_fa=findFA($lpp);}
                                             @tot= split /\t/,$FFAA1[$j][$k];
                                             @tote=split
                                             /[\s]/,$lpp;$tot[0]="$tote[0]\t($lipid_fa)"
                                             ;$utot1= join("\t",@tot);
                        #@tot2= split /\t/,$FFAA1[$j2][$k2];splice @tot2,0,2;$utot2 =
                        join("\t",@tot2);
                        $utot1=~ s/\r|\n//g;
                                        $UTOT[$U]=
                                        "$utot1";$value="$tot[0]\t$tot[1]";#print
                                        "$tot[0]\t$tot[1]\n";
                        if($N==0){$Real=SearchFA($value);}else{$Real=10;}
                                        if($Real>1 && $j==0){print
                                        OUT2"$UTOT[$U]\n";}elsif($Real<=1){print
                                        OUT4"$UTOT[$U]\n";}
                                        $U++;
                #just a print statement and subroutee GET
                }

            }
        #next line of first file
        }

    }
}

}
```

```
sub rearrange
{

 #subprogram that is used for comparinf lipids
 my $Exd; my @exdline; my $rearrange; my @F0line; @F0line= split /\t+/,$fl1;
 @exdline= split /[\_|\:]/,$F0line[1];$Exd= join("\t",@exdline);
 $F0line[1]=$Exd;#print "$Exd\n";
 $F0line[1]=~ s/\r|\n//g;
 $rearrange= join("\t",@F0line);
 #print "$rearrange\n";
 return($rearrange);

}

sub replace
{
 #subprogram to for data processing
 #print "$new, $mass / pis:$pis\n";
 undef @space1;my $char; my $LPn; undef @Exact; my $lpa;
 #$adduct="";
 @space1= split /[\+|\s+|\-|\(|\)|\/|\:|\t+]/,$new;$lnt=@space1;
 #print "Value for neg: $N\n";
$lpa="$space1[0]($space1[1]:$space1[2])";
if($N==1){goto s1;}
else
 { $fa=0;my $a1;$c=0;
   for ($ln1=0;$ln1<=$lnt;$ln1++)
       {
        if ($space1[$ln1]=~m/FA/){$a1.="FA $space1[$ln1+1]:$space1[$ln1+2]/";$c=1}
       }
     if($c>0)
    {
    $lpa.=" ($a1)";#print "neg: $lpa\n";
    }
 }s1:
#print "$N and $lpa\n";
 return($lpa);

}
sub replace2
{
 #subprogramming for data processing
 my @space;my @brks;my @pls;my @sls;my @col;
 my @brke;my @hyp;my $lp1a;my $lp2a;my $lp3a;
 my $lp4a;my $lp5a; my $lp6a;my $lp7a;
 @space= split /\s+/,$am;$lp1a=join('x',@space);@col= split /\:/,$lp1a;$lp2a=join('y',@col);
 @brks= split /\(/,$lp2a;$lp3a=join('z',@brks);@brke= split /\)/,$lp3a;$lp4a=join('q',@brke);
 @pls= split /\+/,$lp4a;$lp5a=join('w',@pls);@hyp= split /\-/,$lp5a;$lp6a=join('v',@hyp);
 @sls= split /\//,$lp6a;$lp7a=join('h',@sls);
 return($lp7a);

}
```

```
sub findFA
{
        #subprogram  that calculates other fatty acid using total carbons (CC) and double
        bonds(DB) of lipid and Cc,Db of FAs detected
    my @space1;my $char;my $faa;my$FA;my @IDs;my $AL;my $Lipid;
     @IDs= split /[\(|\:|\)|\s+|\/]/,$lpp;$AL=@IDs;my $FFA=$FA;
     $Lipid=$IDs[0];$CC=$IDs[1];$DB=$IDs[2];my @Cc;my @Db;my $jj=-1;
       #print "$lpp and $AL and $CC\n";
        for($i=0;$i<=$AL;$i++)
                 {$sick=$IDs[$i];#print"Yes:$lpp,$AL and $sick\n";
                  if ($sick=~m/FA/)
              {$jj=$jj+1;$Cc[$jj]=$IDs[$i+1];$Db[$jj]=$IDs[$i+2];}
                  }
        for($kk=0;$kk<=$jj;$kk++)
        {
         $c=$Cc[$kk];$b=$Db[$kk];
         if($c<$CC && $b<=$DB)
           {$FFA="FA $c:$b";
            $Cr=$CC-$c;$Br=$DB-$b;
            if($Lipid=~m/TAG/)
             {$tf="$Cr:$Br";$tag=tagFA($tf);}
            else {$tag="FA $Cr:$Br";}
            $FFA.="/$tag";

           }
        }#print"this :   $FFA";

    #$ret="$FFA\t$mz1\t$pis1\t$cnt1\n";
        return($FFA);
} #next line of second file

sub tagFA

{
        # this sub program is used to calculate fatty acids of TAGs
        # this is complex way, many FAs can be detected , therefore manually sort the data
        #need to improve further
    for($fi4=1;$fi4<$LF;$fi4++)
    {
        for($fi5=1;$fi5<$LF;$fi5++)
        {
                  $db=$FAtable[$fi4][$fi5];
                  if($tf=~m/$db/)
                   {
                      #print "$db and $tf\n";
                      $FAT.="FA $Fci[0][$fi5]:$Fbi[0][$fi5]/FA $Fcj[$fi4][0]:$Fbj[$fi4][0];";}
        }
    }
return($FAT);
}

sub SearchFA
```

```
{
 #subprogram that search for fatty acids in FA-list files
 my $RT=0;my $prt=0;$fabs=0;
  #print"NXT\n ";
 $fal="FA-list";open(FAL,$fal);@Fal=<FAL>;my @fasp;my $c1,my $c2;my $b1;my $b2;my @fasp2;
 @lipid= split /[\(|\/|\)]/,$value;#print"search2:$CTOT[$C]\n";
foreach $_(@lipid)
  {#print"$_\n";
   chomp $_; if($_=~m/FA/){@fasp=split
   /[\:|\s]/,$_;$fa=$fasp[0];$c1=$fasp[1];$b1=$fasp[2];$fabs=1;}
       foreach $faa(@Fal)
       {chomp $faa;
        @fasp2=split /\:/,$faa;$c2=$fasp2[0];$b2=$fasp2[1];#print "$fa\n";
        #print "$faa :: $rem\n";
          if( $fa=~m/FA/ && $c2==$c1 && $b2==$b1)
             {#print "yes\n";
              $RT++;$prt=$RT;
             }
       }

  }
if($fabs eq 0){$RT=10;}#print " RT:$RT\n";
return($RT);
}
```

```
##########################################
#               Pre-BioLayout.pl         #
##########################################


#!/usr/bin/perl
use List::MoreUtils qw/ uniq /;

 $infile="sta-lipids.txt";#input the name of file that consist lipid names (output from
 Lipid-compare.pl))
 open(FILE,$infile);
 @INF=<FILE>;shift @INF;$i=0;
 open OUT1,"+>","LF-INP";
 #open OUT3,"+>","Calculate-FA";
 #use this to calculate other fatty acid; the symbol "#" at the begining of line should be
 removed to get output

 open OUT2,"+>","LC-INP";
 #open OUT3,"+>","Intlst";
 #print OUT1 "LIPID\tIden-F\tUnIden-F\n";

$ln=@INF;$mn=0;
for($in=0;$in<$ln;$in++)
{
 chomp $INF[$in];$linen=$INF[$in];$jn=0;
 @colsn=split /[\(|\)]/,$linen;$lipidNMn="$colsn[0]";$f1=$colsn[1];$sn=0;
  @fasp=split /[\-|\/|\+]/,$f1;$s=0;my $F1;
  foreach $fa(@fasp){if($fa=~m/FA/){$s++;if($s>1){$F1.="/";}$F1.=$fa;}}
 if($colsn[0]=~m/\+/){@lmm=split /\+/,$colsn[0];$lipidNMn="$lmm[0]";}#."$lmm[1]";}
    if($F1=~m/FA/)
    {
        foreach $line2n(@INF)
        {
          chomp $line2n;
          @cols2n=split /[\(|\)|]/,$line2n;$lipidNM2n="$cols2n[0]";$f2=$cols2n[1];
                  @fasp2=split /[\-|\/|\+]/,$f2;$s2=0;my $F2;
          if($cols2n[0]=~m/\+/){@lmm2=split /\+/,$cols2n[0];$lipidNM2n="$lmm2[0]";}#."$lmm2[1]";}
          if($f2=~m/FA/)
          {#print "$lipidNM and $lipidNM2\n";
             if($lipidNM2n=~m/$lipidNMn/)
              {
               foreach $fa2(@fasp2)
                {
                 if($fa2=~m/FA/)
                                        {$s2++;
                                         if($s2>1){$F2.="/";}
                                         $F2.=$fa2;
                                        }
                                    }
              $Lip[$mn]=$lipidNMn;
              if($sn>=1){$fac[$mn].="/";}
              $fac[$mn].="$F2";$sn++;$in=$jn;$Exp[$mn]="$Lip[$mn]>$fac[$mn]";
             }
```

```
          }$jn++;
        }
    $mn++;}

}
for($wwn=0;$wwn<$mn;$wwn++)
{
#print  "$Lip[$wwn] ($fac[$wwn])\n";
@LPD=split /[\s+|\:]/,$Lip[$wwn];
$LCT=$LPD[1];$LBT=$LPD[2];
@SPF=split /\//,$fac[$wwn];
my @unique = uniq @SPF;$sb=0;my $FAT;my $FcT;my $FbT;
    foreach $fat(@unique)
    {#print "$fat\t";
    @Cb=split
    /[\s+|\:]/,$fat;$FAC[$wwn][$sb]=$Cb[1];$FAB[$wwn][$sb]=$Cb[2];#$FTT="FA($Cb[1]:$Cb[2])";
    #print OUT1 "$Lip[$wwn] ($fat) , $FAC[$wwn][$sb],$FAB[$wwn][$sb]\n";
    $FcT+=$Cb[1];$FbT+=$Cb[2];$FAA[$wwn].="$fat/";
    $sb++;
    }$T=$FAA[$wwn];#print "$FAT,$FcT,$FbT\n";
#print "\n";
if($LCT>=$FcT)
    {
        if($LCT==$FcT)
        {
        if($LBT==$FbT){$Carry=$T;goto Next;}#print "LCT & LBT equal for $Lip[$wwn]\n";
        if($LBT>$FbT || $LBT<$FbT){$Carry=SPLIT($LBT,$LCT,$T);goto Next;}#print "Wrong FAs
        checking individual $Lip[$wwn]\t";
        }
        if($LCT>$FcT)
        {
            if($LBT==$FbT){$RC=$LCT-$FcT;$RB=0;}#print "Double bonds equal but carbons more for
            $Lip[$wwn]\t";
            if($LBT>$FbT){$RC=$LCT-$FcT;$RB=$LBT-$FbT;}#print "Double bonds and carbons more
            for $Lip[$wwn]\t";
            if($LBT<$FbT){$Carry=SPLIT($LBT,$LCT,$T);goto Next;}#print "Wrong FAs checking
            individual $Lip[$wwn]\t";
            $R="FA($RC:$RB)";
            $searche=SEARCH($R);
            if($searche!=1)
            {$Carry=SPLIT($LBT,$LCT,$T);}#print "Wrong FAs calculated checking individual
            $Lip[$wwn]\t";
            else{$Carry="$T FA $RC:$RB";}#print "Finally go it $Carry\n";

        }
    }
    if($FcT>$LCT){$Carry=SPLIT($LBT,$LCT,$T);goto Next;}#print "Wrong FAs checking individual
    $Lip[$wwn]\t";
      Next:

      @FINALFAA = split /\//,$Carry;$Fl=0;#print "$Lip[$wwn]>$Carry\n";
      my @uniq = uniq @FINALFAA;
```

```
    foreach $e(@uniq)
    {print OUT1"$Lip[$wwn]\t$e\n";print OUT2"$Lip[$wwn]\n";#print OUT3"$Lip[$wwn]/";
        $FATS[$wwn][$Fl]=$uniq[$Fl];
        $Fl++;
    }if($Lip[$wwn]!~m/$Lip[$wwn+1]/){if($Carry=~m/[A-Z]/){$tot[$wwn]="$Lip[$wwn]\t$Carry";print
    OUT3"$tot[$wwn]\n";}}
}

sub SPLIT
{
#subprogram for data processing
my $Remb;my $Remb1;my @cc ; my @bb;my @SPLT;
@Faty=split /\//,$T;$x=0;
foreach $Fy(@Faty)
{#print ">$Fy\n";
@SPLT=split /[\s+|\:]/,$Fy;
$cc[$x]=$SPLT[1];$bb[$x]=$SPLT[2];
$x++;
}#print "nof:$x\n";
for($y=0;$y<=$x;$y++)
{my $Rmc;my $Rmb;my $Remb1=$Faty[$y];
$Rmc=$LCT-$cc[$y];$Rmb=$LBT-$bb[$y];
$R="FA($Rmc:$Rmb)";$R1="FA $Rmc:$Rmb";#print "calcul:$R\nsearch\n";
if($Rmc=~m/0/ && $Rmb=~m/0/)
{$Remb.="$Remb1/";goto n2;}#print "equal:$Remb";
        $searche=SEARCH($R);
        if($searche!=1)
        {if($x<=1)
            {@si=split /\s+/,$Remb1;$R="$si[0]($si[1])";$searche=SEARCH($R);#print"$R\n";
             if($searche!=1){undef $Remb1;}
             else{$Remb.="$Remb1/";undef $R1;}
            }
         else {undef $Remb1;}
        }#print "IN SUBSTR :Wrong FA detected Discard\t";
        else{$Remb1.="/$R1";$Remb.="$Remb1/";}#print "GOT\n";
n2:

}
#print "OUT:$Remb\n";
return ($Remb);
}
sub SEARCH
{#sub program to search for fatty acids in fatty acid list file
 my $RT=0;
 #print"NXT\n ";
 $fal="falist";open(FAL,$fal);@Fal=<FAL>;my @fasp;my $c1,my $c2;my $b1;my $b2;my @fasp2;
 @fasp=split /[\(|\:|\)]/,$R;$c1=$fasp[1];$b1=$fasp[2];
 foreach $faa(@Fal)
 {chomp $faa;
  @fasp2=split /[\(|\:|\)]/,$faa;$c2=$fasp2[1];$b2=$fasp2[2];#print "$c1:$b1,$c2:$b2\n";
  #print "$faa :: $rem\n";
  if($c2==$c1 && $b2==$b1)
    {#print "yes\n";
```

```
      $RT=1;goto l;
    }
 }l:
 #print "RT:$RT\n";
 return($RT);
}
```

```
##########################################
#           BioLayout-Lipid-FA.pl        #
##########################################


#!/usr/bin/perl
use List::Util qw/max/;
use List::MoreUtils qw/ uniq /;

print "Input your file name \n";
$inp="Log-LF-INP";#<STDIN>;
chomp $inp;$infile="$inp";#print "$infile\n";
open(SIF,$infile);@SIF=<SIF>;
open OUT1,"+>","LAYFA";
open OUT2,"+>","Log-LAYOUT";
print OUT2"//BioLayout  Express 3D  Version 2.0 Layout  File\n";
my $i=0;my $k=0;$h=0;
foreach $line(@SIF)
{chomp $line;
@temp=split /\s+/,$line;$ID[$i]="$temp[0]";$i++;$ID[$i]="$temp[2] $temp[3]";
print OUT2"$temp[0]\t$temp[2] $temp[3]\n";
print OUT1"$temp[2] $temp[3]\n";
$Fatl[$i]="$temp[2] $temp[3]";
$i++;
}
close OUT1;
#my @unique1 = uniq @Fatl;
my %uniq1;
map { $uniq1{$_}=1} @Fatl;
@unique1 = keys %uniq1;
my %uniq;
map { $uniq{$_}=1} @ID;
@unique = keys %uniq;
$a=0;
open(LAY,"LAYFA");@Lay=<LAY>;$FK=0;my @SIZ;my @FLS;
    foreach $la(@unique1)
    {chomp $la;$in=0;@v1=split /[\(|\:|\)|\s]/,$la;
          if($la=~m/FA/)
        {
        foreach $li(@Lay)
        {chomp $li;@v2=split /[\(|\:|\)|\s]/,$li;if($v1[1]==$v2[1] &&
        $v1[2]==$v2[2]){$in++;}#print">>LA:$v2[1],$v1[2]\n";
        }}
    if($la=~m/FA/){$FLS[$FK]="$la,$in";$SIZ[$FK]=$in;}#print "$la,$in>>\n";
    $FK++;
    }
$ma= max @SIZ;
foreach $lipid(@unique)
{
@class=split /[\(|\)|\s]/,$lipid;$CLS=$class[0];#print "$CLS\n";
@noofcarbons=split /\:/,$class[1];$Size=$noofcarbons[0];$dbz=$noofcarbons[1];
chomp $lipid;#print "$lipid\n";
        if($CLS=~m/FA/)
```

```
        {#print"yesssssssssssss:$CLS.$Size.$dbz\n";
            foreach $ss(@FLS)
            {chomp $ss;@z=split /[\:|\s|\,]/,$ss;$SZ=$z[3];#print"<$ss>??$Size:$dbz ::
            $z[1]:$z[3] \n";
                if($Size=~m/$z[1]/ && $dbz=~m/$z[2]/)
                {#print"$ss:$SZ:$ma\n";
                $Size=(($SZ/$ma)*30);#print"renew:$Size\n";
                    if($Size<=10){$Size=10;}
                print OUT2"//NODECLASS\t$lipid\t$CLS-FA\n";
                print OUT2"//NODESIZE\t$lipid\t$Size\n";
                print OUT2"//NODESHAPE\t$lipid\t0\t0\n";
                print OUT2"//NODEALPHA\t$lipid\t1\n";
                }
            }goto n;
        }
        else
        {$HASH[$h]=$lipid;$SS=0;$KA=0;
            foreach $_(@SIF)
            {chomp $_;@temp=split /\s+/,$_;
             if($lipid=~m/$temp[0]/)
               {$HBIN[$h].="$temp[2] $temp[3],";$KA++;}
            }
        @LTT=split /\,/,$HBIN[$h]; my %uniq2;map { $uniq2{$_}=1} @LTT;@unique2 = keys
        %uniq2;$Size=@unique2;print "$lipid($KA,$Size) =$HBIN[$h]\n\n";
        if($Size<=10){$Size=12;}
        print OUT2"//NODECLASS\t$lipid\t$CLS-LIPID\n";
        print OUT2"//NODESIZE\t$lipid\t$Size\n";
        print OUT2"//NODESHAPE\t$lipid\t0\t0\n";
        print OUT2"//NODEALPHA\t$lipid\t1\n";$h++;
        }

#$px=int($Size/$ma)*10;#print "$Size\n";
#if($Size>10 && $Size<=20){$Size=5;}elsif($Size<=10){$Size=3;}
#if($Size>20 && $Size<=30){$Size=6;}
#if($Size>30 && $Size<=40){$Size=7;}
#if($Size>40 && $Size<=50){$Size=8;}
#if($Size>50){$Size=9;}

n:
}
print OUT2"//CURRENTCLASSSET\tLIPID\n";
print OUT2"//EDGESIZE\t0.4\n";
print OUT2"//EDGECOLOR\t#999999\n";
#print OUT"//NODECLASSCOLOR\tNo Class\t#000090\n";
print OUT2"//NODECLASSCOLOR\tDAG-LIPID\t#ff3333\n";
print OUT2"//NODECLASSCOLOR\tDGDG-LIPID\t#ff3333\n";
print OUT2"//NODECLASSCOLOR\tTAG-LIPID\t#ff3333\n";
print OUT2"//NODECLASSCOLOR\tLPG-LIPID\t#ff3333\n";
print OUT2"//NODECLASSCOLOR\tLPI-LIPID\t#ff3333\n";
print OUT2"//NODECLASSCOLOR\tSQDG-LIPID\t#ff3333\n";
print OUT2"//NODECLASSCOLOR\tPS-LIPID\t#ff3333\n";
print OUT2"//NODECLASSCOLOR\tPI-LIPID\t#ffff33\n";
print OUT2"//NODECLASSCOLOR\tPG-LIPID\t#ffff33\n";
```

```
print OUT2"//NODECLASSCOLOR\tPE-LIPID\t#ffff33\n";
print OUT2"//NODECLASSCOLOR\tPC-LIPID\t#ffff33\n";
print OUT2"//NODECLASSCOLOR\tMGDG-LIPID\t#ffff33\n";
print OUT2"//NODECLASSCOLOR\tFA-LIPID\t#ffff00\n";
print OUT2"//DEFAULTSEARCH\tgoogle";
```

```
##########################################################################################################
##########################################
#
BioLayout-CC.pl                                                                                          #
##########################################################################################################
##########################################

#!/usr/bin/perl
use List::Util qw/max/;
use List::MoreUtils qw/ uniq /;

$inp="LC-INP";#<STDIN>;
chomp $inp;$infile="$inp";print "$infile\n";
open(SIF,$infile);@SIF=<SIF>;
#open OUT1,"+>","LAY1";
open OUT2,"+>","Log-CC-LAYOUT";
print OUT2"//BioLayout  Express 3D  Version 2.0 Layout  File\n";
my $i=0;my $k=0;
foreach $line(@SIF)
{chomp $line;
@temp=split /\s+/,$line;$ID[$i]="$temp[0]";$IA[$i]="$temp[0]";$i++;$ID[$i]="$temp[1]";
print OUT2"$temp[0]\t$temp[1]\n";
#print OUT1"$temp[0]($temp[1])\n";
$Fatl[$i]="$temp[0]($temp[1])";
$i++;
}

#my @unique1 = uniq @Fatl;
my %uniq1;
map { $uniq1{$_}=1} @Fatl;
@unique1 = keys %uniq1;
close OUT1;
my %uniq2;
map { $uniq2{$_}=1} @ID;
@unique2 = keys %uniq2;
close OUT1;
my %uniq;
map { $uniq{$_}=1} @IA;
@unique = keys %uniq;
$a=0;
#open(LAY,"LAY1");@Lay=<LAY>;$FK=0;my @SIZ;my @FLS;
foreach $li(@unique)
{chomp $li;$in=0;   #print"lip:$li\t";
    foreach $la(@unique1)
    {chomp $la;@v1=split /[\(|\:|\)]/,$la;#print "$la\n";
        if($v1[0]=~m/$li/){$in++;}#print">>LA:$la/LI:$li,$in\n";
    }
    if($li=~m/[A-Z]/){$FLS[$FK]="$li,$in";$SIZ[$FK]=$in;print"$li,$in\n";}
    $FK++;
}
$ma= max @SIZ;
foreach $lipid(@unique2)
{chomp $lipid;#print "1>$lipid\n";
```

```
#@class=split /[\(|\)]/,$lipid;$CLS=$class[0];$CLS=$class[1];#print "$CLS\n";
@noofcarbons=split /\:/,$lipid;$Size=$noofcarbons[0];$dbz=$noofcarbons[1];
chomp $lipid;#print "$lipid\n";
        if($lipid=~m/[A-Z]/)
            {#print "2>$lipid\n";
            foreach $ss(@FLS)
            {chomp $ss;@z=split /\,/,$ss;$SZ=$z[1];#print"??$Size:$dbz :: $z[1]:$z[2] \n";
                if($z[0]=~m/$lipid/)
                {#print"$SZ/$ma\n";
                $Size=(($SZ/$ma)*20);#print"adjussiz:$Size\n";
                    if($Size<=10){$Size=10;}
                print OUT2"//NODECLASS\t$lipid\t$lipid-LIPID\n";
                print OUT2"//NODESIZE\t$lipid\t$Size\n";
                print OUT2"//NODESHAPE\t$lipid\t0\t0\n";
                print OUT2"//NODEALPHA\t$lipid\t1\n";
                }
            }goto n;
            }
#$px=int($Size/$ma)*10;#print "$Size\n";
if($Size>10 && $Size<=20){$Size=5;}elsif($Size<=10){$Size=3;}
if($Size>20 && $Size<=30){$Size=6;}
if($Size>30 && $Size<=40){$Size=7;}
if($Size>40 && $Size<=50){$Size=8;}
if($Size>50){$Size=9;}
print OUT2"//NODECLASS\t$lipid\t-CC\n";
print OUT2"//NODESIZE\t$lipid\t$Size\n";
print OUT2"//NODESHAPE\t$lipid\t0\t0\n";
print OUT2"//NODEALPHA\t$lipid\t1\n";
n:
}
print OUT2"//CURRENTCLASSSET\tLIPID\n";
print OUT2"//EDGESIZE\t0.4\n";
print OUT2"//EDGECOLOR\t#999999\n";
#print OUT"//NODECLASSCOLOR\tNo Class\t#000090\n";
print OUT2"//NODECLASSCOLOR\tDGDG-LIPID\t#ff3333\n";
print OUT2"//NODECLASSCOLOR\tTAG-LIPID\t#ff3333\n";
print OUT2"//NODECLASSCOLOR\tLPG-LIPID\t#ff3333\n";
print OUT2"//NODECLASSCOLOR\tLPI-LIPID\t#ff3333\n";
print OUT2"//NODECLASSCOLOR\tSQDG-LIPID\t#ff3333\n";
print OUT2"//NODECLASSCOLOR\tPS-LIPID\t#ff3333\n";
print OUT2"//NODECLASSCOLOR\tPI-LIPID\t#ffff33\n";
print OUT2"//NODECLASSCOLOR\tPG-LIPID\t#ffff33\n";
print OUT2"//NODECLASSCOLOR\tPE-LIPID\t#ffff33\n";
print OUT2"//NODECLASSCOLOR\tPC-LIPID\t#ffff33\n";
print OUT2"//NODECLASSCOLOR\tMGDG-LIPID\t#ffff33\n";
print OUT2"//NODECLASSCOLOR\tFA-LIPID\t#ffff00\n";
print OUT2"//DEFAULTSEARCH\tgoogle";
##########################################################################################
#########################################
```