University of Southampton

Faculty of Natural and Environmental Sciences

Graduate School of the National Oceanography Centre, Southampton

Phytoplankton Lipidomics:

Lipid Dynamics in Response to Microalgal Stressors.

by

Jonathan Eliott Hunter

Thesis for the Degree of Doctor of Philosophy

Supervision: Prof. Rachel A. Mills; Prof. George S. Attard;
Prof. Anthony Postle; Prof. C. Mark Moore.

October 2015

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCE

Ocean and Earth Sciences

Thesis for the Degree of Doctor of Philosophy

PHYTOPLANKTON LIPIDOMICS: LIPID DYNAMICS IN RESPONSE TO MICROALGAL STRESSORS

by Jonathan Eliott Hunter

Phytoplankton growth is sustained by the supply of essential nutrients and balanced by mortality processes such as viral infection, both of which can give rise to stress. Remodelling of cellular lipids in response to such stresses is common in unicellular organisms.

Under phosphorus (P) stress, phytoplankton substitute glycerophospholipids with non-phosphorus analogues, reducing their demand for P. Reported herein, the model marine diatom *Thalassiosira pseudonana* degraded only a small proportion of its original glycerophospholipid. Most of the P remained incorporated in glycerophospholipids, but significant changes in the individual glycerolipid species were observed.

Untargeted lipidomic screening highlighted diglycosylceramides, not previously observed in *T. pseudonana*, that increase with P stress and may be useful as biomarkers. The fatty acids comprising each individual diglyceride lipid were characterised filling a conspicuous gap in our knowledge. Preliminary results suggest partitioning of diacylglycerol lipids between subcellular compartments.

Marine diatoms, rich in lipids such as triacylglycerols are potential feedstocks for biofuels, where nitrogen (N) starvation is common to increase lipid yield. Quantification of individual glycerolipid species under N stress revealed that polyunsaturated glycerophosphatidylcholine species and the predominant chemotype of sulfoquinovosyldiacylglycerol displayed large increases. Total diacylglycerol increased 3 fold under N stress, comprised of increases in saturated/monounsaturated species. This appears to form part of the cell's adaption to N limitation that ultimately leads to the accumulation of triacylglycerides.

These findings provide insight on the diatom lipidic response to nutrient stress and their adaptations to life in low nutrient environments, with additional implications upon biofuel production.

Marine viruses infect phytoplankton influencing host ecology and evolution. *Emiliania huxleyi* has a biphasic life cycle with a diploid and haploid phase. Diploid cells are susceptible to infection by specific coccolithoviruses, yet haploid cells are resistant. Analysis of lipids from cultures of uninfected diploid, infected diploid and uninfected haploid *E. huxleyi* revealed that sialic-glycosphingolipid, previously linked with susceptibility to infection, was absent from the resistant haploid cultures. Additional untargeted analyses unveiled potential biomarkers furthering our understanding of *E. huxleyi* host/virus lipid dynamics and highlight potential novel biomarkers for infection, susceptibility and ploidy.

Contents

List of Figures

List of Tables

Declaration of Authorship

I, Jonathan Eliott Hunter declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Phytoplankton Lipidomics: Lipid Dynamics in Response to Microalgal Stressors.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:

   Hunter, J. E., Frada, M. J., Fredricks, H. F., Vardi, A., and Van Mooy, B. A. S. (2015). Targeted and Untargeted Lipidomics of *Emiliania huxleyi* Viral Infection and Life Cycle Phases Highlights Molecular Biomarkers of Infection, Susceptibility, and Ploidy. *Front. Mar. Sci.*

Signed:

Date:

Acknowledgements

I acknowledge the following: friends and mentors George Attard, Rachel Mills, Tony Postle and Mark Moore for your invaluable advice and encouragement before, during and beyond my Ph. D. research. This has allowed me to make a difficult interdisciplinary challenge into an extremely rewarding experience. To George and Rachel, a special thank you for your years of support (via climbing rope). My advisory panel chairs, Eric Achterberg and Toby Tyrrell.

The mass spectrometry groups in Medicine and Chemistry and staff in Ocean and Earth Science, including Grielof Koster, Victoria Goss, John Langley, Julie Herniman, John Gittins, Sophie Richier, Mark Stinchcombe and Chris Daniels whose methodological assistance, borrowed equipment, support and stimulating discussion throughout has contributed immensely to the data presented within.

My collaborators at the Woods Hole Oceanographic Institution (WHOI, USA), Helen Fredricks, Bethanie Edwards, Jeremy Tagliafere, Justin Ossolinski, James Fulton. Thank you for welcoming me into your group during my exchange visit, resulting in a scientifically extremely rewarding visit and a formative personal experience. To Miguel Frada and Assaf Vardi at the Weizmann Institute (Israel) for their collaboration on the Haploid/Diploid project.

Benjamin Van Mooy (WHOI) and Joost Brandsma (Southampton) in particular, for sustantial input and assistance despite having no formal obligation to do so. Friends and colleagues in Chemistry, Marcus Dymond, Stephanie Tweed, Richard Gillams, Richard Wilson and Duncan Parker for instruction, discussion and friendship throughout my studies and the undergraduate project students that I enjoyed assisting and whom contributed to this research: Samantha Earl, Jamie Burrell, Jonathan Watson.

Jo and Barbara, for your support and advice over the last few years. To my family, John, Joanne and Amy, for their endless and unconditional support. Finally, to Annie, for your constant support, encouragement and understanding throughout everything this thesis has entailed.

This research was financially supported by: The University of Southampton - Vice Chancellors Scholarship Award; Graduate School of the National Oceanography Centre WHOI Exchange Award; the Gordon and Betty Moore Foundation (Grant GBMF3301) and the Wellcome Trust (Grant 057405).

Abbreviations

1N	Haploid
2N	Diploid
[PO_4^{3-}]	Orthophosphate Ion Concentration
AIF	All Ion Fragmentation
AutoMS2	Automatic (Data Dependant) Tandem Mass Spectrometry
BLL	Betaine Like Lipid
Cer	Ceramide
CID	Collision Induced Dissociation
DAG	Diacylglycerol
DGCC	Diacylglycerylcarboxy-N-hydroxymethyl-choline
DGTA	Diacylglycerylhydroxymethyltrimethyl-β-alanine
DGTS	Diacylglyceryltrimethylhomoserine
DGDG	Digalactosyldiacylglycerol
dH_2O	Distilled Water
E. huxleyi	*Emiliania huxleyi*
EhV	*Emiliania huxleyi* Virus
ESI	Electrospray Ionisation
FA	Fatty Acid
Gly-	Glycosyl-
GSL	Glycosphingolipid
HAc	Acetate Adduct
hGSL	Host glycosphingolipid
HPLC	High Performance Liquid Chromatography
IP-DAG	Intact Polar Diacylglycerol
IPL	Intact Polar Lipid
LCB	Long Chain (Sphingoid) Base
L-	Lyso-

MS	Mass Spectrometry
MS2	Tandem Mass Spectrometry
M/Z	Mass / Charge Ratio
MAG	Monoacylglycerol
MGDG	Monogalactosyldiacylglycerol
N	Nitrogen
NL###	Neutral Loss of ### MS/MS Scan
P	Phosphorus
POP	Particulate Organic Phosphorus
P###	Precursor Ion of ### MS/MS Scan
PC	Glycerophosphatidylcholine
PE	Glycerophosphatidylethanolamine
PG	Glycerophosphatidylglycerol
PUFA	Polyunsaturated Fatty Acid
q-ToF	Quadrupole Time of Flight Mass Spectrometry
S.D.	Standard Deviation
sGSL	Sialic glycosphingolipid
SQDG	Sulfoquinovosyldiacylglycerol
T. pseudonana	*Thalassiosira Pseudonana*
TAG	Triacylglycerol
UPLC	Ultra Performance Liquid Chromatography
vGSL	Viral glycosphingolipid
w.r.t.	With Respect To

Chapter 1

Introduction

Organisms akin to modern cyanobacteria can be found in fossil records extending back 3.45 x 10^9 years (**Cavalier-Smith**, 1993; **Falkowski et al.**, 1998). Molecular phylogenetics indicate that via endosymbiosis these prokaryotes gave rise to eukaryotic phytoplankton (**Falkowski et al.**, 1998; **Archibald**, 2009). Phytoplankton drive marine biogeochemistry in the contemporary ocean and contribute nearly half of the biospheric net primary production (**Field et al.**, 1998).

On average, phytoplankton biomass in the oceans turns over around once per week (**Falkowski et al.**, 1998). Growth is sustained by the supply of essential nutrients and light and balanced in the steady state by mortality by grazing, viral infection, autocatalyzed cell death or sinking (**Falkowski et al.**, 1998; **Moore et al.**, 2013). All of these factors can potentially give rise to environmental stresses upon phytoplankton, with subsequent implications for their biology.

This thesis presents investigations of the effects of a number of environmental stresses upon the phytoplankton lipidome (the entirety of the cellular lipids) via lipidomics. Lipidomics is the study of the totality of lipids, the molecules with which they interact and their cellular function (**Watson**, 2006). In particular, phosphorus (P) nutrient stress, nitrogen (N) nutrient stress and viral infection were investigated and the environmental context of each is presented with the thesis structure outline in Section 1.5 . This introduction provides relevant background information and context as well as an overview of

the thesis content and its significance. More detailed literature reviews directly relevant to each chapter are provided in the associated chapter introductions.

1.1 The Lipids of Marine Phytoplankton and Their Importance

Given a vast biochemical diversity of lipids, remodelling of a cell's lipidome in response to environmental stresses is common in unicellular organisms (**Benning et al.**, 1995; **Brandsma**, 2011). In recent decades it has begun to become apparent that lipids function not only as energy depots and structural components but have diverse and complex roles in membrane function and signalling (**Mouritsen**, 2011). Thus the quantitative study of lipids via lipidomics has far reaching biological implications including membrane biophysics, metabolism and inter/intra-cellular signalling (**Mouritsen**, 2011). The lipid metabolism of triacylglycerides in particular represents a field of intense research in order to maximise phytoplankton biodiesel yields (**Hu et al.**, 2008; **Griffiths and Harrison**, 2009; **Sakthivel**, 2011; **Khozin-Goldberg and Cohen**, 2011; **Cagliari et al.**, 2011).

Beyond the direct biological implications of lipids, they bear important secondary implications in an environmental science context. Firstly, membrane lipids comprise 15 to 25% of the carbon in planktonic cells forming a significant component of the biogeochemical cycle of carbon (**Wakeham et al.**, 1997; **Van Mooy et al.**, 2008). Therefore, understanding the nature of this carbon source is of great importance (**Wakeham et al.**, 1997).

Secondly, phosphorus containing membrane phospholipids consume between 18 and 28% of planktonic PO_4^{3-} uptake (**Van Mooy et al.**, 2006), but it is also well known that plankton can substitute phospholipids for non-phosphorus alternatives (**Benning**, 1998; **Van Mooy et al.**, 2006, 2009). This drives deviations from Redfield (C:N:P = 106:16:1) stoichiometry (**Redfield**, 1934) with great importance to macronutrient biogeochemistry in a broad sense (**Fanning**, 1992; **Geider and La Roche**, 2002; **Van Mooy et al.**, 2009; **Moore et al.**, 2013).

Thirdly, lipids derived from phytoplankton are astonishing in their diversity and many bear distinctive structures. Some of these distinctive species are indicative of particular organisms or biogeochemical processes both in the present and/or the past and are termed biomarkers (**Volkman**, 2006). These can be loosely divided into two groups:

Esterified, intact lipids are generally considered to indicate extant, viable microbial communities on account of their lability after cell death (**Brandsma et al.**, 2012). Intact polar lipids such as the glycerophospholipids are rapidly hydrolysed, losing the head group, to yield the more stable 'core' lipid (**Schouten et al.**, 2010). Radio-labelled glycerophosphatidylethanolamine (PE), for example, was reported to be degraded by 70% after 4 days (**Harvey et al.**, 1986; **Schouten et al.**, 2010). This description may be too simplistic however, and evidence suggests that some intact polar lipids, in particular those bearing glycosidic linkages, may be stable for much longer time periods (**Schouten et al.**, 2010). Intact polar lipids have been investigated for a range of purposes including chemotaxonomy (**Brandsma et al.**, 2012), as a biomarker for viral infection (**Vardi et al.**, 2009), nutrient availability (**Van Mooy et al.**, 2009) and microbial biomass (**White et al.**, 1993).

Post mortem, the cellular particulate organic matter (POM) may be subject to a variety of processes in the water column and sediments that could lead to chemical modification of any potential lipid biomarker (**Volkman and Tanoue**, 2002). These may include processes such as heterotrophic microbial respiration, zooplankton digestion, hydrolysis etc. (**Volkman and Tanoue**, 2002). Derivatives of lipid biomarkers that are chemically stable to these processes and present in the sedimentary record on geological timescales are termed palaeoceanographic proxies. These are generally defined by a resilience to degradation processes before, during and after sedimentation. Lipid derived palaeoceanographic proxies constitute a diverse and active field of research (**Volkman et al.**, 1998; **Volkman**, 2006; **Rosell-Melé and McClymont**, 2007; **Eglinton and Eglinton**, 2008).

Notable examples of palaeoceanographic proxies include the long established U^{K}_{37} index. U^{K}_{37} is used for the reconsruction of past sea surface temperatures based upon the ratio of unsaturation of alkenones, derived from coccolithophores (**Brassell et al.**, 1986; **Prahl and Wakeham**, 1987; **Eltgroth et al.**, 2005; **Eglinton and Eglinton**, 2008).

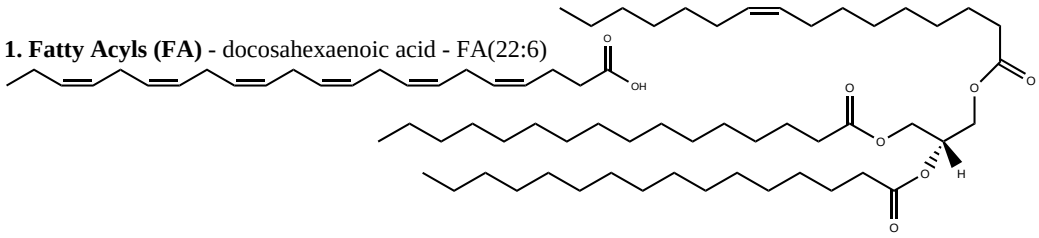
$$U^{K}_{37} = \frac{[C_{37:2}] - [C_{37:4}]}{[C_{37:2} + C_{37:3} + C_{37:4}]} \tag{1.1}$$

The more recent TEX_{86} index is based upon glycerol dialkyl glycerol tetraether (GDGT)

lipids from the *Thaumarchaeota* group of archaea. TEX_{86} is a ratio of GDGT species with differing numbers of cyclic moieties and is a complimentary sea surface temparture proxy to U^{K}_{37} (**Schouten et al.**, 2002; **Eglinton and Eglinton**, 2008).

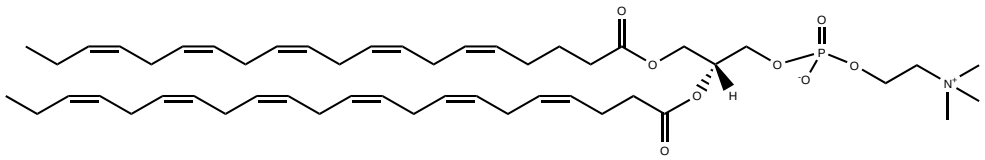
1.2 Lipid Biochemistry

Lipids are classified as "hydrophobic or amphipathic (bearing both polar and apolar moieties) small molecules" (**Fahy et al.**, 2005). They result from two biosynthetic pathways: by condensation of thioesters and/or isoprene units. This definition encompasses a large chemical space (**Liebisch et al.**, 2013), which is subdivided into broad categories as shown in Figure 1.1.

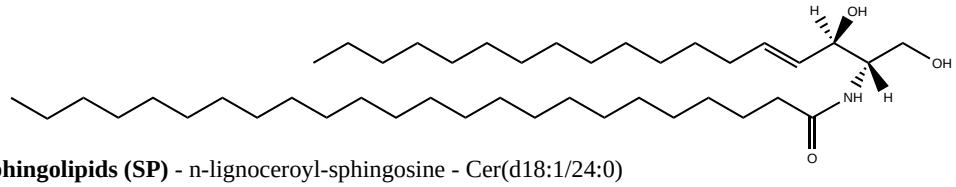


1. Fatty Acyls (FA) - docosahexaenoic acid - FA(22:6)

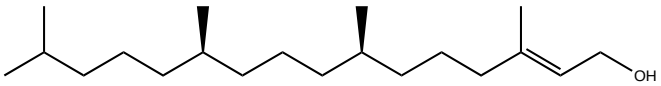
2. Glycerolipids (GL) - 1,2-dipalmitoyl-3-palmitoleoyl-*sn*-glycerol - TAG(16:0/16:0/16:1)



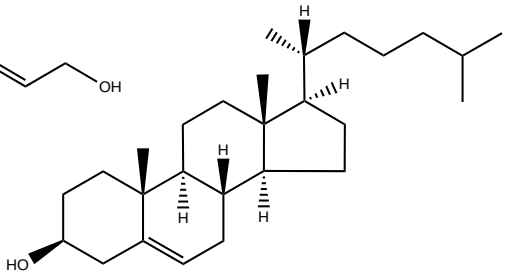
3. Glycerophospholipids (PL) -
1-docosahexaenoyl-2-eicosapentaenoylglycerolphosphatidylcholine - PC(22:6/20:5)



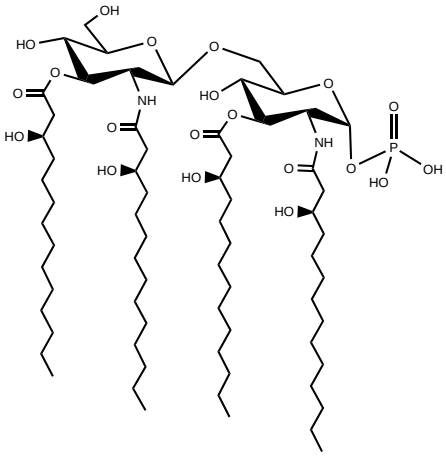
4. Sphingolipids (SP) - n-lignoceroyl-sphingosine - Cer(d18:1/24:0)



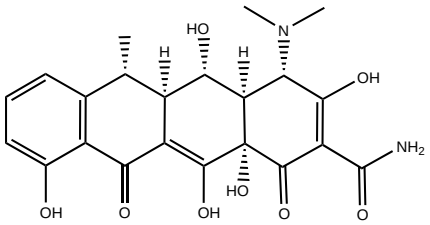
5. Prenol Lipids (PR) - Phytol



6. Sterol Lipids - Cholesterol



7. Saccharolipids (SL) - Lipid A disaccharide-1-phosphate



8. Polyketides (PK) - Doxycycline

Figure 1.1: Representative structures from each of the lipid categories as defined by the Lipid Metabolites and Pathways Strategy (LIPID MAPS) consortium - adapted from references (**Fahy et al.**, 2005, 2009; **Blanksby and Mitchell**, 2010; **Harkewicz and Dennis**, 2011).

Of these categories, the fatty acyls, glycerolipids and glycerophospholipids make up the Intact Polar Diacylglycerols (IP-DAGs) a.k.a. the Intact Polar Lipids (IPLs). They represent the most abundant type of lipid molecules in the surface oceans (**Van Mooy and Fredricks**, 2010). In the model marine diatom *Thalassiosira Pseudonana*, IPLs comprise >80% of the total lipids (**Volkman et al.**, 1989). The IPLs are associated with the cellular membranes (**Van Mooy et al.**, 2009) and are themselves subdivided into IPL classes as illustrated in Figure 1.2.

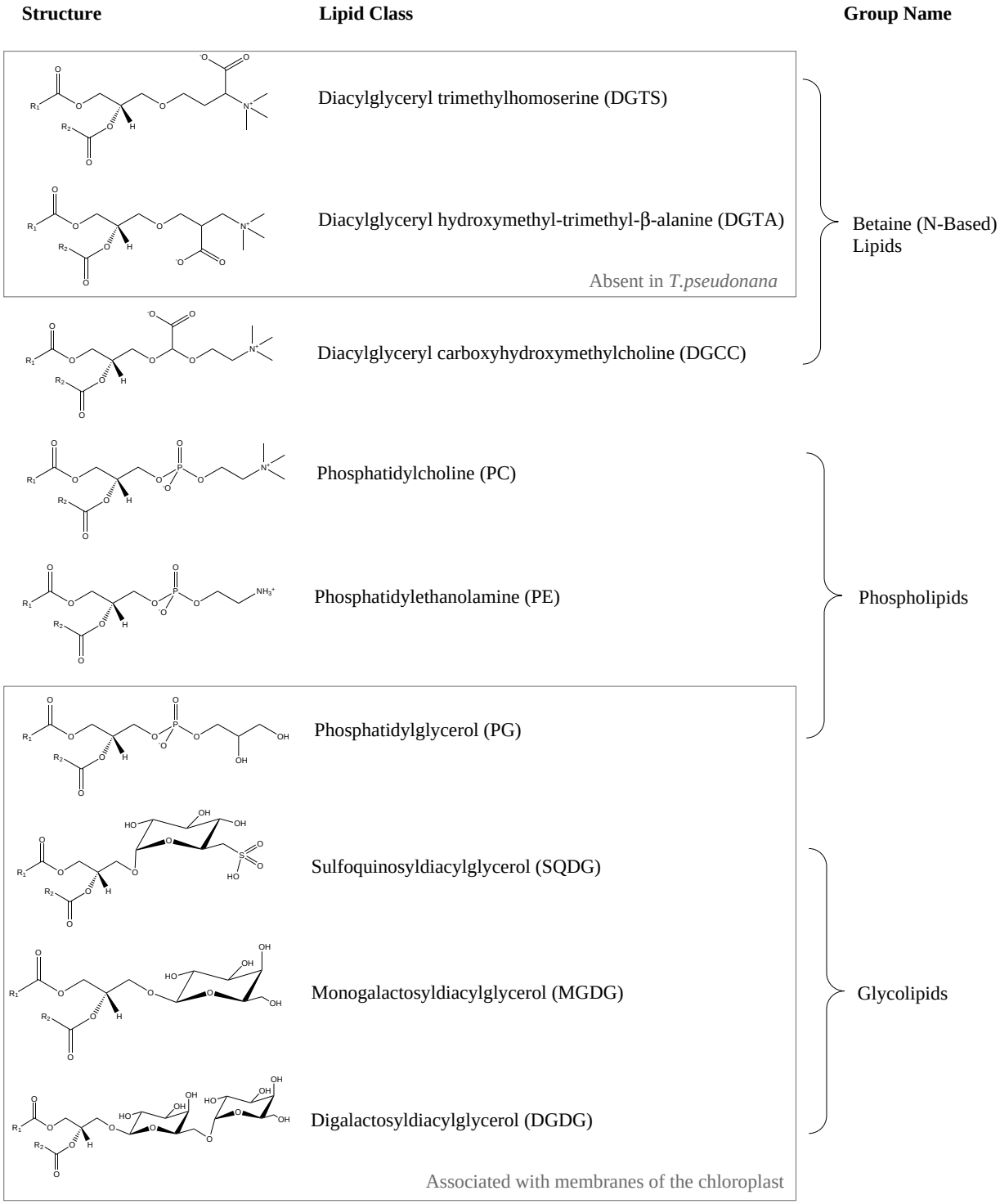


Figure 1.2: The intact polar lipid classes present in *T.pseudonana* and the marine environment. R_1 and R_2 represent fatty acid groups - adapted from references (**Siegenthaler and Murata**, 2004; **Van Mooy et al.**, 2009; **Van Mooy and Fredricks**, 2010).

Each IPL contains a chiral glycerol backbone with three substituents. These classes are denoted based on the nature of the head group at sn-3. Typically, sn-1 and sn-2 bear long chain fatty acyl groups, but may have alkyl ether or 1Z-alkenyl ether (plasmalogen) linked chains. Collectively, these functional groups are denoted radyl. The structure and shorthand notation for description of IPLs is shown in Figure 1.3 (**Fahy et al.**, 2005; **Hölzl and Dörmann**, 2007).

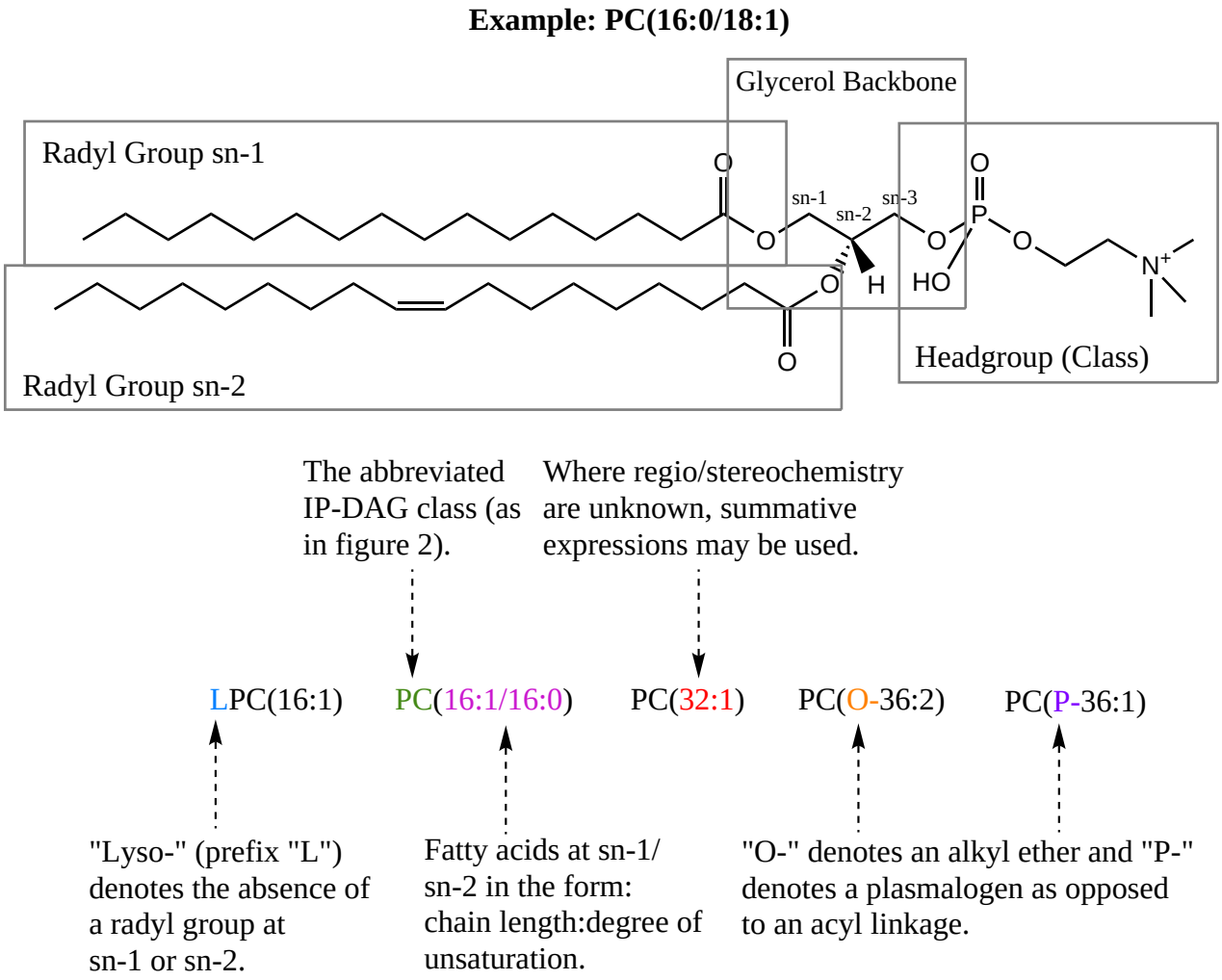


Figure 1.3: IPL structural components and notation - adapted from references (**Fahy et al.**, 2005, 2009).

Beyond the structural IPLs, other types of lipids are important in phytoplankton biology. Many eukaryotic phytoplankton produce substantial quantities (20-50% dry weight) of triacylglycerol (TAG) storage lipids (Figure 1.2 Panel 2) under adverse environmental conditions (**Hu et al.**, 2008; **Sakthivel**, 2011; **Hildebrand et al.**, 2012; **Liu and Benning**, 2013). TAG nomenclature is as presented as for the IPLs, with the exception of an additional fatty acyl group at sn-3 (**Fahy et al.**, 2005).

The sphingolipids are another diverse and complex class of lipids that serve both as structural components of membranes and in a variety of molecular signalling roles (**Haynes et al.**, 2009; **Vardi et al.**, 2009; **Pata et al.**, 2010; **Arakaki et al.**, 2012). Sphingolipids are comprised of a sphingoid (a.k.a long-chain) base with a terminal hydroxyl and an amine moiety at the 2-position. An amide bound fatty acid may link to the amine group of the sphingoid base forming a ceramide (Figure 1.1 Panel 4). Like fatty acids, the sphingoid base may vary in chain length and degree of unsaturation, and commonly bears one or two additional hydroxyl groups (denoted "d" (di-hydroxy) or "t" (tri-hydroxy)). Finally, the sphingolipid may bear a variety of headgroups (such as carbohydrate or phosphatidyl moieties) attached by the terminal hydroxyl, giving rise to a wide variety of sphingolipid species (**Fahy et al.**, 2005; **Haynes et al.**, 2009).

Lipid Metabolism in Eukaryotic Phytoplankton

The majority of our understanding of lipid metabolism in photosynthetic organisms is derived from research on the higher plant model *Arabidopsis thaliana* (**Li-Beisson et al.**, 2010; **Liu and Benning**, 2013). *A. thaliana*, a mustard green and common weed, is the premiere higher plant model organism, due in part to its relatively concise genome, short generation time, small size and prolific self-pollination (**Koornneef and Meinke**, 2010).

Investigations into eukaryotic phytoplankton lipid metabolism have been made, most of which focus upon the green alga *Chlamydomonas reinhardtii* (**Guschina and Harwood**, 2006; **Harwood and Guschina**, 2009). Fatty acids comprise the fundamental building blocks of most of the phytoplankton lipidome and are incorporated in glycerolipids via Acyl-Coenzyme A and Acyl-Acyl Carrier Protein (Figure 1.4, (**Riekhof et al.**, 2005)).

Palmitic acid (16:0) is synthesised de novo in the eukaryotic chloroplast and can then be subject to a series of elongase and desaturase enzymatic processes leading to longer and more unsaturated fatty acids (**Harwood and Guschina**, 2009; **Khozin-Goldberg and Cohen**, 2011). Eukaryotic phytoplankton are rich in long chain (>18C) polyunsaturated fatty acids (PUFA) and commonly >20% of the total fatty acids are comprised by eicosapentaenoic (20:5, EPA) and docosahexaenoic (22:6, DHA) acid (**Harwood and Guschina**, 2009).

Gene annotation in *Chlamydomonas* indicates a broadly comparable, but somewhat simpler glycerolipid metabolic network than in higher plants (**Riekhof et al.**, 2005; **Liu and Benning**, 2013). Glycerolipid biosynthesis involves cooperation between the plastid and extraplastidal compartments with the participation of enzymes localised to the endoplasmic reticulum and chloroplast envelope (**Khozin-Goldberg and Cohen**, 2011). A schematic of the glycerolipid biosynthetic network, based upon known and hypothesised pathways is presented in Figure 1.4. Conversely, several lipid biomarkers are known to be specific to vascular plants and are well established in the study of terrestrial organic material input into marine systems (**Hedges et al.**, 1997) such as plant-wax *n*-alkanes (**Brassell and Eglinton**, 1983) and three-ring diterpenoids (**Simoneit**, 1977; **Hedges et al.**, 1997).

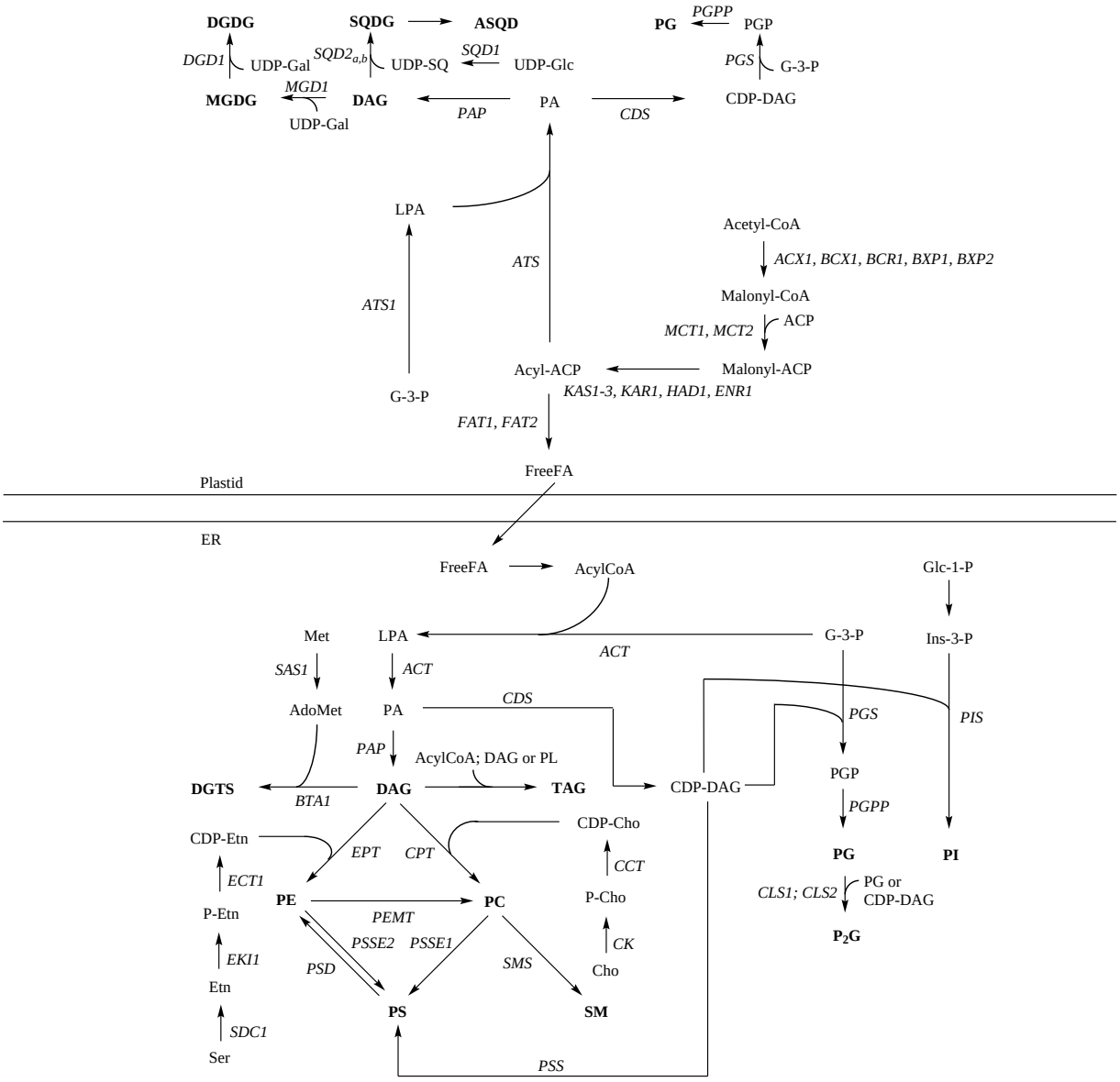


Figure 1.4: Schematic of known and hypothesised glycerolipid biosynthesis pathways in eukaryotic phytoplankton. Endoplasmic Reticulum (ER). Adapted from references (**Sato**, 1988; **Vogel and Eichenberger**, 1992; **Riekhof et al.**, 2005; **Lykidis**, 2007; **Liu and Benning**, 2013).

Abbreviation	Designation
Lipid Products	
ASQD	2'-O- acyl-sulfoquinovosyldiacylglycerol
DAG	Diacylglycerol
DAG	Diacylglycerol
DGDG	Digalactosyldiacylglycerol
DGTS	Diacylglyceryl-N,N,N-trimethylhomoserine
MGDG	Monogalactosyldiacylglycerol
P2G	Cardiolipin
PC	Phosphatidylcholine
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
PI	Phosphatidylinositol
PS	Phosphatidylserine
SM	Sphingomyelin
SQDG	Sulfoquinovosyldiacylglycerol
TAG	Triacylglycerol
Intermediates	
Acetyl-CoA	Acetyl-Coenzyme A
Acyl-ACP	Acyl Carrier Protein
Acyl-CoA	Acyl-Coenzyme A
AdoMet	S-adenosylmethionine
CDP-Cho	Cytidine-5'-diphosphate-choline
CDP-DAG	Cytidine-5'-diphosphate-diacylglycerol
CDP-Etn	Cytidine-5'-diphosphate-ethanolamine
Cho	Choline
Etn	Ethanolamine
FreeFA	Free Fatty Acid
G-3-P	Glycerol-3-phosphate
Glc-1-P	Glucose-1-phosphate
Ins-3-P	Inositol-3-phosphate
LPA	Lysophosphatidic acid
Malonyl-ACP	Malonyl-Acyl Carrier Protein
Malonyl-CoA	Malonyl-Coenzyme A
Met	Methionine
PA	Phosphatidic Acid
P-Cho	Phosphocholine
P-Etn	Phosphoethanolamine
PGP	Phosphatidylglycerophosphate
Ser	Serine
UDP-Gal	Uridine -5'-diphosphate-galactose
UDP-Glc	Uridine-5'diphosphate-glucose
UDP-SQ	Uridine -5'-diphosphate-sulfoquinovose
Enzymes	
ACT	Acyltransferase(s)
ACX	Alpha-carboxyltransferase
ATS	G-3-P:acyl-ACP acyltransferase
BCR	Biotin carboxylase

BCX	Beta-carboxyltransferase
BXP	Biotin carboxy carrier protein
CCT	CTP:choline cytidyltransferase
CDS	CDP-DAG synthetase
CK	Choline Kinase
CPT	CDP-choline:DAGphosphotransferase
DGD	DGDG synthase
ECT	Phosphoethanolamine cytidylyltransferase
EKI	Ethanolamine kinase
ENR	Enoyl-ACP-reductase
EPT	CDP-ethanolamine:DAG-ethanolamine phosphotransferase
FAT	Acyl-ACP thiolase
HAD	3-hydroxyacyl-ACP dehydratase
KAR	3-ketoacyl-ACP reductase
KAS	3-ketoacyl-ACP synthase
MCT	Malonyl-CoA:ACP transacylase
MGD	MGDG synthase
PAP	PA phosphatase
PEMT	Phosphatidylethanolamine methyltransferase
PGPP	Phosphatidylglycerophosphate phosphatase
PGS	Phosphatidylglycerophosphate synthase
PIS	CDP-DAG:inositol phosphotransferase
PSD	Phosphatidylserine decarboxylase
PSS	PS synthase
PSSE	Phosphatidylserine synthase via base-exchange
SAS	Adomet synthetase
SDC	Serine decarboxylase
SMS	SM synthase
SQD1	UDP-sulfoquinovose synthase
SQD2	Sulfolipid synthase

Table 1.1: Abbreviations of lipids, intermediates and enzymes in Figure 1.4 - adapted from references (**Sato**, 1988; **Vogel and Eichenberger**, 1992; **Riekhof et al.**, 2005; **Lykidis**, 2007; **Liu and Benning**, 2013).

Perhaps the most pronounced difference between eukaryotic phytoplankton and plants with regards to lipid biosynthesis is that the plankton produce betaine lipids (**Liu and Benning**, 2013). In *Chlamydomonas* the sole betaine lipid produced is diacylglyceryl-N,N,N-trimethyl-homoserine (DGTS) (**Riekhof et al.**, 2005). In the model marine diatom *T. pseudonana* however, 1,2-Diacylglyceryl-3-O-carboxyhydroxymethylcholine (DGCC) constitutes the only betaine lipid (**Van Mooy et al.**, 2009). The biosynthetic pathway leading to DGCC however, remains a conspicuous gap in our knowledge and nothing appears known other than the reported incorporation of labelled 14C-methionine into its headgroup (**Kato et al.**, 1996, 2003).

1.3 Methods for Lipid Analysis

Historically, lipids have been analysed via a variety of chromatographic-based separation methods such as 1D or 2D thin layer chromatography (**Touchstone**, 1995) or high performance liquid chromatography (HPLC) with assorted detectors (**Picchioni et al.**, 1996; **Hummel et al.**, 2011). Prior to analysis, intact lipids are extracted from cell isolates, most commonly by the method of (**Bligh and Dyer**, 1959). This involves disruption of the cells and dissolution/purification in a chloroform/methanol azeotrope. A biphasic mixutre is formed by the addition of water and the lipid extract is then isolated by way of the chloroform phase. This method remains in use today subject to minor modifications for the extraction of marine plankton (**Sturt et al.**, 2004; **Popendorf et al.**, 2013).

Whilst extraction methods have remained largely unchanged, the analysis of the prepared lipid extracts has advanced considerably with the application and progression of mass spectrometry technologies (**Blanksby and Mitchell**, 2010):

Electron ionisation gas chromatography MS (EI-GC-MS) has been a mainstay in the identification of low molecular weight lipids such as fatty acids and sterols for more than 40 years (**Christie**, 1989; **Blanksby and Mitchell**, 2010). This method is not amenable to the analysis of larger (and less volatile) complex lipids such as the glycerolipids. However, their fatty acids can be analysed by prior derivatisation via transesterifcation to fatty acid methyl esters (FAMES), at the cost of a significant loss of structural information (**Christie**, 1989; **Blanksby and Mitchell**, 2010).

Early analyses of complex lipids was achieved by fast atom bombardment (FAB) MS leading to much of our fundamental understanding of fragmentation in complex lipids (**Jensen et al.**, 1987; **Blanksby and Mitchell**, 2010). This method is severely limited however, by low overall sensitivity, presence of matrix ions and major in-source fragmentation. These limitations were overcome with the implementation of electrospray ionisation (ESI) for glycerophospholipid analysis in the mid 1990s (**Han and Gross**, 1994; **Blanksby and Mitchell**, 2010).

This gave rise to two main approaches for lipidomic analysis, involving the infusion of

sample following separation by high-performance liquid chromatography (HPLC-MS) or via direct infusion a.k.a. "shotgun lipidomics" (**Blanksby and Mitchell**, 2010; **Han and Gross**, 2005). As ESI yields minimal in-source fragmentation it is therefore amenable to tandem MS/MS (a.k.a MS2) approaches (multiple mass-analysers or ion-trap mass spectrometers) (**Han et al.**, 2011). Typical ESI-MS/MS experiments are outlined in figures 1.5 and 1.6.

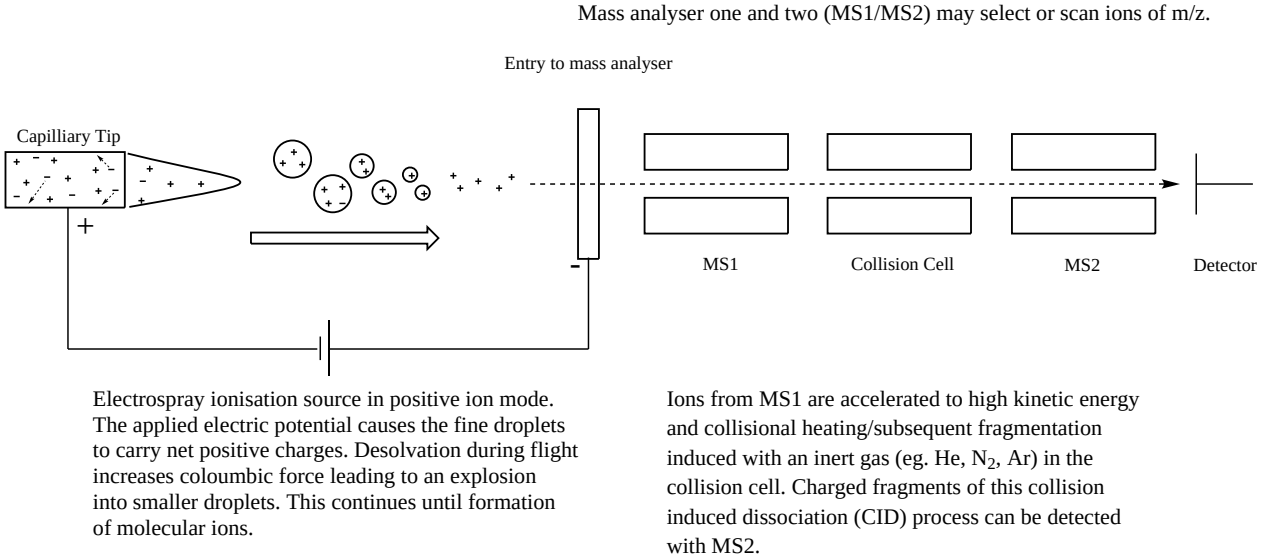


Figure 1.5: Schematic diagram of the principles of ESI-MS/MS adapted from references (**Han and Gross**, 2005; **Han et al.**, 2011).

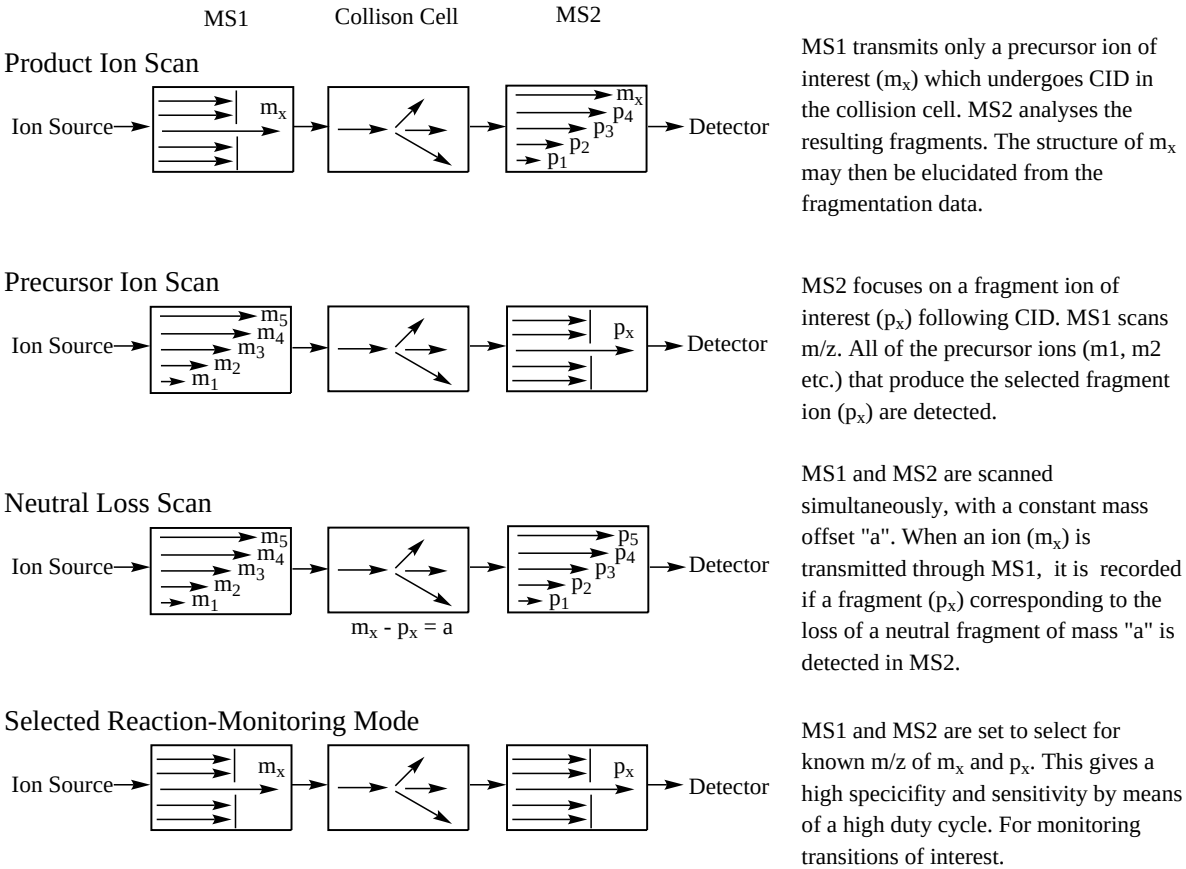


Figure 1.6: Tandem MS/MS modes for lipidomic analysis - adapted from reference (**Han et al.**, 2011).

Some of the most recent technologies include ultra performance liquid chromatography (UPLC) (**Plumb et al.**, 2006) and ion mobility mass spectrometry (IMS) (**Kanu et al.**, 2008). These and high resolution time of flight (ToF)(**Ekroos et al.**, 2002), orbitrap (**Zubarev and Makarov**, 2013) or the older Fourier transform ion cyclotron resonance (FT-ICR-MS) (**Marshall et al.**, 1998) spectrometers have driven the sensitivity and resolution of measurements ever further.

Lipidomics is broadly concerned with the identification and quantification of all lipid species within a given organism or sample. The perfect method would achieve this in a single experiment, but to date no single analytical technology has achieved this outcome (**Hummel et al.**, 2011). Whilst the discussed advances in mass spectrometric technology

have allowed researchers to identify and quantify up to a few hundred lipid species (**Ejsing**, 2007), it is likely that this number represents but a fraction of the full biological lipid diversity (**Blanksby and Mitchell**, 2010; **Dennis et al.**, 2010). Thus, detection and identification of minor lipid species represents one of the major challenges in phytoplankton lipidomics. As these challenges are overcome by technological and methodological advances, the difficulty of quantification of the identified lipid species becomes apparent.

Standards and Quantification

Quantification in mass spectrometry is generally by two approaches, relative or absolute. Relative quantification measures pattern change whereas absolute quantification is concerned with the determination of mole quantities of a given analyte (**Yang and Han**, 2011).

Relative quantification approaches are limited by the comparisons that one can make. The ion intensity observed via mass spectrometry is determined not only by quantity, but also a response factor governed primarily by the ionisation efficiency of a given analyte (**Yang and Han**, 2011). Ionisation efficiency varies with the electrical propensity of an individual analyte, subject to its microenvironment, to gain or lose charge in an electric field. The quantity of ions produced, and hence signal recorded, is therefore influenced by the chemistry of the lipid headgroups and to a lesser extent, the radyl groups (**Han and Gross**, 2005).

As such relative comparisons are quantitative only between molecules of similar ionisation efficiency, for example polar glycerolipids bearing the same headgroup (**Koivusalo et al.**, 2001), within the same matrix such as a directly infused total lipid extract (**Yang and Han**, 2011). In order to quantify in absolute terms, it is required to measure a standard of known quantity and use that measurement to account for the response factor. This can employ external standard methodology, where a calibration curve is established at a series of concentrations. Alternatively, internal standard methodology can be used, where a non-physiological standard is spiked into the sample during extraction (**Yang and Han**, 2011).

The availability of these standards is where the first major issue arises. Two approaches to the preparation of authentic lipid standards are common. Firstly, an individual lipid species may be prepared by chemical synthesis such that it does not overlap with endogenous lipids. This approach is favourable owing to the simplicity and accuracy resulting for being processed and analysed simultaneously via the internal standard approach (**Yang and Han**, 2011). Secondly, lipid external standards can be purified via preparative HPLC (**Ellingson and Zimmerman**, 1987). These are typically utilised via external standard methodology such that there is no concern of potential overlap with endogenous lipids (**Yang and Han**, 2011). In practice, the time and expertise required to effect either process is substantial. As such, appropriate standards for common mammalian lipids such as the glycerophospholipids are readily available as synthetic products from companies such as Avanti Polar Lipids (Alabaster, USA). Many plant-specific lipids however, are only available as purified natural products or not available at all (**Wada and Murata**, 2009).

In Chapter 2, absolute quantification of the betaine lipid DGCC was an important outcome. Achieving this outcome by way of direct infusion methodology proved difficult. Initially, we requested custom synthesis outfit ChemOvation (Horsham, UK) to synthesize an appropriate internal standard. Unfortunately the synthesis failed and we resorted to the application of a purified extract of DGCC provided by collaborator Benjamin Van Mooy at Woods Hole Oceanographic Institution. Whilst this extract overlapped heavily with the DGCC chemotypes present in the organism of interest, *T. pseudonana*, it was fortuitous that *T. pseudonana* is entirely devoid of DGCC during P replete culture (**Van Mooy et al.**, 2009). This allowed us to account for the matrix conditions by addition of the DGCC standard to DGCC free *T. pseudonana* and generate an external standard calibration method for the quantification of DGCC by direct infusion.

1.4 The Study of Lipids in Cultures of Model Organisms

This thesis presents results showing the effects of environmental stresses upon the lipids of batch cultured model phytoplankton species. A culturing approach allows stringent control of all growth parameters and is therefore amenable to direct comparison and statistical analysis, as opposed to in-situ environmental sampling. Two eukaryotic unicellular phytoplankton models were investigated, the marine diatom *T. pseudonana* and the marine coccolithophore *Emiliania huxleyi* (*E. huxleyi*).

Diatoms are eukaryotic microalgae within the phylum heterokontophyta. Characteristically encased in ornate silica frustules, this diverse and populous class represents a significant component of the oceanic phytoplankton and yield up to 25% of the global annual primary production (**Hoek et al.**, 1995). Diatoms therefore represent a significant source of lipids within the marine environment (**Van Mooy and Fredricks**, 2010).

T. pseudonana is a marine diatom used as a model species in many experimental studies as a representative of the diverse and widespread marine genus *Thalassiosira*. A marine strain of *T.pseudonana* (Hustedt) Hasle et Heimdal (CCMP1335/3H), isolated from Moriches Bay, Long Island, New York, U.S.A. became the first diatom to have its entire genome published (**Armbrust et al.**, 2004; **Alverson et al.**, 2011). This established it as a leading model for the study of diatom biogeochemistry. However, it is noted that recent phylogenetic analyses of CCMP1335 have shown that *T. pseudonana* likely descended from a freshwater ancestor. This may serve to confound any extrapolation to the biology of strictly marine species (**Alverson et al.**, 2011).

E. huxleyi is the numerically dominant coccolithophore in the modern oceans and an important component of phytoplankton assemblages, inhabiting all but extreme polar oceans. Moreover, it forms large, dense blooms in high-latitude coastal and shelf ecosystems that exert a critical impact upon the global carbon cycle and the earths climate (**Westbroek et al.**, 1993; **Paasche**, 2001; **Tyrrell and Merico**, 2004). Coccolithophores yield up to 20% of total carbon fixation in some systems (**Poulton et al.**, 2007) and like *T. pseudonana*, a reference genome is available (**Read et al.**, 2013).

1.5 Thesis Structure

This thesis is comprised of four results chapters which are presented as original research manuscripts (Chapter 2 - 5). Each describes the mass spectrometric investigation of the phytoplankton lipidome subject to environmental stressors, focusing on P and N nutrient stress and viral infection. Chapter 6 provides a summary of findings, with a synthesis consideration of the chapters in combination and an outlook for future research. The scientific background of each chapter is outlined briefly in the following sections and in greater detail within each chapter itself.

1.5.1 Chapter 2: Lipid Remodelling by the Diatom *Thalassiosira pseudonana* under Phosphorus Stress – New Insights into Phospholipid Substitution

Phosphorus is essential for the growth of all organisms. For phytoplankton, phosphate (PO_4^{3-}) is the favoured source of dissolved phosphorus, in some cases supplemented with dissolved organic phosphorus (**Paytan and McLaughlin**, 2007; **Van Mooy et al.**, 2009). The Sargasso Sea ($[PO_4^{3-}] = <10 nmol\ L^{-1}$) is an example region of very low P (**Van Mooy et al.**, 2009). The eastern Mediterranean is also severely depleted of PO_4^{3-}. Addition of phosphorus alone does not stimulate growth however as NP co-limitation seems to occur (**Moore et al.**, 2013). This is likely attributable to the increased bioavailability of organic phosphorus compounds, acclimatization mechanisms to phosphorus stress and selection for higher N:P stoichiometry under low nutrient conditions (**Moore et al.**, 2013).

When subjected to P stress phytoplankton must reduce their physiological P demands. In phytoplankton, this is commonly achieved by the substitution of glycerophospholipids with non phosphorus alternatives (**Benning et al.**, 1995; **Van Mooy et al.**, 2009; **Martin et al.**, 2011). This substitution can be observed in situ by phosphate uptake, in the P scarce Sargasso Sea, just $1.3 \pm 0.6\%$ PO_4^{3-} was used for phospholipid synthesis. In comparison, plankton of the relatively P replete South Pacific subtropical gyre used $17 \pm 6\%$ (**Van Mooy et al.**, 2009).

Substitution of the phospholipid PG with the sulfolipid SQDG in phytoplankton under conditions of P scarcity was first observed in 1993 (**Benning et al.**, 1993). This has since been demonstrated to dramatically reduce the phosphate requirement of picocyanobacteria in the oligotropic ocean (**Van Mooy et al.**, 2006). The marine diatom *T. pseudonana* has been used as a model eukaryotic organism to study the effects of P starvation on lipid remodelling (**Martin et al.**, 2011). When grown under low P conditions it synthesises the nitrogen based betaine lipid DGCC as a substitute for the predominant glycerophospholipid PC (**Van Mooy et al.**, 2009). DGCC is normally undetectable under P replete conditions (**Van Mooy et al.**, 2009). In addition to the shift from PC to DGCC lipids in P-starved *T. pseudonana*, a comparable shift reportedly takes place from PG to SQDG (**Van Mooy**

et al., 2009; **Martin et al.**, 2011).

The dynamics of glycerophospholipid substitution in *T. pseudonana* as a consequence of P stress have been studied quantitatively in terms of the major polar headgroup classes (**Martin et al.**, 2011). However, little is known about how the lipid species within each class, varying in fatty acids, change in response to P stress.

Chapter 2 describes the use of direct infusion ESI-MS/MS to quantify each of the polar glycerolipid classes as a whole and the individual species within each class, for *T. pseudonana* cultured under P replete (P+) and (P-) conditions over a period of 96 h. Our findings shed new light on the dynamics of phospholipid substitution at the individual lipid species level, with implications upon lipid substitution as a means for *T. pseudonana* to cope with P stress. Furthermore, the biosynthetic relationship between the various polar lipid types is discussed and potential candidate biomarkers of P stress are highlighted.

1.5.2 Chapter 3: Untargeted Lipidomic Characterisation of the Marine Diatom *Thalassiosira Pseudonana* Subject to Phosphorus Stress - Increase in Diglycosylceramides under Low Phosphorus Conditions

Mass spectrometry based lipidomics often employs two distinct strategies, targeted lipidomics for the quantification of a subset of lipids that are commonly already known, and untargeted (or global) lipidomics which aims to characterise all of the lipids in a system (**Han**, 2009; **Blanksby and Mitchell**, 2010).

T. pseudonana has been used as a model system to study glycerophospholipid substitution dynamics at the level of the total within each of the major glycerolipid classes (**Martin et al.**, 2011). The dynamics of substitution at the level of the individual lipid species within each lipid glycerolipid class, varying in fatty acid substituents, are reported in Chapter 2. This previous work consists exclusively of targeted lipidomic analyses and therefore the behaviour of minor lipid species under P stressed conditions, outside of the major glycerolipid classes, remains unknown. Furthermore, previous reports have not characterised the fatty acid composition of each individual diglyceride lipid species.

Chapter 3 presents the untargeted analysis of the lipidome of *T. pseudonana*, subject to P stress, by ultra performance liquid chromatography - quadrupole time of flight mass spectrometry (UPLC-Q-ToF-MS). Minor lipid species indicative of P stress, including a group of diglycosylceramides, not previously reported in *T. pseudonana* are presented. Furthermore, semi-targeted data dependant MS2 analyses were used to characterise the fatty acid composition of the major diglyceride lipids. Finally, a preliminary analysis of a purified plastid fraction yielded insight into the partitioning of different DAG chemotypes between the plastid fraction and the whole cell.

1.5.3 Chapter 4: Lipid Remodelling by the Diatom *Thalassiosira pseudonana* under Nitrogen Stress – Diglyceride Lipid Dynamics and Relation to Triglyeride Production

Nitrogen, like phosphorus, is an essential elemental nutrient for the growth of all organisms. It is bioavailable to phytoplankton as dissolved nitrate (NO_3^-), nitrite (NO_2^-), ammonium (NH_4^+) and organic forms (**Tyrrell**, 1999; **Zehr and Ward**, 2002). Nitrogen availability limits the growth of phytoplankton, such as marine diatoms, across much of the low-latitude surface ocean (**Hoek et al.**, 1995; **Moore et al.**, 2013). As such, understanding the physiological responses and adaptions of marine diatoms to nitrogen scarcity is important in determining the ecological response under nitrate limitation and impact on global biogeochemical cycles including carbon dioxide drawdown.

Marine diatoms have also received attention as potential feedstocks for biofuel production. This interest results from the ecological success of diatoms and the high proportion of their dry weight contributed by lipids, in particular the major storage lipids triacylglycerides (TAGs) (**Levitan et al.**, 2014).

Much is known about the response of the *T. pseudonana* lipidome to N stress (**Tonon et al.**, 2002, 2005; **Yu et al.**, 2009; **Bromke et al.**, 2013). However, absolute quantification (*i.e.* moles $cell^{-1}$) and between lipid species comparison (relative abundance) have not been investigated. Furthermore, glycerolipids containing polyunsaturated fatty acids such as 20:5 and 22:6 were not previously characterised (**Bromke et al.**, 2013), yet previous research shows the importance of these species, in the glycerolipidome ((**Tonon et al.**, 2005; **Yu et al.**, 2009) and Chapter 2). Finally, the fatty acid composition of each individual glycerolipid involved in N limitation has yet to be explicitly characterised (with the exception of the TAGs (**Yu et al.**, 2009)).

Chapter 4 presents a detailed quantitative study of the diglyceride lipids of *T. pseudonana* subject to N limitation by targeted lipidomics. We describe trends in the response of the diglyceride lipidome to nitrogen limitation and relate these observations to the well-developed literature on triglyceride lipids. These results further our knowledge of the diatom lipidome, its response to N limitation and TAG production. This knowledge bears

implications for the adaption of marine diatoms to N limited conditions throughout the worlds oceans and can inform further optimisation of TAG production efficiency for biofuel applications by genetic modification or other means.

1.5.4 Chapter 5: Targeted and Untargeted Lipidomics of *Emiliania huxleyi* Viral Infection and Life Cycle Phases Highlights Molecular Biomarkers of Infection, Susceptibility, and Ploidy

There are an estimated 10^{30} marine viruses, making them the most abundant biological agent in the oceans. Many of the 10^{23} viral infections that occur per second involve phytoplankton resulting in the death of approximately 20% of the total plankton per day. Marine viruses are therefore, major mortality agents of phytoplankton and result in substantial implications upon biogeochemical cycling (**Fuhrman**, 1999; **Suttle**, 2007; **Brussaard et al.**, 2008). Specific, giant, lytic double-stranded DNA *E. huxleyi* viruses (EhV), belonging to the phycodnaviridae family that infect microalgae (**Van Etten et al.**, 2002) are heavily implicated in the decay of *E. huxleyi* blooms (**Bratbak et al.**, 1993; **Brussaard et al.**, 1996; **Vardi et al.**, 2012).

At the cellular level, as a large dsDNA virus with high metabolic demand for the building blocks of DNA, lipids and protein synthesis, EhV triggers a rapid remodelling of the host metabolism (**Rosenwasser et al.**, 2014; **Schatz et al.**, 2014). In particular, recent studies have highlighted the crucial role of membrane lipids in the progression and regulation of EhV infection (**Vardi et al.**, 2009, 2012; **Rosenwasser et al.**, 2014).

A glycosphingolipid with a sialic acid modified glycosyl headgroup (sGSL) was recently described to have a direct relationship with susceptibility to infection in diploid (2N) cells (**Fulton et al.**, 2014). Interestingly, whilst the 2N form of *E. huxleyi* is generally susceptible to EhV infection, the haploid (1N) form appears completely resistant to EhV (**Frada et al.**, 2008, 2012).

Early targeted analyses have shown similar compositions with respect to the major structural lipids, storage lipids and pigments between 1N and 2N cells in the case of a single strain of *E. huxleyi* (**Bell and Pond**, 1996). Minor or novel lipid classes however, are as yet uncharacterised in the 1N cell lipidome and it is unknown whether the susceptibility marker sGSL is absent from EhV resistant 1N cells.

Chapter 5 presents a detailed characterisation of the lipidomes of cultured *E. huxleyi*: an uninfected 2N strain; the 2N strain under infection with coccolithovirus; and an un-

infected 1N strain. Total lipid extracts derived from these cultures collected over 120 h post-infection were characterised by mass spectrometry. We used two approaches, utilising both targeted analysis for quantification of known GSL/glycerolipid species and untargeted analyses for screening for unknown lipids.

1.5.5 Summary of Thesis Research Aims

- **Chapter 2:** Investigate the lipid dynamics of phospholipid substitution under P stress in the model marine diatom *T. pseudonana*, at the level of the individual lipid species within each lipid headgroup class.

- **Chapter 3:** Characterise the lipidome of *T. pseudonana* subject to P stress, via untargeted lipidomic methods, in order to highlight novel molecular biomarkers. Identify each fatty acid explicitly within each of the major individual lipid species. Gain preliminary insight into the partitioning of lipid species between the chloroplast and whole cell fractions.

- **Chapter 4:** Examine the polar glycerolipid dynamics of *T. pseudonana* subject to N limitation and relate these findings to the literature on the triglyceride response.

- **Chapter 5:** Conduct targeted and untargeted lipidomic analyses of cultured diploid, haploid and virus infected diploid *E. huxleyi*. Characterise the response of known glycosphingolipid biomarkers of infection and susceptibility and look for novel biomarker candidates indicative of infection, susceptibility and ploidy.

1.6 References

Alverson, A., Beszteri, B., Julius, M., and Theriot, E. (2011), The model marine diatom Thalassiosira pseudonana likely descended from a freshwater ancestor in the genus Cyclotella, *BMC Evolutionary Biology*

Arakaki, A., Iwama, D., Liang, Y., Murakami, N., Ishikura, M., Tanaka, T., et al. (2012), Glycosylceramides from marine green microalga Tetraselmis sp., *Phytochemistry*

Archibald, J. M. (2009), The Puzzle of Plastid Evolution, *Current Biology*

Armbrust, E. V., Berges, J. A., Bowler, C., Green, B. R., Martinez, D., Putnam, N. H., et al. (2004), The genome of the diatom Thalassiosira pseudonana: ecology, evolution, and metabolism., *Science*

Bell, M. V. and Pond, D. (1996), Lipid composition during growth of motile and coccolith forms of Emiliania huxleyi, *Phytochemistry*

Benning, C. (1998), Biosynthesis and Function of the Sulfolipid Sulfoquinovosyl Diacylglycerol, *Annual Review of Plant Physiology and Plant Molecular Biology*

Benning, C., Beatty, J. T., Prince, R. C., and Somerville, C. R. (1993), The sulfolipid sulfoquinovosyldiacylglycerol is not required for photosynthetic electron transport in Rhodobacter sphaeroides but enhances growth under phosphate limitation., *Proceedings of the National Academy of Sciences of the United States of America*

Benning, C., Huang, Z. H., and Gage, D. A. (1995), Accumulation of a novel glycolipid and a betaine lipid in cells of Rhodobacter sphaeroides grown under phosphate limitation., *Archives of Biochemistry and Biophysics*

Blanksby, S. J. and Mitchell, T. W. (2010), Advances in Mass Spectrometry for Lipidomics, *Annual Review of Analytical Chemistry*

Bligh, E. G. and Dyer, W. J. (1959), A Rapid Method of Total Lipid Extraction and Purification, *Canadian Journal of Biochemistry and Physiology*

Brandsma, J. (2011), The origin and fate of intact polar lipids in the marine environment (Ph. D. Thesis, Utrecht University)

Brandsma, J., Hopmans, E. C., Philippart, C. J. M., Veldhuis, M. J. W., Schouten, S., and Sinninghe Damsté, J. S. (2012), Low temporal variation in the intact polar lipid composition of North Sea coastal marine water reveals limited chemotaxonomic value, *Biogeosciences*

Brassell, S. and Eglinton, G. (1983), The Potential of Organic Geochemical Compounds as Sedimentary Indicators of Upwelling, in E. Suess and J. Thiede, eds., Coastal Upwelling Its Sediment Record SE - 27, volume 10B of *NATO Conference Series* (Springer US)

Brassell, S. C., Eglinton, G., Marlowe, I. T., Pflaumann, U., and Sarnthein, M. (1986), Molecular stratigraphy: a new tool for climatic assessment, *Nature*

Bratbak, G., Egge, J. K., and Heldal, M. (1993), Viral mortality of the marine alga Emiliania huxleyi (Haptophyceae) and termination of algal blooms, *Marine Ecology Progress Series*

Bromke, M. A., Giavalisco, P., Willmitzer, L., and Hesse, H. (2013), Metabolic Analysis of Adaptation to Short-Term Changes in Culture Conditions of the Marine Diatom Thalassiosira pseudonana, *PLoS ONE*

Brussaard, C., Gast, G., van Duyl, F., and Riegman, R. (1996), Impact of phytoplankton bloom magnitude on a pelagic microbial food web, *Marine Ecology Progress Series*

Brussaard, C. P. D., Wilhelm, S. W., Thingstad, F., Weinbauer, M. G., Bratbak, G., Heldal, M., et al. (2008), Global-scale processes with a nanoscale drive: the role of marine viruses, *ISME J*

Cagliari, A., Margis, R., dos Santos Maraschin, F., Turchetto-Zolet, A. C., Loss, G., and Margis-Pinheiro, M. (2011), Biosynthesis of Triacylglycerols (TAGs) in plants and algae, *International Journal of Plant Biology*

Cavalier-Smith, T. (1993), Kingdom protozoa and its 18 phyla., *Microbiological Reviews*

Christie, W. W. (1989), Gas chromatography and lipids, volume 39 (Oily Press Ayr)

Dennis, E. A., Deems, R. A., Harkewicz, R., Quehenberger, O., Brown, H. A., Milne, S. B., et al. (2010), A Mouse Macrophage Lipidome, *Journal of Biological Chemistry*

Eglinton, T. I. and Eglinton, G. (2008), Molecular proxies for paleoclimatology, *Earth and Planetary Science Letters*

Ejsing, C. S. (2007), Molecular characterisation of the lipidome by mass spectrometry (Ph. D. Thesis, Technische Universität Dresden)

Ekroos, K., Chernushevich, I. V., Simons, K., and Shevchenko, A. (2002), Quantitative Profiling of Phospholipids by Multiple Precursor Ion Scanning on a Hybrid Quadrupole Time-of-Flight Mass Spectrometer, *Analytical Chemistry*

Ellingson, J. S. and Zimmerman, R. L. (1987), Rapid separation of gram quantities of phospholipids from biological membranes by preparative high performance liquid chromatography., *Journal of Lipid Research*

Eltgroth, M. L., Watwood, R. L., and Wolfe, G. V. (2005), Production and Cellular Localization of Neutral Long-Chain Lipids in the Haptophyte Algae Isochrysis Galbana and Emiliania Huxleyi., *Journal of Phycology*

Fahy, E., Subramaniam, S., Brown, H. A., Glass, C. K., Merrill, A. H., Murphy, R. C., et al. (2005), A comprehensive classification system for lipids, *Journal of Lipid Research*

Fahy, E., Subramaniam, S., Murphy, R. C., Nishijima, M., Raetz, C. R. H., Shimizu, T., et al. (2009), Update of the LIPID MAPS comprehensive classification system for lipids, *Journal of Lipid Research*

Falkowski, P. G., Barber, R. T., and Smetacek, V. (1998), Biogeochemical Controls and Feedbacks on Ocean Primary Production, *Science*

Fanning, K. A. (1992), Nutrient provinces in the sea: Concentration ratios, reaction rate ratios, and ideal covariation, *Journal of Geophysical Research: Oceans*

Field, C. B., Behrenfeld, M. J., Randerson, J. T., and Falkowski, P. (1998), Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components, *Science*

Frada, M., Probert, I., Allen, M. J., Wilson, W. H., and de Vargas, C. (2008), The "Cheshire Cat" escape strategy of the coccolithophore Emiliania huxleyi in response to viral infection, *Proceedings of the National Academy of Sciences*

Frada, M. J., Bidle, K. D., Probert, I., and de Vargas, C. (2012), In situ survey of life cycle phases of the coccolithophore Emiliania huxleyi (Haptophyta), *Environmental Microbiology*

Fuhrman, J. A. (1999), Marine viruses and their biogeochemical and ecological effects., *Nature*

Fulton, J. M., Fredricks, H. F., Bidle, K. D., Vardi, A., Kendrick, B. J., DiTullio, G. R., et al. (2014), Novel molecular determinants of viral susceptibility and resistance in the lipidome of Emiliania huxleyi, *Environmental Microbiology*

Geider, R. and La Roche, J. (2002), Redfield revisited: variability of C:N:P in marine microalgae and its biochemical basis, *European Journal of Phycology*

Griffiths, M. and Harrison, S. (2009), Lipid productivity as a key characteristic for choosing algal species for biodiesel production, *Journal of Applied Phycology*

Guschina, I. A. and Harwood, J. L. (2006), Lipids and lipid metabolism in eukaryotic algae, *Progress in Lipid Research*

Han, X. (2009), Lipidomics: Developments and applications, *Journal of Chromatography. B, Analytical Technologies in the Biomedical and Life Sciences*

Han, X. and Gross, R. W. (1994), Electrospray ionization mass spectroscopic analysis of human erythrocyte plasma membrane phospholipids, *Proceedings of the National Academy of Sciences*

Han, X. and Gross, R. W. (2005), Shotgun lipidomics: Electrospray ionization mass spectrometric analysis and quantitation of cellular lipidomes directly from crude extracts of biological samples, *Mass Spectrometry Reviews*

Han, X., Yang, K., and Gross, R. W. (2011), Multi-dimensional mass spectrometry-based shotgun lipidomics and novel strategies for lipidomic analyses., *Mass Spectrometry Reviews*

Harkewicz, R. and Dennis, E. A. (2011), Applications of Mass Spectrometry to Lipids and Membranes, *Annual Review of Biochemistry*

Harvey, H., Fallon, R. D., and Patton, J. S. (1986), The effect of organic matter and oxygen on the degradation of bacterial membrane lipids in marine sediments, *Geochimica et Cosmochimica Acta*

Harwood, J. L. and Guschina, I. A. (2009), The versatility of algae and their lipid metabolism, *Biochimie*

Haynes, C. A., Allegood, J. C., Park, H., and Sullards, M. C. (2009), Sphingolipidomics: Methods for the Comprehensive Analysis Of Sphingolipids, *Journal of Chromatography. B, Analytical Technologies in the Biomedical and Life Sciences*

Hedges, J. I., Keil, R. G., and Benner, R. (1997), What happens to terrestrial organic matter in the ocean?, *Organic Geochemistry*

Hildebrand, M., Davis, A. K., Smith, S. R., Traller, J. C., and Abbriano, R. (2012), The place of diatoms in the biofuels industry, *Biofuels*

Hoek, C., Mann, D. G., and Jahns, H. M. (1995), Algae, (Cambridge University Press)

Hölzl, G. and Dörmann, P. (2007), Structure and function of glycoglycerolipids in plants and bacteria, *Progress in Lipid Research*

Hu, Q., Sommerfeld, M., Jarvis, E., Ghirardi, M., Posewitz, M., Seibert, M., et al. (2008), Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances, *The Plant Journal*

Hummel, J., Segu, S., Li, Y., Irgang, S., Jueppner, J., and Giavalisco, P. (2011), Ultra Performance Liquid Chromatography and High Resolution Mass Spectrometry for the Analysis of Plant Lipids, *Frontiers in Plant Science*

Jensen, N., Tomer, K., and Gross, M. (1987), FAB MS/MS for phosphatidylinostitol,-glycerol,-ethanolamine and other complex phospholipids, *Lipids*

Kanu, A. B., Dwivedi, P., Tam, M., Matz, L., and Hill, H. H. (2008), Ion mobility - mass spectrometry, *Journal of Mass Spectrometry*

Kato, M., Kobayashi, Y., Torii, A., and Yamada, M. (2003), Betaine Lipids in Marine Algae, in N. Murata, M. Yamada, I. Nishida, H. Okuyama, J. Sekiya, and W. Hajime, eds., Advanced Research on Plant Lipids SE - 3 (Springer Netherlands)

Kato, M., Sakai, M., Adachi, K., Ikemoto, H., and Sano, H. (1996), Distribution of betaine lipids in marine algae, *Phytochemistry*

Khozin-Goldberg, I. and Cohen, Z. (2011), Unraveling algal lipid metabolism: Recent advances in gene identification, *Biochimie*

Koivusalo, M., Haimi, P., Heikinheimo, L., Kostiainen, R., and Somerharju, P. (2001), Quantitative determination of phospholipid compositions by ESI-MS: effects of acyl chain length, unsaturation, and lipid concentration on instrument response, *Journal of Lipid Research*

Koornneef, M. and Meinke, D. (2010), The development of Arabidopsis as a model plant, *The Plant Journal*

Levitan, O., Dinamarca, J., Hochman, G., and Falkowski, P. G. (2014), Diatoms: a fossil fuel of the future, *Trends in Biotechnology*

Li-Beisson, Y., Shorrosh, B., Beisson, F., Andersson, M. X., Arondel, V., Bates, P. D., et al. (2010), Acyl-Lipid Metabolism, *The Arabidopsis Book / American Society of Plant Biologists*

Liebisch, G., Vizcaíno, J. A., Köfeler, H., Trötzmüller, M., Griffiths, W. J., Schmitz, G., et al. (2013), Shorthand notation for lipid structures derived from mass spectrometry, *Journal of Lipid Research*

Liu, B. and Benning, C. (2013), Lipid metabolism in microalgae distinguishes itself., *Current Opinion in Biotechnology*

Lykidis, A. (2007), Comparative genomics and evolution of eukaryotic phospholipid biosynthesis, *Progress in Lipid Research*

Marshall, A. G., Hendrickson, C. L., and Jackson, G. S. (1998), Fourier transform ion cyclotron resonance mass spectrometry: a primer, *Mass Spectrometry Reviews*

Martin, P., Van Mooy, B. A. S., Heithoff, A., and Dyhrman, S. T. (2011), Phosphorus supply drives rapid turnover of membrane phospholipids in the diatom Thalassiosira pseudonana, *The ISME Journal*

Moore, C. M., Mills, M. M., Arrigo, K. R., Berman-Frank, I., Bopp, L., Boyd, P. W., et al. (2013), Processes and patterns of oceanic nutrient limitation, *Nature Geosci*

Mouritsen, O. G. (2011), Lipidology and lipidomics–quo vadis? A new era for the physical chemistry of lipids, *Physical Chemistry Chemical Physics*

Paasche, E. (2001), A review of the coccolithophorid Emiliania huxleyi (Prymnesiophyceae), with particular reference to growth, coccolith formation, and calcification-photosynthesis interactions, *Phycologia*

Pata, M. O., Hannun, Y. A., and Ng, C. K.-Y. (2010), Plant sphingolipids: decoding the enigma of the Sphinx, *New Phytologist*

Paytan, A. and McLaughlin, K. (2007), The Oceanic Phosphorus Cycle, *Chemical Reviews*

Picchioni, G. A., Watada, A. E., and Whitaker, B. D. (1996), Quantitative high-performance liquid chromatography analysis of plant phospholipids and glycolipids using light-scattering detection, *Lipids*

Plumb, R. S., Johnson, K. A., Rainville, P., Smith, B. W., Wilson, I. D., Castro-Perez, J. M., et al. (2006), UPLC/MSE; a new approach for generating molecular fragment information for biomarker structure elucidation, *Rapid Communications in Mass Spectrometry*

Popendorf, K. J., Fredricks, H. F., and Van Mooy, B. A. (2013), Molecular Ion-Independent Quantification of Polar Glycerolipid Classes in Marine Plankton Using Triple Quadrupole MS, *Lipids*

Poulton, A. J., Adey, T. R., Balch, W. M., and Holligan, P. M. (2007), Relating coccolithophore calcification rates to phytoplankton community dynamics: Regional differences and implications for carbon export, *Deep Sea Research Part II: Topical Studies in Oceanography*

Prahl, F. G. and Wakeham, S. G. (1987), Calibration of unsaturation patterns in long-chain ketone compositions for palaeotemperature assessment, *Nature*

Read, B. A., Kegel, J., Klute, M. J., Kuo, A., Lefebvre, S. C., Maumus, F., et al. (2013), Pan genome of the phytoplankton Emiliania underpins its global distribution, *Nature*

Redfield, A. C. (1934), On the proportions of organic derivatives in sea water and their relation to the composition of plankton (University Press of Liverpool, UK)

Riekhof, W. R., Sears, B. B., and Benning, C. (2005), Annotation of Genes Involved in Glycerolipid Biosynthesis in Chlamydomonas reinhardtii: Discovery of the Betaine Lipid Synthase BTA1Cr, *Eukaryotic Cell*

Rosell-Melé, A. and McClymont, E. L. (2007), Chapter Eleven Biomarkers as Paleoceanographic Proxies, in Developments in Marine Geology, volume 1 (Elsevier)

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, *The Plant Cell*

Sakthivel, R. (2011), Microalgae lipid research, past, present: A critical review for biodiesel production, in the future, *Journal of Experimental Sciences*

Sato, N. (1988), Dual Role of Methionine in the Biosynthesis of Diacylglyceryltrimethylhomoserine in Chlamydomonas reinhardtii, *Plant Physiology*

Schatz, D., Shemi, A., Rosenwasser, S., Sabanay, H., Wolf, S. G., Ben-Dor, S., et al. (2014), Hijacking of an autophagy-like process is critical for the life cycle of a DNA virus infecting oceanic algal blooms, *New Phytologist*

Schouten, S., Hopmans, E. C., Schefuß, E., and Sinninghe Damsté, J. S. (2002), Distributional variations in marine crenarchaeotal membrane lipids: a new tool for reconstructing ancient sea water temperatures?, *Earth Planet Sc. Lett.*

Schouten, S., Middelburg, J. J., Hopmans, E. C., and Sinninghe Damsté, J. S. (2010), Fossilization and degradation of intact polar lipids in deep subsurface sediments: A theoretical approach, *Geochimica et Cosmochimica Acta*

Siegenthaler, P.-A. and Murata, N. (2004), Lipids in Photosynthesis: Structure, Function and Genetics (Springer Netherlands)

Simoneit, B. R. T. (1977), Diterpenoid compounds and other lipids in deep-sea sediments and their geochemical significance, *Geochimica et Cosmochimica Acta*

Sturt, H. F., Summons, R. E., Smith, K., Elvert, M., and Hinrichs, K.-U. (2004), Intact polar membrane lipids in prokaryotes and sediments deciphered by high-performance liquid chromatography/electrospray ionization multistage mass spectrometrynew biomarkers for biogeochemistry and microbial ecology, *Rapid Communications in Mass Spectrometry*

Suttle, C. A. (2007), Marine viruses - major players in the global ecosystem, *Nat. Rev. Micro.*

Tonon, T., Harvey, D., Larson, T. R., and Graham, I. A. (2002), Long chain polyunsaturated fatty acid production and partitioning to triacylglycerols in four microalgae, *Phytochemistry*

Tonon, T., Sayanova, O., Michaelson, L. V., Qing, R., Harvey, D., Larson, T. R., et al. (2005), Fatty acid desaturases from the microalga Thalassiosira pseudonana, *The FEBS journal*

Touchstone, J. C. (1995), Thin-layer chromatographic procedures for lipid separation, *Journal of Chromatography B: Biomedical Sciences and Applications*

Tyrrell, T. (1999), The relative influences of nitrogen and phosphorus on oceanic primary production, *Nature*

Tyrrell, T. and Merico, A. (2004), Emiliania huxleyi: bloom observations and the conditions that induce them, in H. R. Thierstein and J. R. Young, eds., Coccolithophores SE - 4 (Springer Berlin Heidelberg)

Van Etten, J. L., Graves, M. V., Müller, D. G., Boland, W., and Delaroque, N. (2002), Phycodnaviridae large DNA algal viruses, *Archives of Virology*

Van Mooy, B. A. S. and Fredricks, H. F. (2010), Bacterial and eukaryotic intact polar lipids in the eastern subtropical South Pacific: Water-column distribution, planktonic sources, and fatty acid composition, *Geochimica et Cosmochimica Acta*

Van Mooy, B. A. S., Fredricks, H. F., Pedler, B. E., Dyhrman, S. T., Karl, D. M., Koblížek, M., et al. (2009), Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity, *Nature*

Van Mooy, B. A. S., Moutin, T., Duhamel, S., Rimmelin, P., and Van Wambeke, F. (2008), Phospholipid synthesis rates in the eastern subtropical South Pacific Ocean, *Biogeosciences*

Van Mooy, B. A. S., Rocap, G., Fredricks, H. F., Evans, C. T., and Devol, A. H. (2006), Sulfolipids dramatically decrease phosphorus demand by picocyanobacteria in oligotrophic marine environments, *Proceedings of the National Academy of Sciences*

Vardi, A., Haramaty, L., Van Mooy, B. A. S., Fredricks, H. F., Kimmance, S. A., Larsen, A., et al. (2012), Hostvirus dynamics and subcellular controls of cell fate in a natural coccolithophore population, *Proceedings of the National Academy of Sciences*

Vardi, A., Van Mooy, B. A. S., Fredricks, H. F., Popendorf, K. J., Ossolinski, J. E., Haramaty, L., et al. (2009), Viral glycosphingolipids induce lytic infection and cell death in marine phytoplankton., *Science*

Vogel, G. and Eichenberger, W. (1992), Betaine Lipids in Lower Plants. Biosynthesis of DGTS and DGTA in Ochromonas danica (Chrysophyceae) and the Possible Role of DGTS in Lipid Metabolism, *Plant and Cell Physiology*

Volkman, J. (2006), Lipid Markers for Marine Organic Matter (Springer Berlin / Heidelberg)

Volkman, J. and Tanoue, E. (2002), Chemical and Biological Studies of Particulate Organic Matter in the Ocean, *Journal of Oceanography*

Volkman, J. K., Barrett, S. M., Blackburn, S. I., Mansour, M. P., Sikes, E. L., and Gelin, F. (1998), Microalgal biomarkers: A review of recent research developments, *Organic Geochemistry*

Volkman, J. K., Jeffrey, S. W., Nichols, P. D., Rogers, G. I., and Garland, C. D. (1989), Fatty acid and lipid composition of 10 species of microalgae used in mariculture, *Journal of Experimental Marine Biology and Ecology*

Wada, H. and Murata, N. (2009), Lipids in Photosynthesis: Essential and Regulatory Functions, Advances in Photosynthesis and Respiration (Springer Netherlands)

Wakeham, S. G., Hedges, J. I., Lee, C., Peterson, M. L., and Hernes, P. J. (1997), Compositions and transport of lipid biomarkers through the water column and surficial sediments of the equatorial Pacific Ocean, *Deep Sea Research Part II: Topical Studies in Oceanography*

Watson, A. D. (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*

Westbroek, P., Brown, C. W., van Bleijswijk, J., Brownlee, C., Brummer, G. J., Conte, M., et al. (1993), A model system approach to biological climate forcing. The example of Emiliania huxleyi, *Global and Planetary Change*

White, D. C., Meadows, P., Eglinton, G., and Coleman, M. L. (1993), In situ Measurement of Microbial Biomass, Community Structure and Nutritional Status [and Discussion], *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*

Yang, K. and Han, X. (2011), Accurate Quantification of Lipid Species by Electrospray Ionization Mass Spectrometry Meets a Key Challenge in Lipidomics, *Metabolites*

Yu, E. T., Zendejas, F. J., Lane, P. D., Gaucher, S., Simmons, B. A., and Lane, T. W. (2009), Triacylglycerol accumulation and profiling in the model diatoms Thalassiosira pseudonana and Phaeodactylum tricornutum (Baccilariophyceae) during starvation, *Journal of Applied Phycology*

Zehr, J. P. and Ward, B. B. (2002), Nitrogen Cycling in the Ocean: New Perspectives on Processes and Paradigms, *Applied and Environmental Microbiology*

Zubarev, R. A. and Makarov, A. (2013), Orbitrap Mass Spectrometry, *Analytical Chemistry*

Chapter 2

Lipid Remodelling by the Diatom *Thalassiosira pseudonana* under Phosphorus Stress – New Insights into Phospholipid Substitution

Jonathan E. Hunter[1,2], Joost Brandsma[3], Marcus Dymond[4], Grielof Koster[3], C. Mark Moore[1], Anthony D. Postle[3], Rachel A. Mills[1] and George S. Attard[5].

1. Ocean and Earth Science, University of Southampton, National Oceanography Centre Southampton, European Way, Southampton, SO14 3ZH, United Kingdom
2. Institute for Life Sciences, University of Southampton, SO17 1BJ, United Kingdom
3. Faculty of Medicine, University of Southampton, Southampton General Hospital, Tremona Road, Southampton, SO16 6YD, United Kingdom
4. School of Pharmacy and Biomolecular Sciences, University of Brighton, Brighton, BN2 4GJ, United Kingdom
5. Chemistry, University of Southampton, Southampton, SO17 1BJ, United Kingdom

2.1 Author Contributions

Jonathan E. Hunter designed the experiments, developed methods, carried out culturing, sample preparation, analysis, data interpretation and wrote the manuscript. The remaining co-authors assisted with experimental design, interpretation and drafting the manuscript.

2.2 Abstract

Under phosphorus (P) stress, phytoplankton substitute the phosphorus-containing headgroups of membrane lipids with non-phosphorus analogues, reducing their cellular demand for P. We find that this switch occurs in the marine diatom *Thalassiosira pseudonana*, primarily by the cessation of glycerophospholipid synthesis, accompanied by synthesis of the betaine lipid substitute diacylglycerylcarboxyhydroxymethylcholine (DGCC). A small proportion ($\sim 1/3$) of the original glycerophospholipid appeared to be broken down, the majority of P tied up in the glycerophospholipids remained as such.

Furthermore, we present data on how the individual lipid molecular species of *T. pseudonana* differ during exponential growth under P replete (P+) and P depleted (P-) conditions. For example, significant changes in the fatty acid composition of the decreasing glycerophosphatidylcholine (PC) pool were observed. PC species containing long-chain polyunsaturated fatty acids (PUFAs) decreased less than other PC species under P- conditions. The DGCC compositions synthesised under P- conditions mirrored those of the PC pool.

Several glycerophosphatidylethanolamine species were detected uniquely under P- conditions, some of which appear to bear very long chain PUFAs and may therefore be candidate molecular biomarkers. Variation in the relative abundances of diacylglycerol, mono/di-galactosyldiacylglycerol and sulfoquinovosyldiacylglycerol lipid species are also discussed.

These findings provide new insight on the response of the diatom lipidome to nutrient stress, yielding insight into this effective adaptation to life in low nutrient environments.

Keywords: *Thalassiosira pseudonana*, Lipid, Glycerophospholipid, Diatom, Lipidomics, Phosphorus, Stress, Limitation, ESI-MS, Biomarker.

2.3 Introduction

Diatoms are a diverse and numerous group of eukaryotic microalgae responsible for up to 25% of the global annual primary production (**Hoek et al.**, 1995). As lipids comprise 25 - 45% of the total cellular mass (dry weight) in diatoms (**Levitan et al.**, 2014), they represent a major source of organic carbon in the environment; e.g. in the equatorial Pacific Ocean, lipids account for 23% of the organic, total planktonic carbon (**Wakeham et al.**, 1997). Furthermore, these lipids, as well as their degradation products, contribute to the organic matter in sediments, and have the potential to provide biomarkers of palaeoceanographic conditions.

Lipids are a chemically and functionally varied group of biological molecules (**Vance and Vance**, 2008), broadly defined by their hydrophobic [and/or] amphipathic (bearing both polar and apolar moieties) character (**Fahy et al.**, 2005). Polar glycerolipids form the basic building blocks of most cellular membranes, and consist of a glycerol backbone with fatty acids attached at the sn-1 and/or sn-2 positions and a polar headgroup moiety at the sn-3 position. Each of the two hydrophobic fatty acid moieties can differ in carbon chain length, degree of unsaturation or stereochemistry and the nature of the linkage to the glycerol backbone. In combination with a range of different polar headgroups this results in a large compositional complexity (**Fahy et al.**, 2005; **Liebisch et al.**, 2013).

Lipid biosynthesis involves a network of reactions that, through the exchange of headgroups and radyl moieties, generates a rich variety of individual chemical species whose relative as well as absolute amounts change over the course of the cell cycle and in response to external stimuli and environmental conditions. Remodelling of a cells lipidome (the entirety of its cellular lipids) in response to environmental conditions is common in unicellular organisms (**Benning et al.**, 1995; **Martin et al.**, 2011) and serves to regulate the biophysical characteristics of a given membrane or to yield nutritional, metabolic or other benefits (**Liu and Benning**, 2013). One such remodelling mechanism utilizes the substitution of membrane glycerophospholipids with non-phosphorus glycerolipid counterparts when an organism is subjected to phosphorus (P) stress or starvation (**Benning**

et al., 1995; **Martin et al.**, 2011). This response allows a phytoplankter to reduce its P demands in P-limited environments (**Van Mooy et al.**, 2009). Oceanic regions such as the Sargasso Sea or the eastern Mediterranean Sea contain very low bioavailable phosphate (PO_4^{3-}) concentrations (typically <10 nmol L^{-1})(**Thingstad et al.**, 2005; **Mather et al.**, 2008; **Van Mooy et al.**, 2009). These high nitrogen (N) : P ratios are generally considered to result from high levels of nitrogen fixation (**Wu et al.**, 2000; **Sanudo-Wilhelmy et al.**, 2001). Microorganisms in these low P environments can potentially gain considerable advantage from any reduction in their physiological P requirements by allowing prioritisation of cellular P use for non-substitutable functions, such as nucleic acid synthesis, over phospholipid biosynthesis (**Van Mooy et al.**, 2009; **Guschina and Harwood**, 2006). For example, plankton communities in the Sargasso Sea were found to utilise just 1.3 ± 0.6% (**Van Mooy et al.**, 2009) of PO_4^{3-} uptake for phospholipid synthesis, versus 17 ± 6% in plankton communities from the South Pacific subtropical gyre, which has a higher PO_4^{3-} concentration (average 151.4 nmol L^{-1})(**Van Mooy et al.**, 2009).

The marine diatom *Thalassiosira pseudonana* (*T. pseudonana*) has been used as a model organism to study the effects of P starvation on lipid remodelling (**Van Mooy et al.**, 2009; **Martin et al.**, 2011). The total fatty acid profile of *T. pseudonana* has been reported (**Volkman et al.**, 1989). During exponential growth polyunsaturated fatty acids (PUFA) (predominantly 16:3, 18:4, 20:5 and 22:6 species) accounted for 53% of the total FA pool, while monounsaturated FAs (mainly the 16:1 species) contributed 20% and saturated FAs (mostly 14:0 and 16:0 species) made up the remaining 27% (**Volkman et al.**, 1989; **Zhukova**, 2004).

When grown under phosphorus replete growth conditions (P+) *T. pseudonana* predominately synthesizes the membrane glycerophospholipid classes glycerophosphatidylcholine (PC), glycerophosphatidylglycerol (PG) and glycerophosphatidylethanolamine (PE), in common with other eukaryotic phytoplankton (**Van Mooy et al.**, 2009). In addition, three non-phosphorus glyceroglycolipid classes, monogalactosyldiacylglycerol (MGDG), digalactosyldiacylglycerol (DGDG) and sulfoquinovosyldiacylglycerol (SQDG), are also present (**Martin et al.**, 2011). The distribution of these lipids within a cell is heterogenenous;

for example PG, MGDG, DGDG and SQDG appear to be associated primarily with the thylakoid membranes of the chloroplast (**Siegenthaler and Murata**, 2004).

In contrast, when *T. pseudonana* is grown under P- conditions it synthesises the nitrogen based betaine lipid diacylglycerylcarboxyhydroxymethylcholine (DGCC), which is normally undetectable under P+ conditions (**Van Mooy et al.**, 2009). This increase in DGCC is concomitant with an absolute decrease in PC lipids, and it has been suggested that these two physicochemically similar zwitterionic lipid classes can substitute for each other without significant loss of membrane function (**Eichenberger and Gribi**, 1997; **Van Mooy et al.**, 2009; **Martin et al.**, 2011). In *T. pseudonana*, the sole representative of the betaine lipids is DGCC (**Van Mooy et al.**, 2009). Other phytoplankton commonly synthesise other betaine lipid classes including diacylglyceryl hydroxymethyl-trimethyl--alanine (DGTA) and diacylglyceryltrimethylhomoserine (DGTS), and all three have been shown to be abundant in the marine environment (**Van Mooy et al.**, 2009; **Schubotz et al.**, 2009; **Van Mooy and Fredricks**, 2010; **Popendorf et al.**, 2011; **Brandsma et al.**, 2012).

In addition to the shift from PC to DGCC lipids in P-starved *T. pseudonana*, a comparable shift takes place from PG lipids to the sulphur-containing glyceroglycolipid SQDG (**Van Mooy et al.**, 2009; **Martin et al.**, 2011). The cellular abundance of the other main non-phosphorus glyceroglycolipids, MGDG and DGDG, does not appear to increase in P-starved *T. pseudonana* (**Martin et al.**, 2011). As a consequence of this lipid remodelling, the ratios of total DGCC:PC and SQDG:PG show a considerable increase in magnitude as a function of increasing P stress (**Van Mooy et al.**, 2009). These lipid ratios have been suggested as biomarkers of P limitation in planktonic communities in both marine and freshwater environments (**Van Mooy et al.**, 2009; **Bellinger et al.**, 2014). The implications of P stress upon the pennate marine diatom *Phaeodactylum triconutum* were recently reported, resulting in PG to SQDG and PC to DGTA substitution, reflecting a contrasting system (**Abida et al.**, 2014).

The kinetics of the apparent lipid headgroup substitution observed during P starvation were shown to be very rapid (**Martin et al.**, 2011). After 48 h of growth under P-

conditions, the molar percentage contribution of the glycerophospholipids to the total polar lipid pool of a *T. pseudonana* culture fell from 45.0 ± 0.9% at 0 h to 21.0 ± 4.5%. DGCC rose from below detection to 11.5 ± 1.7% and SQDG from 38.0 ± 2.0% to 52.0 ± 6.0% (**Martin et al.**, 2011). It was estimated that the amount of P that was not channelled into the synthesis of glycerophospholipids corresponded to four haploid genomes (**Martin et al.**, 2011). The response of P-starved *T. pseudonana* to the resupply of P is equally rapid, occurring over a 12 - 24 h period, and reverses the trend seen in P- cultures, such that the amount of glycerophospholipids increases while the amount of DGCC decreases (**Martin et al.**, 2011).

Transcriptomic and proteomic investigations of *T. pseudonana* show differential transcription and expression of a homolog of the sqdB gene and of a gene coding for an N-acetylglucosaminyltransferase complex subunit, which is required for phosphatidylinositol/sulfolipid biosynthesis (**Dyhrman et al.**, 2012). Transcription and expression of both these genes are upregulated in P- conditions, consistent with lipid remodelling resulting in an increase in the SQDG:PG ratio. Interestingly, these studies of *T. pseudonana* did not reveal any protein candidates or transcripts that might be implicated in the dramatic increase in DGCC that is observed in P- cultures.

The changes in the lipids of *T. pseudonana* as a consequence of P stress have been studied quantitatively in terms of the major polar headgroup classes. However, little is known about how the lipid species within each class, varying in fatty acids, change in response to P stress. Here we describe the use of mass spectrometry to quantify each of the polar lipid classes as a whole, and the individual species within each class, for *T. pseudonana* cultured under P+ and P- conditions over a period of 96 h. Our findings shed new light on the dynamics of glycerophospholipid substitution at the individual lipid species level, with implications upon lipid substitution as a means for *T. pseudonana* to cope with P stress. Furthermore, the biosynthetic relationship between the various polar lipid types is discussed and potential candidate biomarkers are highlighted.

2.4 Results

2.4.1 Culture Nutrient Concentrations

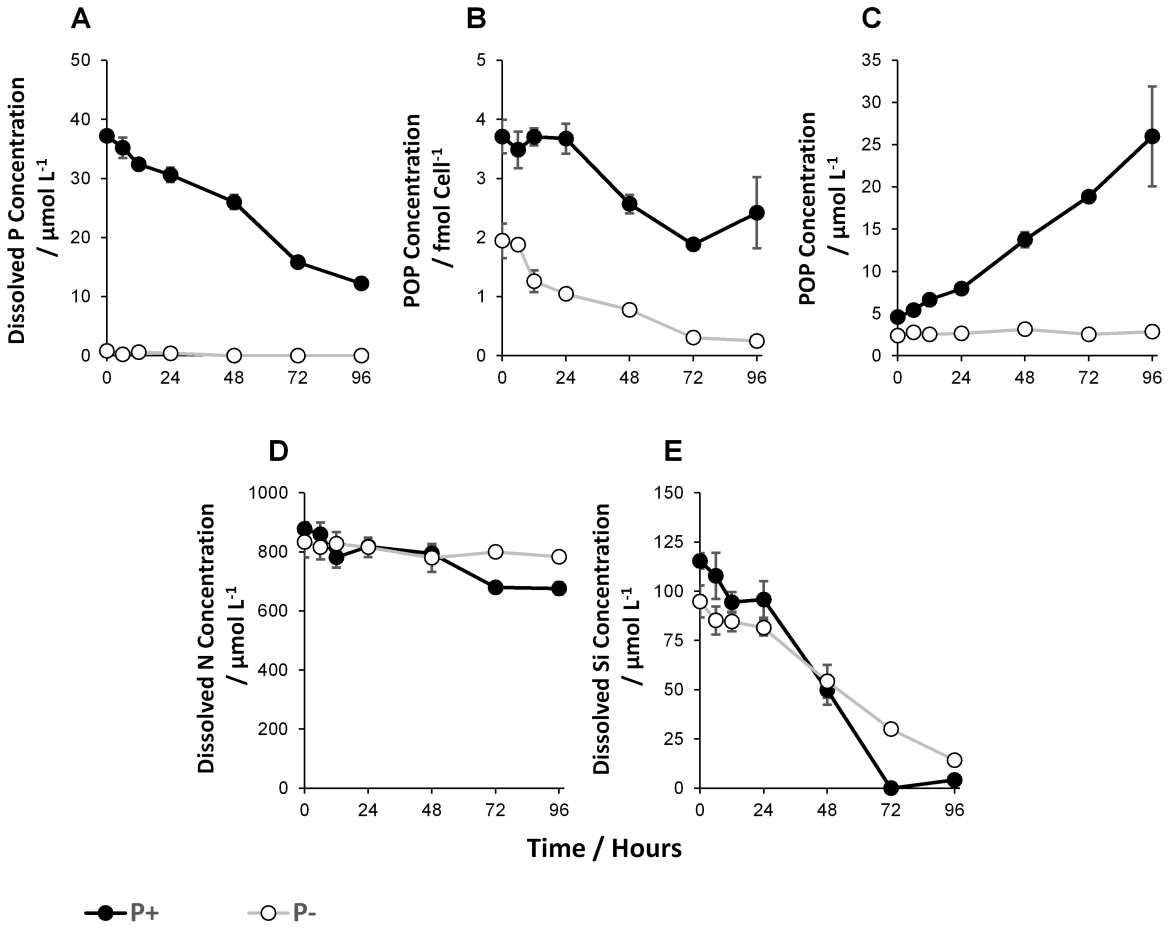


Figure 2.1: Macronutrient concentrations in the cultures through time. Dissolved phosphorus concentration in the growth media (A); particulate organic phosphorus (POP) per cell (B); particulate organic phosphorus concentration per litre growth media (C); dissolved nitrogen (D) and dissolved silicon concentration in the growth media (E). Data are the mean of n = 3 biological replicates, with error bars of 1 standard deviation. Statistical significance is indicated in accompanying text, *p<0.05 **p<0.005 by two-tailed, paired equal-variance T-test.

Dissolved macronutrients and particulate organic phosphorus (POP) were quantified throughout the time course (Figure 2.1). The dissolved phosphate ($[PO_4^{3-}]$) concentra-

tion (Figure 2.1A) in the phosphorus replete (P+) cultures at 0 h was 37.20 ± 0.92 μmol L^{-1} and decreased through time to a minimum of 12.20 ± 0.60 μmol L^{-1} at 96 h. [PO_4^{3-}] in the phosphorus stressed (P-) cultures at 0 h was 0.80 ± 0.00 μmol L^{-1}, dropping to 0.20 ± 0.00 μmol L^{-1} after 6 h.

POP quantity per cell (Figure 2.1B) was 0.10 ± 0.03** fold less in the P- cultures than the P+ cultures at 96 h. A discrepancy was evident however, between the P+ and P- values at zero time. When considering POP quota per cell at 0 h, a P+ value of 3.71 ± 0.28 fmol $cell^{-1}$ and a P- value of 1.95 ± 0.29 fmol $cell^{-1}$ were observed.

POP concentration per litre culture media (Figure 2.1C) increased through time in the P+ cultures, from 4.57 ± 0.23 μmol L^{-1} at 0 h to 25.97 ± 5.92 μmol L^{-1} at 96 h. POP concentration per litre remained constant through time in the P- cultures, with a minima of 2.37 ± 0.31 μmol L^{-1} at 0 h and a maxima of 3.13 ± 0.15 μmol L^{-1} at 48 h.

Dissolved NO_3^- (Figure 2.1D) remained in excess at >675 μmol L^{-1} throughout, in both P+ and P- cultures. Dissolved Si (orthosilicate ions [SiO_4^{4-}]) (Figure 2.1E) dropped from 115.40 ± 3.99 μmol L^{-1} at 0 h to below detection (<6.00 μmol L^{-1}) after 72 h in the P+ cultures, but only fell from 94.80 ± 8.12 μmol L^{-1} to 30.00 ± 2.16 μmol L^{-1} in the P- cultures over the same time period.

2.4.2 Culture Growth Dynamics

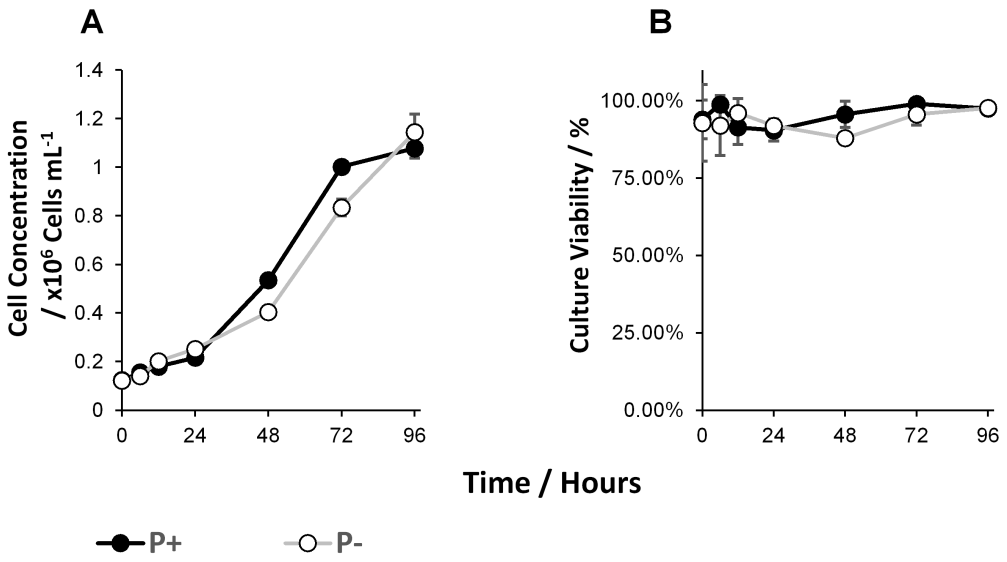


Figure 2.2: Cell concentration growth curve (A) and culture viability through time (B). Data are the mean of n = 3 biological replicates, with error bars of 1 standard deviation. Statistical significance is indicated in accompanying text, *p<0.05 **p<0.005 by two-tailed, paired equal-variance T-test.

During exponential growth of the cultures (Figure 2.2A) between 0 and 72 h, the P+ cultures grew at a rate of 0.029 ± 0.0013 (23.76 ± 1.08 h doubling time). Exponential growth during the same period in the P- cultures was marginally slower at 0.026 ± 0.0016* (27.09 ± 1.78 h doubling time). Both the P+ and P- cultures appeared to be transitioning into stationary phase at 96 h with maximum populations of 1.08 x 10^6 ± 4.11 x 10^4 cells mL^{-1} and 1.14 x 106 ± 7.35 x 104 mL^{-1} respectively. Culture viability was high throughout at >90 %, in both P+ and P- cultures, with the exception of the P- cultures at 48 h (87.95 ± 1.78 %) (Figure 2.2B).

2.4.3 Lipid Class Level Response to Phosphorus Stress

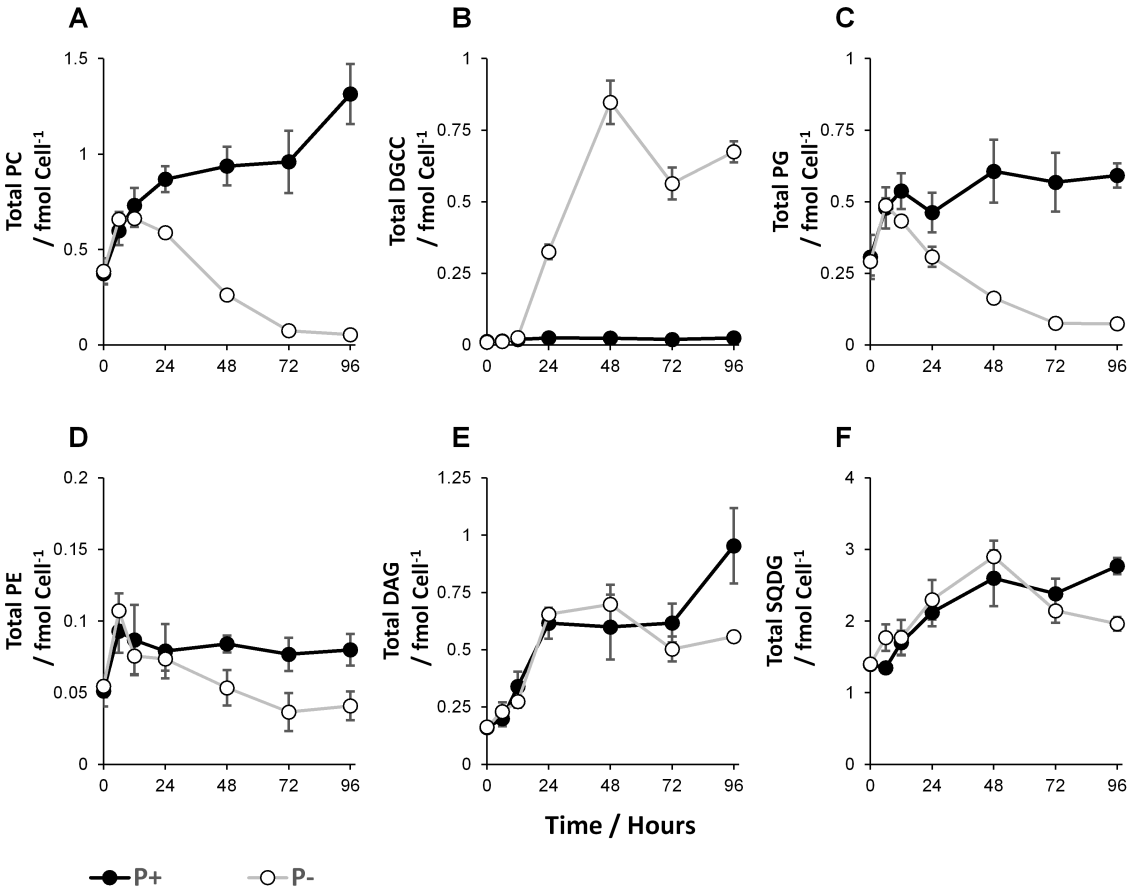


Figure 2.3: Variation in total lipid content per cell through time for each of the lipid classes: Glycerophosphatidylcholine (PC, A); Diacylglycerylcarboxyhydroxymethylcholine (DGCC, B); Glycerophosphatidylglycerol (PG, C); Glycerophosphatidylethanolamine (PE, D); Diacylglycerol (DAG, E); and sulfoquinovosyldiacylglycerol (SQDG, F). Data are the mean of n = 3 biological replicates, with error bars of 1 standard deviation. Statistical significance is indicated in accompanying text, *p<0.05 **p<0.005 by two-tailed, paired equal-variance T-test.

Glycerophosphatidylcholine (PC, Figure 2.3A) was the predominant glycerophospholipid. Between 0 h and 12 h, PC quantity increased between 0.39 ± 0.067 fmol cell^{-1} and 0.66 ± 0.040 fmol cell^{-1} respectively in the P- cultures, in line with the P+ control cultures. PC quantity in the P- cultures then decreased dramatically from 24 h onwards, reaching a

minimum of 0.054 ± 0.0074 fmol cell^{-1} after 96 h, 0.041 ± 0.0075** fold less than the P+ control.

The betaine lipid Diacylglycerylcarboxyhydroxymethylcholine (DGCC, Figure 2.3B), was absent from the P+ cultures throughout and the P- cultures before 12 h. Concomitant with the decline in PC, DGCC quantities increased through time in the P- cultures following 12 h to reach a maximum of 0.85 ± 0.075 fmol cell^{-1} at 48 h and 0.67 ± 0.036 fmol cell^{-1} at 96 h.

Glycerophosphatidylglycerol (PG, Figure 2.3C) in the P- cultures increased in quantity between 0 h and 6 h from 0.29 ± 0.049 fmol cell^{-1} to 0.49 ± 0.025 fmol cell^{-1}, in line with the P+ cultures. This increase was followed by a decline from 12 h onwards in the P- cultures resulting in a minimum of 0.074 ± 0.019 fmol cell^{-1} at 96 h, equivalent to a 0.12 ± 0.033** fold reduction.

Glycerophosphatidylethanolamine (PE, Figure 2.3D), like PC and PG, showed an initial increase in the P- cultures from 0.055 ± 0.0078 fmol cell^{-1} at 0 h to 0.11 ± 0.012 fmol cell^{-1} at 6 h, in line with the P+ cultures. This was followed by a decline in PE quantity in the P- cultures from 12 h onwards to a minimum of 0.037 ± 0.013 fmol cell^{-1} at 72 h, equivalent to a 0.48 ± 0.19* fold reduction.

Diacylglycerol (DAG, Figure 2.3E) quantity per cell did not vary significantly between P+ and P- cultures, with the exception of a 0.58 ± 0.10* fold reduction in the P- cultures at 96 h.

The sulfolipid sulfoquinovosyldiacylglycerol (SQDG, Figure 2.3F), was 1.31 ± 0.14* fold greater in quantity in the P- cultures than the P+ control at 6 h. SQDG quantities were then statically similar in P+ and P- between 12 and 72 h, followed by a 0.71 ± 0.045** fold reduction in the P- cultures, relative to P+ at 96 h.

2.4.4 Glycerophospholipid Substitution Dynamics

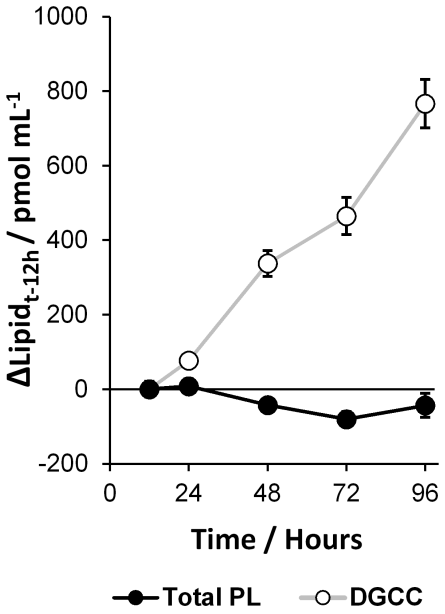


Figure 2.4: Change in total lipid quantity (total phospholipids (PL) and DGCC) per mL culture ($Lipid_{t\text{-}12h}$), between time t and 12 h, in the P- cultures. $y < 0$ indicates a net degradation and loss of the total phospholipids, $y = 0$ indicates a constant quantity (no net synthesis or degradation), $y > 0$ indicates biosynthesis and net increase in total lipid quantity. Values are relative to 12 h as the initiation of DGCC biosynthesis and maximum PL in the P- cultures. Data are the mean of $n = 3$ biological replicates, with error of 1 standard deviation. Statistical significance is indicated in accompanying text, *$p<0.05$ **$p<0.005$ by two-tailed, paired equal-variance T-test.

At 48 and 72 h in the P- cultures, the change in total phospholipid ($PL_{t\text{-}12h}$) was -42.27 ± 18.19* pmol mL^{-1} and -80.47 ± 20.03** pmol mL^{-1} respectively, indicating a net degradation of the total phospholipids at these times (Figure 4). The maximum decrease at 72 h corresponded to 34.10 ± 8.75 % of the maximum total phospholipid at 12 h. The change in total DGCC ($DGCC_{t\text{-}12h}$) increased rapidly through time reaching a maximum of 772.13 ± 64.76 pmol mL^{-1} at 96h. At 72 h, the molar decrease in $PL_{t\text{-}12h}$ was equivalent to 17.12 ± 4.63 % of the increase in DGCC.

2.4.5 Fatty Acid Level Variability during P-replete Growth

Time / Hours

C	DB	0	6	12	24	48	72	96
32	1	11.8 ±0.5 3.2%	12.5 ±1.7 2.1%	15.8 ±0.8 2.2%	27.4 ±3.1 3.2%	37.8 ±6.1 4.0%	59.2 ±10.9 6.2%	62.7 ±0.5 4.8%
34	4	19.6 ±3.1 5.2%	38.3 ±6.7 6.4%	35.4 ±1.6 4.9%	42.1 ±4.0 4.8%	67.8 ±10.5 7.2%	119.4 ±14.7 12.5%	110.4 ±15.3 8.4%
34	5	23.9 ±3.8 6.4%	16.8 ±2.5 2.8%	17.3 ±0.7 2.4%	19.1 ±2.1 2.2%	22.5 ±2.9 2.4%	45.1 ±6.3 4.7%	45.0 ±5.2 3.4%
36	5	14.3 ±2.4 3.8%	52.4 ±7.1 8.8%	65.2 ±2.8 9.0%	74.6 ±5.4 8.6%	67.4 ±7.3 7.2%	73.8 ±13.1 7.7%	76.3 ±6.6 5.8%
36	6	38.5 ±7.1 10.3%	29.8 ±3.5 5.0%	34.3 ±3.7 4.7%	58.1 ±4.8 6.7%	77.5 ±6.6 8.3%	87.0 ±18.9 9.0%	102.7 ±19.3 7.8%
36	7	21.8 ±4.0 5.8%	10.0 ±0.8 1.7%	8.8 ±1.5 1.2%	19.9 ±3.2 2.3%	35.2 ±3.1 3.8%	38.8 ±8.0 4.0%	45.7 ±9.8 3.5%
38	6	14.5 ±1.5 3.9%	29.3 ±5.8 4.9%	47.4 ±3.4 6.5%	58.1 ±2.7 6.7%	38.8 ±3.9 4.2%	31.8 ±7.5 3.3%	104.4 ±17.5 8.1%
38	7	15.6 ±2.8 4.2%	23.5 ±3.3 3.9%	21.9 ±2.2 3.0%	28.2 ±1.8 3.3%	28.3 ±1.8 3.0%	19.9 ±3.5 2.1%	69.2 ±14.2 5.2%
38	9	36.2 ±7.4 9.6%	77.4 ±11.1 12.9%	65.3 ±13.5 8.9%	79.0 ±11.2 9.1%	142.4 ±22.9 15.2%	134.4 ±21.4 14.0%	159.3 ±39.8 12.0%
40	10	18.3 ±3.5 4.9%	105.4 ±10.7 17.7%	147.6 ±22.7 20.1%	134.7 ±13.4 15.5%	95.0 ±5.9 10.2%	38.4 ±6.0 4.0%	48.1 ±17.3 3.6%
42	11	18.1 ±4.4 4.8%	70.4 ±9.6 11.8%	128.7 ±28.9 17.5%	168.6 ±9.8 19.4%	105.0 ±12.4 11.2%	39.7 ±7.3 4.1%	53.7 ±15.5 4.0%

PC Quantity (P+ Cultures) / amol cell^{-1}
± Std. Deviation
Relative Abundance (Within Class) / %

Figure 2.5: Individual PC lipid molecular species quantity per cell, through time, in the P+ control cultures. Molecular species contributing less than 5% of the total quantity of PC throughout the time course were excluded for brevity. Full datasets are included as Supplementary Figures 7.2-7.9. Darker shades indicate a greater quantity per cell of the given lipid species. Data are the mean of n = 3 biological replicates, with error of 1 standard deviation.

Variation in the quantity of individual PC lipid species (varying in fatty acyl identity) through time, in the P+ control cultures is shown in Figure 2.5. The top five molecular species, as ranked by their maximum quantity per cell, behaved as follows:

The polyunsaturated PC(42:11) increased from 18.11 ± 4.39 amol cell^{-1} at 0 h to a maximum of 168.57 ± 9.83 amol cell^{-1} at 24 h, before decreasing to 53.69 ± 15.53 amol cell^{-1} at 96 h. PC(38:9) increased through time from 36.21 ± 7.44 amol cell^{-1} at 0 h to 159.26 ± 39.81 amol cell^{-1} at 96 h. PC(40:10) behaved similarly to PC(42:11), displaying an increase from 18.35 ± 3.52 at 0 h to 147.55 ± 22.74 amol cell^{-1} at 12 h, followed by a decrease to 48.06 ± 17.26 amol cell^{-1} at 96 h. PC(34:4) increased from 19.54 ± 3.14 amol cell^{-1} at 0 h to a maximum of 119.37 ± 14.69 amol cell^{-1} at 72 h. PC(38:6) increased from 14.54 ± 1.51 amol cell^{-1} to 58.11 ± 2.71 amol cell^{-1} between 0 and 48 h, before falling to 31.83 ± 7.55 amol cell^{-1} at 72 h, and then increasing greatly to 104.39 ± 17.52 amol cell^{-1} at 96 h. These data indicate a great degree of variability in PC lipid quantity through time in the P+ control cultures, with no clear pattern underlying the observed variation.

2.4.6 Fatty Acid Level Response to Phosphorus Stress

A

Lipid Class / Treatment

C	DB	PC P+	PC P-	DGCC P+	DGCC P-	PE P+	PE P-
29	5					0.2 ±0.4* 0.3%	4.0 ±1.1* 7.5%
31	6					0.1 ±0.2* 0.2%	3.1 ±0.6* 5.8%
32	1			0.0 ±0.0* %	54.7 ±4.5* 6.6%		
34	4	67.8 ±10.5* 7.2%	7.8 ±0.6* 3.0%				
36	2					5.5 ±3.2 6.8%	0.8 ±0.6 1.5%
	5	67.4 ±7.3* 7.2%	17.7 ±1.0* 6.7%	0.0 ±0.0* %	116.7 ±7.7* 14.1%		
	6	77.5 ±6.6* 8.3%	23.9 ±1.5* 9.1%	0.0 ±0.0* %	67.1 ±6.3* 8.1%	6.3 ±2.6* 7.4%	1.3 ±1.0* 2.3%
38	6	38.8 ±3.9* 4.2%	13.3 ±0.6* 5.1%	0.0 ±0.0* %	57.0 ±6.1* 6.9%		
	9	142.4 ±22.9* 15.2%	17.0 ±1.1* 6.5%				
40	10	95.0 ±5.9* 10.2%	52.3 ±0.4* 20.0%	0.0 ±0.0* %	92.2 ±5.5* 11.2%	13.8 ±2.2* 16.5%	7.5 ±3.1* 13.7%
42	11	105.0 ±12.4* 11.2%	52.2 ±1.1* 19.9%	0.0 ±0.0* %	89.8 ±6.5* 10.9%	32.3 ±9.0* 38.1%	12.6 ±2.5* 23.7%
46	5					0.9 ±0.4* 1.1%	2.9 ±0.3* 5.6%
52	12					0.2 ±0.2* 0.3%	3.8 ±0.8* 7.4%
54	13					0.7 ±0.2* 0.8%	3.8 ±1.9* 6.8%

B

Lipid Class / Treatment

C	DB	PG P+	PG P-	DAG P+	DAG P-	SQDG P+	SQDG P-	MGDG P+	MGDG P-	DGDG P+	DGDG P-
L20	1			34.6 ±4.6 5.9%	32.9 ±7.2 4.7%						
L22	1			29.6 ±1.5* 5.1%	65.3 ±14.7* 9.3%						
28	0					1,080.8 ±154.3 41.7%	1,262.4 ±41.5 43.7%				
30	0					386.5 ±72.4 14.8%	484.2 ±63.6 16.7%				
	1	33.1 ±8.1* 5.4%	6.8 ±2.1* 4.2%	87.7 ±20.9 14.7%	98.2 ±20.9 14.0%	268.9 ±17.8* 10.5%	318.4 ±23.9* 11.0%	12.3%	18.1%	18.3%	16.8%
32	0									5.4%	0.4%
	1	278.9 ±45.3* 46.1%	59.0 ±4.5* 36.0%	201.6 ±52.2 33.5%	270.5 ±27.6 38.8%	384.4 ±101.7 14.6%	405.0 ±32.8 14.0%	37.9%	41.5%	44.7%	43.9%
	2	44.9 ±8.9* 7.4%	12.0 ±0.9* 7.3%	69.8 ±16.8 11.8%	96.4 ±8.8 13.8%			18.2%	18.6%	15.9%	25.6%
	3							8.2%	4.9%	4.1%	6.2%
	6							9.1%	4.0%		
34	3	10.2 ±2.9 1.8%	8.2 ±3.1 5.0%								
	4	22.1 ±5.6 3.6%	18.1 ±2.0 11.1%								
36	5	84.9 ±22.6* 13.9%	19.9 ±5.1* 12.1%								
	6	55.1 ±10.5* 9.2%	12.3 ±3.7* 7.5%								
	7									7.2%	4.2%

Quantity / amol cell^{-1}
± Std. Deviation
Relative Abundance (Within Class) / %

Figure 2.6: Individual lipid molecular species quantity per cell, at 48 h in the P+ and P- cultures. Molecular species contributing less than 5% of the total quantity in their lipid head group class in both treatments were excluded for brevity. Full datasets are included as Supplementary Figures 7.2-7.9. Red indicates a decrease in quantity per cell of the given lipid species between P- and P+ cultures and green an increase. *Lipid species varying statistically significantly in quantity per cell between P- and P+ cultures. Data are the mean of n = 3 biological replicates, with error bars of 1 standard deviation. Statistical significance is indicated in accompanying text, *$p<0.05$ **$p<0.005$ by two-tailed, paired equal-variance T-test.

Variation in individual lipid species quantity, between P+ and P- cultures at 48 h is shown in Figure 2.6. All PC species showed a decrease in quantity per cell in the P- cultures, by comparison to the P+ control cultures. Variation in the magnitude of this decrease is evident from the data: PC(34:4) and PC(38:9) showed the largest fold decreases of 0.12 ± 0.020** and 0.12 ± 0.021** fold respectively. PC(36:5), PC(36:6) and PC(38:6) displayed intermediate fold decreases between the P- and P+ cultures, of 0.26 ± 0.032**, 0.31 ± 0.032** and 0.34 ± 0.037** fold respectively. Highly unsaturated PC(42:11) and PC(40:10) species displayed the smallest fold decreases of 0.50 ± 0.060** and 0.55 ± 0.035** fold respectively.

Betaine lipid DGCC species occurred only in the P- cultures and were entirely absent from the P+ control cultures. At 48 h, the five most abundant DGCC species in the P- cultures were DGCC(36:5), DGCC(40:10), DGCC(42:11), DGCC(36:6) and DGCC(38:6), with quantities of 116.69 ± 7.65, 92.22 ± 5.54, 89.77 ± 6.49, 67.11 ± 6.30 and 57.00 ± 6.10 amol $cell^{-1}$ respectively.

The glycerophospholipid PG behaved in line with PC and showed a decrease in quantity per cell for its molecular species in the P- cultures, relative to the P+ control cultures. PG(30:1) displayed the largest fold decrease at 0.20 ± 0.080* fold, followed by PG(32:1) at 0.21 ± 0.038** fold, PG(36:6) at 0.22 ± 0.079**, PG(36:5) at 0.23 ± 0.087* fold and PG(32:2) at 0.27 ± 0.057 fold**. PG(34:3) and PG(34:4) species were statistically similar in quantity per cell between treatments.

Glycerophospholipid PE species decreased in the P- cultures relative to the P+ control cultures. PE(36:6), PE(42:11) and PE(40:10) decreased by 0.21 ± 0.18*, 0.39 ± 0.13* and 0.55 ± 0.24* fold respectively under P stress. In contrast to PC and PG however, a number of minor PE species were present under P- conditions, but absent in the P+ cultures, representing a per cell increase under P stress. PE(29:5), PE(54:13), PE(52:12), PE(31:6), PE(46:5), were present under P stress at 3.97 ± 1.14* amol $cell^{-1}$, 3.80 ± 1.90* amol $cell^{-1}$, 3.79 ± 0.83** amol $cell^{-1}$, 3.08 ± 0.56** amol $cell^{-1}$ and 2.88 ± 0.28** amol $cell^{-1}$ respectively.

The majority of DAG molecular species did not appear to vary and only one displayed

a statistically significant variation between the P- and P+ cultures. LDAG(22:1) increased 2.21 ± 0.51* fold under P stress.

Within the sulfolipid SQDG class, only SQDG(30:1) was differentially abundant between P- and P+ cultures, increasing by 1.18 ± 0.12* fold under P stress.

The glycolipids MGDG and DGDG, for which suitable standards for internal standard quantification were not available, were quantified only in relative terms within each lipid head group class. MGDG(30:1) increased 1.48 ± 0.18** fold in its percentage relative abundance of the total MGDG, between the P- and P+ cultures. MGDG(32:6) and MGDG(32:3) decreased 0.44 ± 0.15* and 0.60 ± 0.088** fold respectively. DGDG(32:2) increased its relative abundance by 1.61 ± 0.49* fold and DGDG(32:0) decreased by 0.068 ± 0.064** fold under P stress.

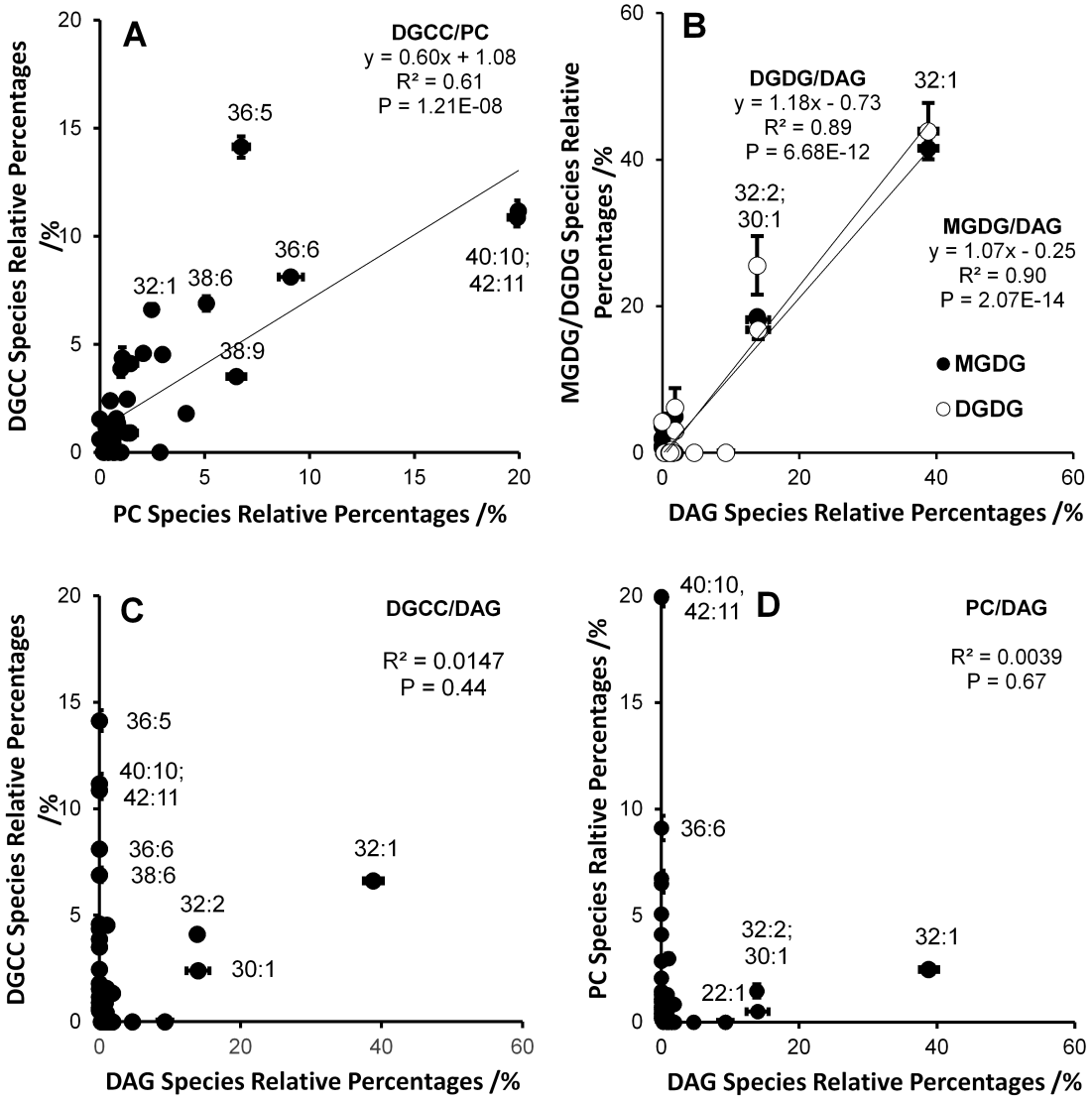


Figure 2.7: Bivariate analyses comparing % relative abundance of individual lipid species within two lipid classes after 48h P starvation, as a measure of compositional similarity. Comparisons are between DGCC/PC (A), MGDG/DAG and DGDG/DAG (B), DGCC/DAG (C) and PC /DAG (D). An identical fatty acyl composition between two lipid classes would yield a regression coefficient (m = 1) and a coefficient of determination ($R^2 = 1$). As the % relative abundances sum to 100 in both cases, the similarity can be simply assessed upon R^2. Data are the mean of n = 3 biological replicates, with error bars of 1 standard deviation. P represents the statistical significance of the linear regression.

Bivariate analyses (Figure 2.7) were used to compare the relative abundance of individual lipid species within pairs of lipid classes as a measure of compositional similarity with respect to their diglyceride identities. The substitute betaine lipid DGCC was correlated with PC (Figure 2.7A) with $R^2 = 0.61$, $P < 0.005$. The composition of the glyceroglycolipids MGDG and DGDG was correlated with DAG (Figure 2.7B) with $R^2 = 0.90$, $P < 0.005$ and $R^2 = 0.89$, $P < 0.005$ respectively. PC and its substitute DGCC were not correlated with DAG (Figure 2.7C and D), yielding $R^2 = 0.015$ and $R^2 = 0.0039$ respectively.

2.5 Discussion

2.5.1 Culture Nutrient Concentrations

Macronutrient concentrations in the experimental cultures were measured throughout the time course (Figure 2.1). Dissolved P in the P+ control was in excess throughout the course of the experiment, such that the control cultures did not experience P stress. In the P- cultures, dissolved P was scarce, thus subjecting the P- cultures to P stress.

A discrepancy in the POP data, whereby the POP concentration per cell and per mL at 0 h was higher in the P+ cultures than the P- cultures was evident. This is attributable to carryover of dissolved P from the high [PO_4^{3-}] growth medium used in the P+ cultures onto the filter. Because of the relatively low cell numbers involved, this would have a bigger effect on the P quota values calculated at the early time points. Regardless, P stress is supported by the POP data, whereby a constant POP concentration in the media was observed through time in the P- cultures, resulting in a large reduction in per cell POP concentration.

Dissolved N was in excess throughout the experiment in both treatments, negating any N stress. We note that after 96 h incubation, the cultures are likely to be under Si stress, particularly the P+ cultures .

2.5.2 Culture Growth Dynamics

T. pseudonana growth rates (Figure 2.2A) were observed to have been reduced slightly for the P- cultures relative to the P+ controls. The lack of growth limitation over 96 hours and similarity of the growth curve to the P+ control, despite clear evidence of P stress, is a reflection of the overall capacity of this organism to adapt to low P environments (**Dyhrman et al.**, 2012). The cultures appeared healthy throughout as determined by the viability measurements (Figure 2.2B). The plateau in culture growth between 72 and 96 h, concomitant with exhaustion of dissolved Si in the culture media likely indicates the transition to a Si limited stationary growth phase.

2.5.3 Lipid Class Level Response to Phosphorus Stress

At the total lipid class level (Figure 2.3), we observed initial increases in glycerophospholipid (PC, PG, PE) quantities per cell up to 12 h, in the P- cultures. This delay in the lipid remodelling response to P stress is attributable to the utilisation of carry over of dissolved P from the P-replete seed cultures. After 12 h, a rapid decline in per-cell glycerophospholipids was observed, concomitant with the detection, from absence, of the betaine substitute lipid DGCC. These observations are consistent with previous reports (**Martin et al.**, 2011).

The dynamics of PE appeared distinct from PC and PG with a lesser magnitude of decrease under P stress. This difference is attributable to the increase of select PE molecular species discussed in the following sections.

DAG, which to our knowledge has not been previously characterised in the *T. pseudonana* P stress system, did not vary significantly in total quantity between the P- and P+ cultures during the time window attributable to P stress (12 - 72 h).

The sulfolipid SQDG, did not vary statistically significantly between the P- and P+ cultures, between 12 and 72 h. SQDG therefore, did not vary significantly under P stress and did not appear to act as a substitute lipid for PG in this case, as found in other systems (**Van Mooy et al.**, 2009).

2.5.4 Glycerophospholipid Substitution Dynamics

The difference in total phospholipid per unit volume culture, between time (t) and 12 h in the P- cultures was used to assess the change in bulk lipid, regardless of culture growth (Figure 2.4). Phospholipid net breakdown, resulting from a decrease in total phospholipid quantity per cell at a greater rate than culture growth by cell division, was statistically significant at 48 and 72 h. The maximum breakdown observed at 72 h was a modest $\sim 1/3$ of the maximum total phospholipid at 12 h. The molar decrease in total phospholipid was equivalent to $\sim 1/6$ of DGCC synthesis at 72 h.

The observation of a decrease in the amounts of lipids belonging to the phospholipid classes (PC, PE and PG) under P- conditions, and the accompanying increase in DGCC,

has previously been interpreted as indicative of a breakdown in phospholipids to liberate, and hence recycle, phosphorus (**Martin et al.**, 2011). The resulting remodelling of the cell surface lipids has been postulated to be one of the five major strategies that *T. pseudonana* has evolved to deal with P deficiency (**Dyhrman et al.**, 2012).

Based upon this analysis, the dynamics of phospholipid substitution in *T. pseudonana* follow a cessation of phospholipid net synthesis with a minor degree of phospholipid net breakdown. The majority of the cellular phospholipids are retained and diluted in per cell terms with the progression of culture growth by cell division. Therefore, the majority of the original lipid bound P remains as such and does not appear to be made available for other cellular processes, as previously hypothesised (**Martin et al.**, 2011).

The molar quantity of phospholipid breakdown accounts for a minor proportion of the biosynthesised DGCC. This suggests that the recycling of diglyceride moieties from phospholipid breakdown could form at most a minor source for incorporation into the newly biosynthesised DGCC.

2.5.5 Fatty Acid Level Variability during P-replete Growth

The composition of individual lipid species, within PC in the P+ control cultures, was analysed (Figure 2.5). We observed a great degree of variability, in some cases in excess of five-fold, in the quantity of individual lipid species through time. The polyunsaturated PC(40:10) and PC(42:11) increased dramatically from 0 to 12 h, before decreasing to 96 h. Conversely, a number of more saturated species such as PC(34:4) increased gradually throughout, reaching their greatest concentrations per cell at late time 72/96 h.

The observed increase in the polyunsaturated species (PC(40:10), PC(42:11)) was concomitant with the exponential growth phase and declined later with the transition to stationary phase. Conversely, species containing more saturated/monounsaturated fatty acids, such as PC(34:4) showed a steady increase through time, reaching maxima with the transition to stationary growth phase. These observations are in agreement with previous reports on the relationship between eukaryotic phytoplankton glycerolipids and growth phase (**Hodgson et al.**, 1991).

The complexity and degree of the temporal changes in the individual lipid species composition, within the P+ control cultures, highlight the lipidomic plasticity of *T. pseudonana* and its highly dynamic nature, both often poorly accounted for in biomarker studies.

2.5.6 Fatty Acid Level Response to Phosphorus Stress

The change in individual lipid species quantity per cell, under P stress, for each of the lipid classes is presented (Figure 2.6). PC species were all observed to decrease in quantity per cell in the P- cultures, relative to the P+ control cultures, at 48 h. Substantial variation in the magnitude of this decrease for each species was evident. PC(40:10) and PC(42:11) displayed the smallest fold decreases, resulting in a shift in PC composition under P stress to a greater proportion of these highly unsaturated species. The implications of this shift in composition are unclear but may be related to the increase of the same lipid species produced during exponential growth under P+ conditions, as previously discussed.

Reflecting the total lipid class level results, DGCC species were detected in the P- cultures and entirely absent in the P+ control cultures. The DGCC individual lipid species composition and its relationship with PC is discussed later.

The lipid species within the glycerophospholipid PG class decreased unanimously under P stress, with the exception of the relatively minor PG(34:3) and PG(34:4) species, that showed no significant change in quantity per cell between treatments. The fold change of the major PG lipid species was statistically similar throughout. As such, PG behaved broadly in line with the more abundant PC.

The predominate PE lipid species in the P+ cultures were PE(36:6), PE(42:11) and PE(40:10). These showed a decrease in quantity per cell in the P- cultures, of similar magnitudes to those observed in PC and PG. Interestingly, five PE species that were absent in the P+ cultures, were detected in P-. This equates to a per-cell increase under P stress, in contrast to the behaviour of the remaining PE species. The species involved (PE(29:5), PE(31:6), PE(46:5), PE(52:12), PE(54:13)) were present in low abundance and appear to be unique to the PE class by these analyses.

The reported total fatty acid profile of *T. pseudonana* (P+ conditions) includes fatty

acids up to 22:6 (**Volkman et al.**, 1989). Therefore, the diglycerides of these PE species require carbon chains that are longer than those reported previously. The very long chain polyunsaturated fatty acids (VLC-PUFA), required to account for these lipids (eg. 28:7) have, however, been characterised in other phytoplankton (**Rezanka et al.**, 2008; **Rezanka and Sigler**, 2009).

PE lipids are well documented to be associated with proteins involved in autophagy. Autophagy is a common response of eukaryotic cells to nutrient starvation, involving the degradation of cellular components to free up intracellular nutrients (**Duszenko et al.**, 2011). Autophagy related protein ATG8 bears a conjugated PE molecule. A putative orthologue to ATG8 is one of several in the genome of *T. pseudonana* (**Rigden et al.**, 2009; **Duszenko et al.**, 2011). Autophagy protein conjugated PE has been reported in *Emiliania huxleyi* subject to viral infection (**Schatz et al.**, 2014). We speculate therefore, that the PE species that increasing in absolute abundance under phosphorus starvation may be associated with autophagy related protein(s). Further work would be required to elucidate this link. Isolation and direct lipidomic characterisation of the autophagosomes generated under P stress could assess this hypothesis. Subject to further characterisation and investigation, these unusual PE species may have potential as molecular biomarkers for phosphorus starvation and/or autophagy in diatoms.

The quantity of lipid molecular species within the DAG class was statistically similar between P- and P+ cultures, with the exception of LDAG(22:1) which showed a ~ 2 fold increase under P stress. The role of this lysolipid is unknown and it did not appear in any of the other lipid classes analysed. Like DAG, SQDG molecular species quantity did not generally vary subject to P stress, with the exception of SQDG(30:1) which showed a minor increase in per-cell quantity. Taken together with the observation of no response at the total class level, DAG and SQDG were not significantly affected by P stress.

The glyceroglycolipids MGDG and DGDG were quantified in relative terms within each head group class only as previously discussed. Several statistically significant compositional changes were observed, the greatest of which was a ~ 1.6 fold increase in the relative abundance of DGDG(32:2) and a ~ 0.07 fold decrease in DGDG(32:0). No overall trend

was apparent, however. Previous research has indicated that the total class level quantity per cell of MGDG and DGDG does not respond to P stress (**Martin et al.**, 2011). Despite this, the observation of compositional changes of the individual lipid molecular species indicates that P stress does impact upon the glyceroglycolipids of *T. pseudonana*.

The relative abundances of individual lipid species within pairs of lipid classes were compared by bivariate analyses in order to assess compositional similarity with respect to their diglyceride identities (Figure 2.7). The composition of DGCC lipid species was found to be similar to that of PC. This observation provides further evidence for the role of DGCC as a substitute lipid for PC in *T. pseudonana* (**Van Mooy et al.**, 2009; **Martin et al.**, 2011).

When comparing PC/DAG, no correlation between their molecular species compositions was observed. DAG is a known precursor in eukaryotes to glycerophospholipids (PC, PE)(**Lykidis**, 2007). Under P- conditions, an uncoupling of the DAG composition with the glycerophospholipids could be expected as they are not being synthesised. However, DAG composition did not vary between P+ and P- conditions as previously discussed. Therefore, we expect the composition of lipid species in the DAG to reflect that of PC, the most abundant of its biosynthetic products, under P+ conditions but that was not the case.

Furthermore, no correlation was observed between DGCC/DAG. As the precursor to PC biosynthesis, DAG is a possible candidate for a role in the unknown biosynthesis pathway of the chemically similar DGCC. This could occur via *de novo* synthesis or represent a recycling of the diglyceride from the observed PC breakdown via phospholipase C. Again, our observations do not support these hypotheses. As we know that PC is synthesised from DAG under P+ conditions, yet we cannot observe this via the correlation of its individual lipid species, a role of DAG in the biosynthesis of DGCC cannot be ruled out, based on these results.

In contrast, the observed DAG molecular composition was correlated with the glycolipids MGDG and DGDG. Like the phospholipids discussed previously, MGDG and DGDG are also derived from a DAG precursor (**Riekhof et al.**, 2005). Thus, it appears that the

observed DAG is utilised primarily by monogalactosyldiacylglycerol synthase as a biosynthetic precursor to MGDG and DGDG (**Armbrust et al.**, 2004; **Riekhof et al.**, 2005). These observations lead to the formation of the hypothesis that there are two (or more) separate pools of DAG. Firstly, the larger and/or slower turned over DAG pool indicative of MGDG/DGDG synthesis, observed by characterisation of the total lipid extract. Secondly, the smaller and/or more rapidly turned over DAG pool indicative of PC/PE synthesis that is conspicuously not observed. A preliminary investigation into this hypothesis is presented in Chapter 3.

2.6 Conclusions

We have presented findings consistent with previous reports on the response of the polar lipids of *T. pseudonana* to P stress, at the level of lipid class. Specifically, the phospholipids PC, PG and PE were observed to decrease in quantity per cell with increasing incubation time, while the betaine lipid substitute DGCC increased.

Beyond the previous knowledge, it was observed that P stress resulted in a net cessation of glycerophospholipid synthesis, with a minor degree of breakdown. The majority of the original glycerophospholipid remained intact and was diluted through culture growth by cell division. As a result, most of the P bound in the glycerophospholipid headgroups remains as such and is unavailable for other cellular processes as has been hypothesised. The broken down glycerophospholipid was equivalent to only a small proportion of the biosynthesised DGCC, suggesting phospholipid breakdown cannot form a major source of recycled diglyceride for DGCC biosynthesis.

Further insight was gained by investigation of the individual molecular species within each lipid headgroup class. During P-replete growth, significant variability was observed within the per cell quantities of PC molecular species, with the progression of time. The complexity and degree of these temporal changes highlight the plasticity of the lipidome, often overlooked in biomarker studies. During P-stress, the quantity per cell of individual molecular species within the PC, PG and PE classes decreased in line with the total class data. The magnitude of this decrease varied between individual PC species, notably the highly polyunsaturated PC(40:10) and PC(42:11) showed the smallest decreases. In contrast, several minor PE species, some of which appear to bear very long chain (C >22:6) polyunsaturated fatty acids, increased under P stress. Subject to further characterisation, these may have utility as molecular biomarkers.

Variation in the relative abundance of molecular species of the glyceroglycolipids MGDG and DGDG were observed. Therefore, despite previous observations on the stability of the total MGDG and DGDG quantity per cell to P stress, we find that there are implications upon the glyceroglycolipids.

Finally, the similarity of the lipid molecular species composition between lipid classes was assessed. The compositions of PC and DGCC were found to be similar, providing further evidence for the role of DGCC as a substitute for PC. The molecular composition of DAG, a common precursor to many of the polar lipid classes investigated, was distinct from PC and DGCC, being strongly correlated instead to the glyceroglycolipids MGDG and DGDG. Thus, the observed DAG appears to be utilised primarily by monogalactosyldiacylglycerol synthase as a precursor to MGDG and DGDG.

2.7 Experimental Procedures

All data were gathered in biological triplicate (*ie.* samples from three separate cultures) and were represented as the mean average and 1 standard deviation. Technical replicates (*ie.* repeat measurements) were not acquired. The precision of the targeted lipidomic measurements was accounted for by univariate statistics (unpaired, equal variance T-test). The quantification of glycerolipids by DI-ESI-MS/MS as herein, based upon internal standards within the same lipid headgroup class gives quantitative accuracy of approximately 5% (**Yang and Han**, 2011). The analytical variability (determined to be $\sim$ 5-20%) during the duration of the analysis was determined based upon quality control mixtures of standards as described in Supplementary Figure 7.1.

2.7.1 Culturing

Axenic culture of *T. pseudonana* (1085/12 also designated CCMP1335/3H) was obtained from the Culture Collection of Algae and Protozoa, Scottish Association for Marine Science, U.K. All culture manipulations were performed under a sterile, laminar flow environment. Borosilicate culture vessels were washed with methanol/chloroform before use. Solvents were HPLC grade (Fisher Scientific).

F/2+Si growth media (**Guillard**, 1975), based on artificial seawater was used (**Kester et al.**, 1967). All reagents were analytical or biological grade. Seed culture (175 mL) was grown to mid-log phase concentration of 1.13 x 10^6 cells mL^{-1} over 4 days incubation at 18°C; 12:12h light/dark cycle; 123mol quanta m^{-2} s^{-1} illumination; 70 rpm gentle orbital agitation.

The seed culture was split (2 x 79.6 mL) and cells isolated from the media by filtration (Millipore Steritop, 0.22 μm pore size). Cells were washed on the filter with 50 mL of P+/P- media, depending on treatment, then resuspended and split to form 3 x 300 mL for each P+ and P-. These experimental cultures were incubated as above and sampled after 0; 6; 12; 24; 48; 72 and 96 hours for size distribution/cell count; viability; dissolved/particulate macronutrients and lipid extracts.

2.7.2 Cell Size Distribution/Counting

An aliquot (950 μL) of culture was mixed with freshly prepared paraformaldehyde solution (170 μL, 34% w/v, dH_2O) and stored at 4°C for <24 h before analysis. Cell size distributions were generated with a Beckmann Coulter Multisizer 3 Coulter Counter. A 70μm aperture and 3% NaCl electrolyte were used and the samples diluted to ensure <10% aperture coincidence concentration. The Coulter Counter was calibrated with 5.023μm polystyrene latex standard beads prior to use (Beckmann Coulter via Meritics Ltd., Dunstable, U.K.). Size distributions were used to generate cell concentration values, between the limits of 3 and 9 μm particle diameter.

2.7.3 Viability Assay

Experimental culture (50μL) was incubated with SYTOX-Green dye (Invitrogen Life Technologies, Paisley, U.K.) at a concentration of 0.5μM for 5 minutes in the dark. 18 μL of this solution was then imaged with a Cellometer Vision Duo (Nexcelcom Bioscience via. Peqlab, Sarisbury Green, U.K; X100-F101 Optics; SD100 Slides). All cells were counted manually under brightfield mode and stained, non-viable cells under fluorescence mode (470/535nm excitation/emission).

2.7.4 Nutrient Quantification

Experimental culture (10 mL) was syringe filtered over pre-combusted (450°C, 12h) GF/F filters. The filtrate was stored at -20 °C. Filters were dried at 60°C for 24 h and stored in a desiccator. Particulate phosphorus was determined following an oxidation procedure as described in reference (**Raimbault et al.**, 1999), samples were centrifuged (1000 x G, 10 mins, 18 °C) before sampling to remove particulates. Nutrient samples were diluted: 1/60 (dissolved nutrients (filtrate)) and 3/20 (oxidised particulate nutrients) in milliQ dH_2O and characterised by segmented flow autoanalysis on an AutoAnalyzer 3 (Seal Analytical, Fareham, U.K.) for phosphorus, nitrate/nitrite and silicon.

2.7.5 Lipid Extraction

Lipid samples (20 mL) from the experimental cultures were isolated by syringe filtration as above. The filtrate was discarded and the filters stored at -78 °C until extraction. Total lipid extracts were prepared using a Bligh-Dyer extraction procedure (**Bligh and Dyer**, 1959) as modified by **Popendorf et al.** (2013). The isolated cells (on the filters) were transferred to PTFE capped glassware and resuspended in phosphate buffered saline (0.8 mL). Chloroform (1 mL) was added. Non-physiological internal standards were made up in 1:0.5 chloroform:methanol solution and added at quantities adjusted per sample to maintain constant standard:cell equivalents ratios as follows: PC(12:0/12:0)(56.2 amol cell^{-1}); PG(12:0/12:0)(77.0 amol cell^{-1}); PE(12:0/12:0)(47.6 amol cell^{-1}); DAG(20:4/18:0)(99.7 amol cell^{-1}) and SQDG (mixed extract, predominantly 34:3)(1.27 fmol cell^{-1}). Phospholipid/DAG standards were acquired from Avanti Polar Lipids (Alabaster, U.S.A.), SQDG spinach leaf extract was provided by Lipid Products (Surrey, U.K.).

Methanol (2 mL), chloroform (1 mL) and water (1 mL) were then added, vortexing thoroughly between each addition. The mixture was centrifuged (100 x G, 10 mins, 18°C) and the bottom chloroform phase isolated. A variable volume of this phase was isolated per sample ensuring a constant quantity of lipid cell equivalents/standard quantity was infused into the mass spectrometer during analysis (hence accounting for ion suppression and enabling external standard quantification of DGCC). The optimal quantity for analysis under instrument/conditions below was found to be 0.8 x 10^6 cells (+standards as described). This fraction was dried under N_2 and stored at -20°C until analysis.

2.7.6 ESI-MS/MS Analysis

Mass Spectrometric analysis was performed on a Waters Micromass Quattro Ultima triple quadrupole instrument. Dried samples were dissolved in 250μL 66% methanol, 30% dichloromethane, 4% ammonium acetate (300 mM in H_2O). The sample solution was directly infused into the instrument at 6μL min^{-1} and analysed as indicated in Table 2.1.

Spectra were processed by despiking, baseline subtraction, isotopic correction and assignment by a visual basic macro (**Postle et al.**, 2011). The spectrum of the SQDG

standard in isolation was taken in triplicate and used to perform a subtraction for overlapping minor peaks based on the dominant peak at 834 m/z (SQDG34:3, not detectable in *T. pseudonana* under P+/P- conditions).

MS	**MS/MS Mode**	**Lipid**	**Range (M/Z)**	**Molecular Ion**	**Collision Energy (eV)**
ES+	Full Positive	-	2:1300	$[M]^+$	-
ES-	Full Negative	-	2:1300	$[M]^-$	-
ES+	Precursors of 104	DGCC	450:1000	$[M]^+$	40
ES+	Precursors of 184	PC	450:1000	$[M]^+$	30
ES+	Neutral Loss of 35	DAG	350:750	$[M+NH4]^+$	15
ES+	Neutral Loss of 141	PE	450:900	$[M+H]^+$	25
ES+	Neutral Loss of 189	PG	450:900	$[M+NH4]^+$	25
ES+	Neutral Loss of 197	MGDG	450:1000	$[M+NH4]^+$	22
ES+	Neutral Loss of 261	SQDG	450:1000	$[M+NH4]^+$	25
ES+	Neutral Loss of 359	DGDG	600:1500	$[M+NH4]^+$	24

Table 2.1: Mass Spectrometry Analytical Conditions

2.7.7 DGCC External Standard Quantification

DGCC extract (purified by preparative HPLC from *T. pseudonana* under P starvation) was provided by Benjamin van Mooy (Woods Hole Oceanographic Institute, U.S.A.). External standard calibrations were generated from addition of DGCC extract (0.094; 0.19; 0.38; 0.75; 1.5 nmol) to P-replete grown (hence no intrinsic DGCC) *T. pseudonana* total lipid extracts prepared at optimal, constant cellular lipid/standard concentration as previously discussed. DGCC total counts were normalised to the chemically similar PC internal standard for quantification of experimental samples: $\text{Counts}_{\text{DGCC}}/\text{Counts}_{\text{PC}} = 18.594 \text{ x } \text{Quantity}_{\text{DGCC}}$. Highly linear over this range ($R^2 = 0.992$).

2.8 Acknowledgements

The authors would like to acknowledge John R. Gittins, Stephanie Tweed, Sophie Richier, Mark Stinchcombe and Victoria Goss for methodological assistance. We thank Benjamin van Mooy and Helen Fredricks (Woods Hole Oceanographic Institution, U.S.A.) for providing the DGCC standard. This work was funded by the University of Southampton - Vice Chancellors Scholarship Award. The purchase of the mass spectrometer was supported by the Wellcome Trust (Grant 057405).

2.9 References

Abida, H., Dolch, L.-J., Mei, C., Villanova, V., Conte, M., Block, M. A., et al. (2014), Membrane glycerolipid remodeling triggered by nitrogen and phosphorus starvation in Phaeodactylum tricornutum, *Plant Physiology*

Armbrust, E. V., Berges, J. A., Bowler, C., Green, B. R., Martinez, D., Putnam, N. H., et al. (2004), The genome of the diatom Thalassiosira pseudonana: ecology, evolution, and metabolism., *Science*

Bellinger, B. J., Van Mooy, B. A. S., Cotner, J. B., Fredricks, H. F., Benitez-Nelson, C. R., Thompson, J., et al. (2014), Physiological modifications of seston in response to physicochemical gradients within Lake Superior, *Limnol. Oceanogr*

Benning, C., Huang, Z. H., and Gage, D. A. (1995), Accumulation of a novel glycolipid and a betaine lipid in cells of Rhodobacter sphaeroides grown under phosphate limitation., *Archives of Biochemistry and Biophysics*

Bligh, E. G. and Dyer, W. J. (1959), A Rapid Method of Total Lipid Extraction and Purification, *Canadian Journal of Biochemistry and Physiology*

Brandsma, J., Hopmans, E. C., Philippart, C. J. M., Veldhuis, M. J. W., Schouten, S., and Sinninghe Damsté, J. S. (2012), Low temporal variation in the intact polar lipid composition of North Sea coastal marine water reveals limited chemotaxonomic value, *Biogeosciences*

Duszenko, M., Ginger, M. L., Brennand, A., Gualdrón-López, M., Colombo, M.-I., Coombs, G. H., et al. (2011), Autophagy in protists, *Autophagy*

Dyhrman, S. T., Jenkins, B. D., Rynearson, T. A., Saito, M. A., Mercier, M. L., Alexander, H., et al. (2012), The Transcriptome and Proteome of the Diatom Thalassiosira pseudonana Reveal a Diverse Phosphorus Stress Response, *PLoS ONE*

Eichenberger, W. and Gribi, C. (1997), Lipids of Pavlova lutheri: Cellular site and metabolic role of DGCC, *Phytochemistry*

Fahy, E., Subramaniam, S., Brown, H. A., Glass, C. K., Merrill, A. H., Murphy, R. C., et al. (2005), A comprehensive classification system for lipids, *Journal of Lipid Research*

Guillard, R. R. L. (1975), Culture of phytoplankton for feeding marine invertebrates., in W. L. Smith and M. H. Chanley, eds., Culture of Marine Invertebrate Animals (Plenum Press, New York, USA.),

Guschina, I. A. and Harwood, J. L. (2006), Lipids and lipid metabolism in eukaryotic algae, *Progress in Lipid Research*

Hodgson, P., Henderson, R., Sargent, J., and Leftley, J. (1991), Patterns of variation in the lipid class and fatty acid composition of Nannochloropsis oculata (Eustigmatophyceae) during batch culture, *Journal of Applied Phycology*

Hoek, C., Mann, D. G., and Jahns, H. M. (1995), Algae, (Cambridge University Press)

Kester, D. R., Duedall, I. W., Connors, D. N., and Pytkowicz, R. M. (1967), Preparation of Artificial Seawater, *Limnology and Oceanography*

Levitan, O., Dinamarca, J., Hochman, G., and Falkowski, P. G. (2014), Diatoms: a fossil fuel of the future, *Trends in Biotechnology*

Liebisch, G., Vizcaíno, J. A., Köfeler, H., Trötzmüller, M., Griffiths, W. J., Schmitz, G., et al. (2013), Shorthand notation for lipid structures derived from mass spectrometry, *Journal of Lipid Research*

Liu, B. and Benning, C. (2013), Lipid metabolism in microalgae distinguishes itself., *Current Opinion in Biotechnology*

Lykidis, A. (2007), Comparative genomics and evolution of eukaryotic phospholipid biosynthesis, *Progress in Lipid Research*

Martin, P., Van Mooy, B. A. S., Heithoff, A., and Dyhrman, S. T. (2011), Phosphorus supply drives rapid turnover of membrane phospholipids in the diatom Thalassiosira pseudonana, *The ISME Journal*

Mather, R. L., Reynolds, S. E., Wolff, G. A., Williams, R. G., Torres-Valdes, S., Woodward, E. M. S., et al. (2008), Phosphorus cycling in the North and South Atlantic Ocean subtropical gyres, *Nature Geoscience*

Popendorf, Tanaka, T., Pujo-Pay, M., Lagaria, A., Courties, C., Conan, P., et al. (2011), Gradients in intact polar diacylglycerolipids across the Mediterranean Sea are related to phosphate availability, *Biogeosciences*

Popendorf, K. J., Fredricks, H. F., and Van Mooy, B. A. (2013), Molecular Ion-Independent Quantification of Polar Glycerolipid Classes in Marine Plankton Using Triple Quadrupole MS, *Lipids*

Postle, A. D., Henderson, N. G., Koster, G., Clark, H. W., and Hunt, A. N. (2011), Analysis of lung surfactant phosphatidylcholine metabolism in transgenic mice using stable isotopes, *Chemistry and Physics of Lipids*

Raimbault, P., Diaz, F., and Pouvesle, W. (1999), Simultaneous determination of particulate organic carbon, nitrogen and phosphorus collected on filters, using a semi-automatic wet-oxidation method, *Marine Ecology Progress Series*

Rezanka, T., Nedbalova, L., and Sigler, K. (2008), Identification of very-long-chain polyunsaturated fatty acids from Amphidinium carterae by atmospheric pressure chemical ionization liquid chromatographymass spectroscopy, *Phytochemistry*

Rezanka, T. and Sigler, K. (2009), Odd-numbered very-long-chain fatty acids from the microbial, animal and plant kingdoms, *Progress in Lipid Research*

Riekhof, W. R., Sears, B. B., and Benning, C. (2005), Annotation of Genes Involved in Glycerolipid Biosynthesis in Chlamydomonas reinhardtii: Discovery of the Betaine Lipid Synthase BTA1Cr, *Eukaryotic Cell*

Rigden, D. J., Michels, P., and Ginger, M. L. (2009), Autophagy in protists: examples of secondary loss, lineage-specific innovations, and the conundrum of remodeling a single mitochondrion, *Autophagy*

Sanudo-Wilhelmy, S. A., Kustka, A. B., Gobler, C. J., Hutchins, D. A., Yang, M., Lwiza, K., et al. (2001), Phosphorus limitation of nitrogen fixation by Trichodesmium in the central Atlantic Ocean, *Nature*

Schatz, D., Shemi, A., Rosenwasser, S., Sabanay, H., Wolf, S. G., Ben-Dor, S., et al. (2014), Hijacking of an autophagy-like process is critical for the life cycle of a DNA virus infecting oceanic algal blooms, *New Phytologist*

Schubotz, F., Wakeham, S. G., Lipp, J. S., Fredricks, H. F., and Hinrichs, K.-U. (2009), Detection of microbial biomass by intact polar membrane lipid analysis in the water column and surface sediments of the Black Sea., *Environmental Microbiology*

Siegenthaler, P.-A. and Murata, N. (2004), Lipids in Photosynthesis: Structure, Function and Genetics (Springer Netherlands)

Thingstad, T. F., Krom, M. D., Mantoura, R. F. C., Flaten, G. A. F., Groom, S., Herut, B., et al. (2005), Nature of phosphorus limitation in the ultraoligotrophic eastern Mediterranean., *Science*

Van Mooy, B. A. S. and Fredricks, H. F. (2010), Bacterial and eukaryotic intact polar lipids in the eastern subtropical South Pacific: Water-column distribution, planktonic sources, and fatty acid composition, *Geochimica et Cosmochimica Acta*

Van Mooy, B. A. S., Fredricks, H. F., Pedler, B. E., Dyhrman, S. T., Karl, D. M., Koblížek, M., et al. (2009), Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity, *Nature*

Vance, J. E. and Vance, D. E. (2008), Biochemistry of Lipids, Lipoproteins and Membranes (Elsevier)

Volkman, J. K., Jeffrey, S. W., Nichols, P. D., Rogers, G. I., and Garland, C. D. (1989), Fatty acid and lipid composition of 10 species of microalgae used in mariculture, *Journal of Experimental Marine Biology and Ecology*

Wakeham, S. G., Hedges, J. I., Lee, C., Peterson, M. L., and Hernes, P. J. (1997), Compositions and transport of lipid biomarkers through the water column and surficial sediments of the equatorial Pacific Ocean, *Deep Sea Research Part II: Topical Studies in Oceanography*

Wu, J., Sunda, W., Boyle, E. A., and Karl, D. M. (2000), Phosphate Depletion in the Western North Atlantic Ocean, *Science*

Yang, K. and Han, X. (2011), Accurate Quantification of Lipid Species by Electrospray Ionization Mass Spectrometry Meets a Key Challenge in Lipidomics, *Metabolites*

Zhukova, N. V. (2004), Changes in the Lipid Composition of Thalassiosira pseudonana during Its Life Cycle, *Russian Journal of Plant Physiology*

Chapter 3

Untargeted Lipidomic Characterisation of the Marine Diatom *Thalassiosira Pseudonana* Subject to Phosphorus Stress - Increase in Diglycosylceramides under Low Phosphorus Conditions

Jonathan E. Hunter[1,2], Joost Brandsma[3], Jamie Burell[4], Marcus Dymond[5], Grielof Koster[3], C. Mark Moore[1], Anthony D. Postle[3], George S. Attard[4] and Rachel A. Mills[1].

1. Ocean and Earth Science, University of Southampton, National Oceanography Centre Southampton, European Way, Southampton, SO14 3ZH, United Kingdom
2. Institute for Life Sciences, University of Southampton, SO17 1BJ, United Kingdom
3. Faculty of Medicine, University of Southampton, Southampton General Hospital, Tremona Road, Southampton, SO16 6YD, United Kingdom
4. Chemistry, University of Southampton, Southampton, SO17 1BJ, United Kingdom
5. School of Pharmacy and Biomolecular Sciences, University of Brighton, Brighton, BN2 4GJ, United Kingdom

3.1 Author Contributions

Jonathan E. Hunter designed the experiments, developed methods, carried out culturing, sample preparation, analysis, data interpretation and wrote the manuscript. Jamie Burell developed the subcellular fractionation method used to isolate a plastid fraction. The remaining co-authors assisted with experimental design, interpretation and drafting the manuscript.

3.2 Abstract

Unicellular organisms remodel lipids in response to environmental conditions, such as the substitution of glycerophospholipids with non-phosphorus alternatives under phosphorus (P) stress. The marine diatom *Thalassiosira Pseudonana* (*T. pseudonana*) has been used as a model system to study glycerophospholipid substitution dynamics by targeted lipidomic analyses. However, the behaviour of minor lipid species under P stressed conditions, outside of the major glycerolipid classes, remains unknown. Furthermore, previous reports have not characterised the fatty acid composition of each individual diglyceride lipid species.

We report herein, the untargeted analysis of the lipidome of *T. pseudonana*, subject to P stress, by ultra performance liquid chromatography - quadrupole time of flight mass spectrometry (UPLC-Q-ToF-MS). Several diacylglycerylcarboxyhydroxymethylcholine (DGCC) chemotypes increased strongly under P stress in line with previous observations. A triacylglycerol species, present only at trace levels under other stress regimes was also increased subject to P stress.

A group of diglycosylceramides, not previously detected in *T. pseudonana*, were linked with P stress. These species may prove useful as lipid biomarkers for P starvation in diatoms, with potential applications in the elucidation of P related biogeochemical processes.

Targeted MS2 analyses were used to characterise the fatty acyl compositions of each of the major diglyceride lipid species. This provides new insight into the *T. pseudonana* lipidome filling a conspicuous gap in an otherwise well characterised system.

Finally, a preliminary qualitative study of the subcellular partitioning of precursor DAG lipids supported the hypothesis that DAG lipids were comprised of two subcellular pools: one a putative MGDG/DGDG precursor pool localised to the chloroplast, the other a putative PC/PE precursor localised elsewhere.

Keywords: Thalassiosira Pseudonana, Phosphorus Stress, Untargeted Lipidomics, Lipids.

3.3 Introduction

Oceanic microbial primary producers, also known as phytoplankton, are responsible for the fixation of $\sim$ 45 gigatons of organic carbon per annum (**Falkowski et al.**, 1998), a significant proportion of which is utilised in the production of cellular lipids. For example, in the equatorial Pacific Ocean $\sim$ 11-23% of the total planktonic organic carbon is comprised by lipids (**Wakeham et al.**, 1997). Lipids are a diverse group of organic biomolecules defined by their "hydrophobic [and/or] amphipathic (bearing both polar and apolar moieties)" character (**Fahy et al.**, 2005). They have range of physiological functions from structural and energy storage, to involvement in signalling and membrane functionality (**Mouritsen**, 2011).

Unicellular organisms commonly remodel their lipidome (the entirety of its cellular lipids) in response to environmental conditions (**Benning et al.**, 1995; **Van Mooy et al.**, 2009; **Martin et al.**, 2011). One such remodelling involves the substitution of glycerophospholipids with non-phosphorus alternatives under phosphorus (P) stress (**Benning et al.**, 1995; **Van Mooy et al.**, 2009; **Martin et al.**, 2011), allowing the phytoplankter to reduce its demand for P (**Van Mooy et al.**, 2009). The dynamics of glycerophospholipid substitution in phytoplankton have been well characterised both in laboratory culture studies ((**Benning et al.**, 1993; **Van Mooy et al.**, 2009; **Martin et al.**, 2011) and Chapter 2) and in the marine environment (**Van Mooy et al.**, 2006; **Schubotz et al.**, 2009; **Van Mooy et al.**, 2009; **Popendorf et al.**, 2011b,a).

Lipidomics seeks to quantitatively describe the entirety of the cellular lipid complement (the lipidome), how this changes with time, as well as the functions of individual species or classes (**Blanksby and Mitchell**, 2010). In the past two decades, lipidomics has been driven by insights into the importance of lipid-lipid and lipid-protein interactions and the availability of powerful analytical tools, in particular mass spectrometry (**van Meer**, 2005; **Wenk**, 2005; **Guschina and Harwood**, 2006; **Ivanova et al.**, 2009; **Blanksby and Mitchell**, 2010). The study of phytoplankton lipids is no exception and much insight has been gained into their diversity and metabolism (**Guschina and Harwood**, 2006). Mass

spectrometry based lipidomics often employs two distinct strategies, targeted lipidomics for the quantification of a subset of lipids that are commonly already known, and untargeted (or global) lipidomics which aims to characterise all of the lipids in a system (**Han**, 2009; **Blanksby and Mitchell**, 2010).

The marine diatom *Thalassiosira Pseudonana* (*T. pseudonana*) has been used as a model system to study glycerophospholipid substitution dynamics at the level of the total within each of the major glycerolipid classes (**Martin et al.**, 2011). The dynamics of substitution at the level of the individual lipid species within each lipid glycerolipid class, varying in fatty acid substituents, have also been reported (Chapter 2). Previous work consists exclusively of targeted lipidomic analyses and therefore the behaviour of minor lipid species under P stressed conditions, outside of the major glycerolipid classes, remains unknown. Furthermore, previous reports have not characterised the fatty acid composition of each individual diglyceride lipid species.

In addition, previous observations showed that the composition of individual lipid species within the precursor lipid diacylglycerol (DAG) was highly correlated with that of the glyceroglycolipids mono/di-galactosyldiacylglycerol (M/DGDG) (Chapter 2). This suggests that the observed DAG pool was being utilised for the synthesis of MGDG and DGDG, thought to be associated primarily with the chloroplast (**Siegenthaler and Murata**, 2004). DAG is also expected to be a putative precursor to synthesis of the extraplastidic lipids glycerophosphatidylcholine (PC) and glycerophosphatidylethanolamine (PE) in eukaryotic phytoplankton (**Riekhof et al.**, 2005) but no comparable relationship was observed. (Chapter 2). This uncoupling of DAG with the abundant glycerophospholipids implies two distinct pools of DAG, of which only the plastid pool is being observed in the total lipid extract. This hypothesis is in line with genetic models of lipid metabolism (**Armbrust et al.**, 2004; **Riekhof et al.**, 2005) but has not been observed directly in *T. pseudonana*.

We report herein, the untargeted analysis of the lipidome of *T. pseudonana*, subject to P stress, by ultra performance liquid chromatography - quadrupole time of flight mass spectrometry (UPLC-Q-ToF-MS). Minor lipid species indicative of P stress, including a

group of diglycosylceramides, not previously reported in *T. pseudonana* are presented. Furthermore, semi-targeted data dependant MS2 analyses were used to characterise the fatty acid composition of the major diglyceride lipids. Finally, a preliminary analysis of a purified plastid fraction yielded insight into the partitioning of different DAG chemotypes between the plastid fraction and the whole cell.

3.4 Results

3.4.1 Untargeted Lipidomic Screening of Phosphrous Stressed *T. pseudonana*

Following data processing by peak picking and filtering, a total of 998 and 163 distinct ions were observed in positive and negative ion mode respectively (pooled from both P+ and P- samples). Ions were quantified by differential abundance between the P- and P+ cultures resulting in fold changes, but no absolute quantities, and ranked as shown in (Figure 3.1).

A

	Normalised Abundance:					Database Assignment and Corroborating MS2 Evidence:		
MZ (Da)	R.T.	P Val.	P+	P-	Adduct	Assignment	PPM Diff.	MS2 Fragments
774.5897	11.5	1.04E-05	0.00	1.00	$(M+H)^+$	DGCC(20:5/16:0)	2.39	D184, 20:5 and 16:0 NL Fatty Ketene
536.3573	5.0	2.19E-03	0.00	1.00	$(M+H)^+$	LDGCC(20:5)	-1.57	D104, 20:5 NL Fatty Ketene
800.6057	11.9	5.98E-05	0.00	1.00	$(M+H)^+$	DGCC(22:6/16:0)	2.71	D104, 22:6 and 16:0 NL Fatty Ketene
562.3729	5.4	8.74E-03	0.00	1.00	$(M+H)^+$	LDGCC(22:6)	-1.65	D104, 22:6 NL Fatty Ketene
820.5735	10.6	1.02E-04	0.00	1.00	$(M+H)^+$	DGCC(20:5/20:5)	1.59	D104, 20:5 NL Fatty Ketene
846.5897	11.0	1.48E-04	0.00	1.00	$(M+H)^+$	DGCC(22:6/20:5)	2.20	D104, 22:6 NL Fatty Ketene
772.5719	10.8	2.91E-03	0.00	1.00	$(M+H)^+$	DGCC(20:5/16:1)	-0.32	D104, 20:5 NL Fatty Ketene
748.5718	11.0	1.81E-03	0.00	1.00	$(M+H)^+$	DGCC(18:4/16:0)	-0.50	D104, 18:4 and 16:0 NL Fatty Ketene
726.5891	11.7	8.79E-04	0.00	1.00	$(M+H)^+$	DGCC(16:1/16:0)	1.68	D104, 16:1 and 16:0 NL Fatty Ketene
746.5589	10.5	2.22E-03	0.00	1.00	$(M+H)^+$	DGCC(20:5/14:0)	3.10	D104, 20:5 and 14:0 NL Fatty Ketene
510.3410	4.8	9.59E-03	0.00	1.00	$(M+H)^+$	LDGCC(18:4)	-3.04	D104, 18:4 NL Fatty Ketene
724.5720	11.1	4.12E-03	0.00	1.00	$(M+H)^+$	DGCC(16:1/16:1)	-0.20	D104, 16:1 NL Fatty Ketene
796.7424	20.7	7.97E-04	0.00	1.00	$(M+NH_4)^+$	TAG(16:0/16:0/14:0)	4.41	16:0 and 14:0 NL FA+NH3
798.5871	11.1	1.03E-03	0.00	1.00	$(M+H)^+$	DGCC(20:5/18:2)	-0.96	D104, 20:5 and 18:2 NL Fatty Ketene
794.5575	10.1	2.88E-02	0.00	1.00	$(M+H)^+$	DGCC(20:5/18:4)	1.23	D104, 20:5 NL Fatty Ketene

B

	Normalised Abundance:					Database Assignment and Corroborating MS2 Evidence:		
MZ (Da)	R.T.	P Val.	P+	P-	Adduct	Assignment	PPM Diff.	MS2 Fragments
818.5771	11.5	4.41E-04	0.00	1.00	$(M+HAc-CH_3)^-$	DGCC(20:5/16:0)	-2.18	20:5 and 16:0 FA Daughter Fragments
844.5949	11.9	7.12E-04	0.00	1.00	$(M+HAc-CH_3)^-$	DGCC(22:6/16:0)	0.49	22:6 and 16:0 FA Daughter Fragments
890.5801	11.1	1.77E-04	0.00	1.00	$(M+HAc-CH_3)^-$	DGCC(22:6/20:5)	1.42	20:5 FA Daughter Fragment
770.5778	11.6	1.87E-03	0.00	1.00	$(M+HAc-CH_3)^-$	DGCC(16:1/16:0)	-1.39	16:1 and 16:0 FA Daughter Fragments
816.5621	10.8	2.20E-05	0.00	1.00	$(M+HAc-CH_3)^-$	DGCC(20:5/16:1)	-1.35	20:5 and 16:1 FA Daughter Fragments
792.5624	11.1	4.49E-04	0.00	1.00	$(M+HAc-CH_3)^-$	DGCC(18:4/16:0)	-1.02	18:4 and 16:0 FA Daughter Fragments
1174.7555	12.9	1.30E-05	0.00	1.00		Unknown		
1014.7112	13.6	6.17E-03	0.00	1.00		$(Gly)_2Cer(d18:3/24:0)$	1.34	See Figure 3
532.3505	5.1	4.67E-02	0.00	1.00	$(M+HAc-CH_3)^-$	LDGCC(16:1)	2.49	16:1 FA Daughter Fragment
796.5931	11.9	3.46E-04	0.00	1.00	$(M+HAc-CH_3)^-$	DGCC(18:2/16:0)	-1.72	18:2 and 16:0 FA Daughter Fragments
595.5395	12.7	4.14E-03	0.00	1.00		Unknown		
838.5529	10.1	4.59E-03	0.00	1.00	$(M+HAc-CH_3)^-$	DGCC(20:5/18:4)	6.38	Coelution with Identified +ve Ion
623.5695	13.6	8.15E-03	0.14	0.86		Unknown		
819.5296	11.3	3.72E-02	0.22	0.78	$(M-H)^-$	SQDG(34:1)O0	-0.26	DB Match and Diagnostic R.T. Only
694.6305	15.4	1.28E-02	0.27	0.73		Unknown		

Figure 3.1: Figure legend overleaf.

Figure 3.1: Untargeted screen of the *T. pseudonana* lipidome subject to P stress. Detected ions were ranked based upon normalised differential abundance $Quantity_{P-}/(Quantity_{P+} + Quantity_{P-})$ therefore those at the top are most strongly increased subject to P stress. Panel A displays positive ions and Panel B displays negative ions. R.T. represents chromatographic retention time, P value was determined by unpaired, two sample equal variance T-test. Assignment represents the lipid identity, PPM Diff. the difference between the observed and predicted M/Z, and MS2 fragments outlines the observed fragments under AutoMS2 in support of the designated assignment. L- refers to a lyso-species (bearing 1 rather than 2 fatty acids) and NL = neutral loss. Data represent the mean of biological triplicate samples. Assignments represent the primary fatty acyl configuration, as determined by the abundance of the fatty acyl fragments in the MS2 spectra.

Lipid species were ranked according to their differential abundance, in descending order from the most strongly increased under P stress (Figure 3.1). The betaine lipid diacylglycerylcarboxyhydroxymethylcholine (DGCC) dominated the positive ion results comprising 14 of the top 15 ions. DGCC species containing eicosapentaenoic (20:5), docosahexaenoic (22:6) and palmitic (16:0) fatty acids, displayed the greatest differential increases and were hence the highest ranked. As indicated by the normalised abundance equal to one in the P- case, all of the top 15 lipid species had a presence absence relationship between P- and the P+ control cultures.

The top five species, in descending order, were: DGCC(20:5/16:0), LDGCC(20:5), DGCC(22:6/16:0), LDGCC(22:6) and DGCC(20:5/20:5). TAG(16:0/16:0/14:0) was the exception to the dominance of DGCC in the top 15 lipid species in positive ion mode and also displayed a presence/absence increase subject to P stress.

In negative ion mode: DGCC(20:5/16:0), DGCC(22:6/16:0), DGCC(22:6/20:5), DGCC(16:1/16:0), DGCC(20:5/16:1) and DGCC(18:4/16:0) occupied the top six ranks. Four unknown ions, whose identities could not be discerned by these methods, were present. SQDG(34:1) displayed a modest increase of 3.47 $\pm$ 3.25 fold under P- conditions.

Most remarkably, a 1014.7112 Da species, corresponding to a diglycoceramide with a dihydroxy(18:3) long chain base and a 24:0 fatty amide ($(Gly)_2$Cer(d18:3/24:0), Figure 3.2A) with a formic acid adduct, displayed an absence presence response to P stress. This lipid species, assigned based upon its MS2 fragmentation spectra in both negative and positive ion mode, is displayed in Figure 3.2B and C respectively.

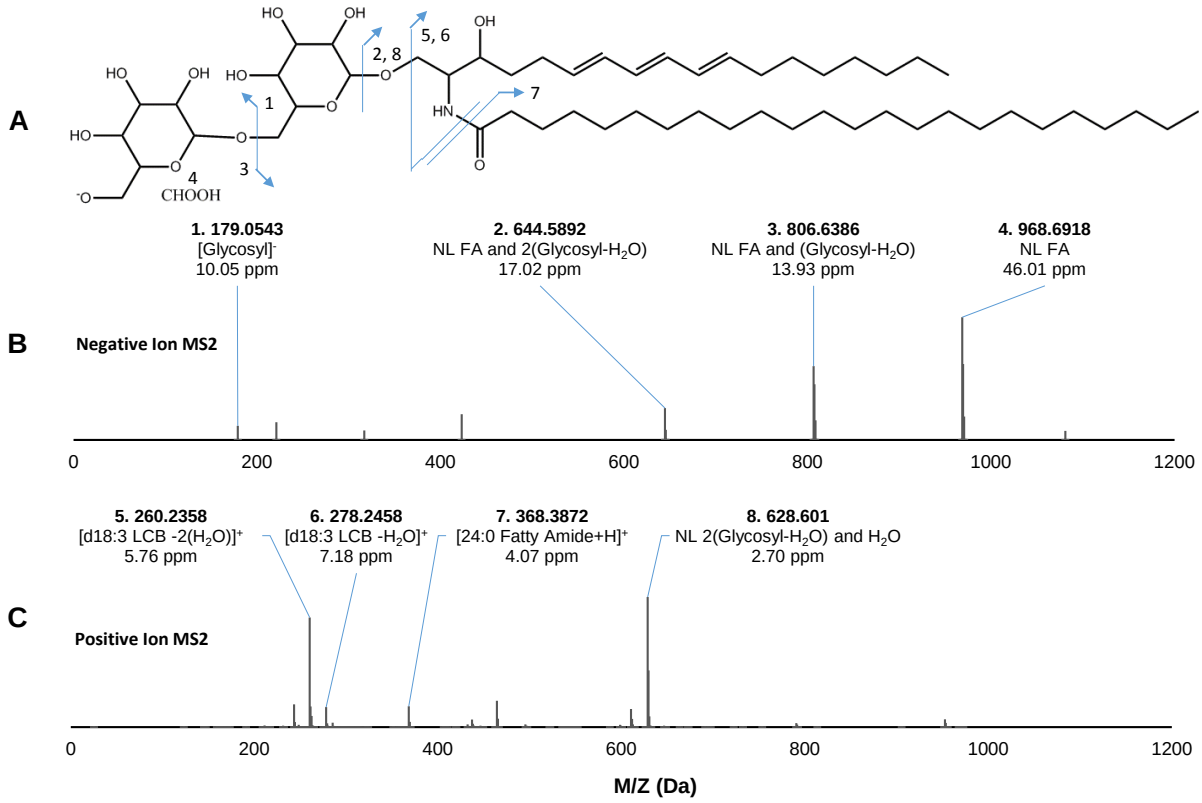


Figure 3.2: Chemical structure assignment of $(Gly)_2Cer(d18:3/24:0)$ (A) and supporting MS2 fragmentation data in negative (B) and positive (C) ion mode. Spectra from a single representative P- sample. Hydroxyl group and unsaturation regio- and stereochemistry was not resolved. Arrows indicate fragmentation and the direction of the charged ion position.

The negative ion MS2 (Figure 3.2B) revealed fragment ions of 968.6918, 806.6386, 644.5892 and 179.0543 Da, corresponding to a neutral loss of the formic acid (CHOOH) adduct, neutral loss of CHOOH and a glycosyl-H_2O, neutral loss of CHOOH and two glycosyl-H_2O units and a glycosyl fragment respectively. Knowledge of the CHOOH adduct and the retention time was used to identify the $[M+H]^+$ equivalent in positive ion mode and its respective MS2 fragmentation spectrum.

In positive ion mode, MS2 fragment ions of 628.6010, 368.3872, 278.2458 and 260.2358 corresponded to a neutral loss of two glycosyl moieties, a 24:0 fatty amide fragment ion and a d18:3 long chain base minus one and two H_2O respectively.

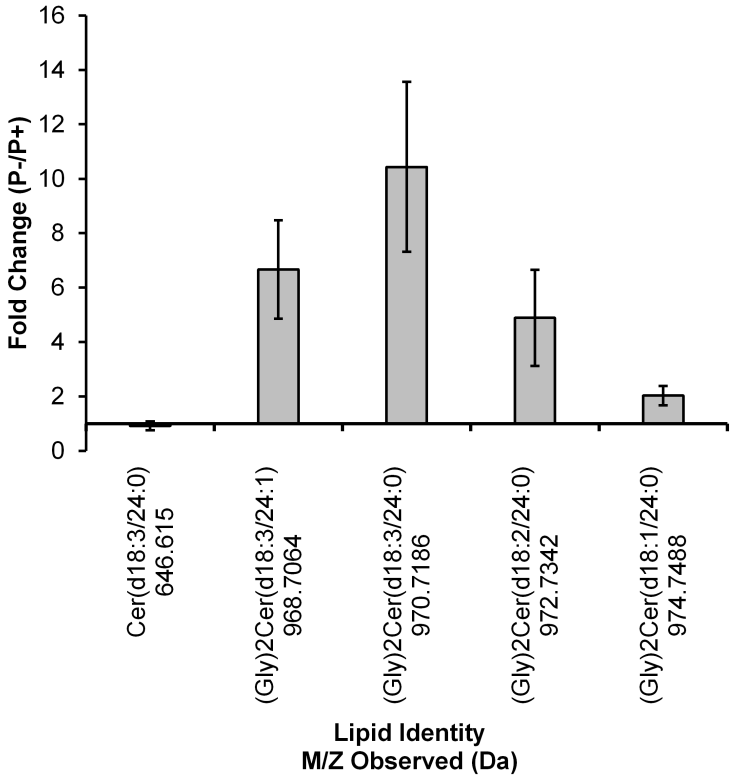


Figure 3.3: Fold change, subject to P stress, in observed ions related to $(Gly)_2Cer(d18:3/24:0)$, in positive ion mode. Long chain base and fatty acid amide fragments were observed in support of each of the following assignments in positive MS2. Data represent the mean of biological triplicate samples. Assignments represent the primary fatty acyl configuration, as determined by the abundance of the fatty acyl fragments in the MS2 spectra. Statistical significance is indicated in accompanying text, *$p<0.05$ **$p<0.005$ by two-tailed, paired equal-variance T-test.

Of the observed ions related to the novel $(Gly)_2Cer(d18:3/24:0)$ lipid species (Figure 3), it responded the strongest to P stress with a 10.43 ± 3.12** fold increase between the P- and P+ cultures. Each of the glycosphingolipids of the same type, displayed a statistically significant increase subject to P stress, albeit of varying magnitudes. $(Gly)_2Cer(d18:3/24:1)$, $(Gly)_2Cer(d18:2/24:0)$ and $(Gly)_2Cer(d18:1/24:0)$ increased by 6.67 ± 1.81**, 4.89 ± 1.76** and 2.03 ± 0.36** fold respectively, between the P- and P+ cultures.

The ceramide equivalent lipid Cer(d18:3/24:0), without the diglycosyl headgroup, was also detected. In contrast, its differential abundance between P- and P+ cultures showed

no significant change.

3.4.2 Fatty Acyl Composition of the Major Individual Glycerolipid Species

Lipid (Sum FA)	Primary	Secondary	Lipid (Sum FA)	Primary	Secondary
LDGCC(16:1)	16:1		PC(34:4)	18:4/16:0	
LDGCC(16:0)	16:0		DGCC(34:1)	18:1/16:0	
LMGDG(16:1)	16:1		SQDG(28:0)	14:0/14:0	
LDGCC(18:4)	18:4		PC(34:3)	18:3/16:0	18:2/16:1
LPC(18:4)	18:4		PC(34:2)	18:2/16:0	18:1/16:1
LPE(22:6)	22:6		PG(34:4)	18:4/16:0	18:3/16:1
LDGCC(20:5)	20:5		PC(34:1)	18:1/16:0	
LPC(20:5)	20:5		PG(34:3)	18:2/16:1	18:3/16:0
LSQDG(14:1)	14:1		MGDG(34:7)	18:4/16:3	
DAG(30:1)	16:1/14:0		PG(34:1)	18:1/16:0	18:0/16:1
LDGCC(22:6)	22:6		SQDG(29:0)	15:0/14:0	
LPC(22:6)	22:6		DGCC(36:7)	20:5/16:2	
DAG(32:2)	16:2/16:0	16:1/16:1	DGCC(36:6)	20:5/16:1	
DAG(32:1)	16:1/16:0		PC(36:8)	18:4/18:4	
DAG(32:0)	16:0/16:0	18:0/14:0	DGCC(36:5)	20:5/16:0	
DAG(34:4)	18:4/16:0		PC(36:7)	20:5/16:2	
DAG(34:2)	18:2/16:0		PC(36:6)	20:5/16:1	
DAG(34:1)	18:1/16:0	18:0/16:1	PC(36:5)	20:5/16:0	18:4/18:1
DAG(34:0)	18:0/16:0		SQDG(30:1)	16:1/14:0	
DAG(36:2)	18:2/18:0	18:1/18:1	PE(40:10)	20:5/20:5	
DAG(36:0)	18:0/18:0		PG(36:6)	20:5/16:1	
PE(32:2)	16:1/16:1		SQDG(30:0)	16:0/14:0	
DGCC(30:1)	16:1/14:0		PG(36:5)	20:5/16:0	
PC(30:1)	16:1/14:0		DGCC(38:9)	20:5/18:4	
PG(30:1)	16:1/14:0		SQDG(31:1)	16:1/15:0	
PG(30:0)	16:0/14:0		DGCC(38:8)	20:5/18:3	
MGDG(30:3)	16:3/14:0		DGCC(38:7)	20:5/18:2	
MGDG(30:1)	16:1/14:0		PC(38:9)	20:5/18:4	
DGCC(32:4)	18:4/14:0		SQDG(32:4)	18:4/14:0	
DGCC(32:2)	16:1/16:1		PC(38:7)	20:5/18:2	
PC(32:4)	16:3/16:1	18:4/14:0	PC(38:6)	20:5/18:1	
DGCC(32:1)	16:1/16:0		SQDG(32:2)	16:1/16:1	
PC(32:3)	16:3/16:0		PE(42:11)	22:6/20:5	
PC(32:2)	16:2/16:0	16:1/16:1	SQDG(32:1)	16:1/16:0	
PC(32:1)	16:1/16:0		SQDG(32:0)	16:0/16:0	
MGDG(31:0)	16:3/16:4		DGCC(40:10)	20:5/20:5	
PE(36:6)	20:5/16:1		PC(40:10)	20:5/20:5	
MGDG(32:6)	16:3/16:3	16:4/16:2	SQDG(34:4)	18:4/16:0	
PG(32:2)	16:1/16:1		SQDG(34:2)	18:2/16:0	
MGDG(32:5)	16:3/16:2	16:4/16:1	SQDG(34:0)	18:0/16:0	
PG(32:1)	16:1/16:0		DGCC(42:11)	22:6/20:5	
MGDG(32:4)	18:4/14:0	16:3/16:1	PC(42:11)	22:6/20:5	
MGDG(32:3)	16:3/16:0	16:2/16:1	SQDG(36:6)	18:3/18:3	
MGDG(32:2)	16:1/16:1	16:2/16:0	SQDG(36:5)	18:3/18:2	
DGCC(34:5)	20:5/14:0		SQDG(36:4)	18:3/18:1	
MGDG(32:1)	16:1/16:0		DGDG(30:1)	16:1/14:0	
DGCC(34:4)	18:4/16:0		DGDG(32:3)	16:2/16:1	
PC(34:7)	18:4/16:3		DGDG(32:2)	16:1/16:1	16:2/16:0
DGCC(34:3)	18:3/16:0		DGDG(32:1)	16:1/16:0	
PC(34:5)	18:4/16:1		DGDG(36:7)	20:5/16:2	
DGCC(34:2)	18:2/16:0				

Figure 3.4: Fatty acid (FA) composition of individual glycerolipid species in *T. pseudonana* in the P+ samples. DGCC species were analysed based upon the P- samples. Primary and secondary FA compositions represent a qualitative assessment of the major fragments in each MS2 spectrum.

The fatty acid compositions of the common glycerolipids in *T. pseudonana* are displayed in Figure 3.4 and were determined by a mixture of autoMS2 and targeted MS2 analyses. Glycerolipid species were comprised of 14:0, 14:1, 16:0, 16:1, 16:2, 16:3, 18:1, 18:2, 18:3, 18:4, 20:5 and 22:6 fatty acids. Previous work observed that trace lipid species, giving a neutral loss of 141 under low resolution MS2, differentially increased under P stress (Chapter 2). In this study, we made efforts to identify these lipid species, using the same samples used in that report. However, we were unable to identify any ions based upon the predicted accurate masses of the PE species (PE(29:5), PE(52:12), PE(54:13)). Furthermore, none of the automatically acquired MS2 spectra contained both a low resolution mass match to the reported species and a high resolution neutral loss of 141.0191.

3.4.3 Subcellular Partitioning of Polyunsaturated DAG Lipid Species

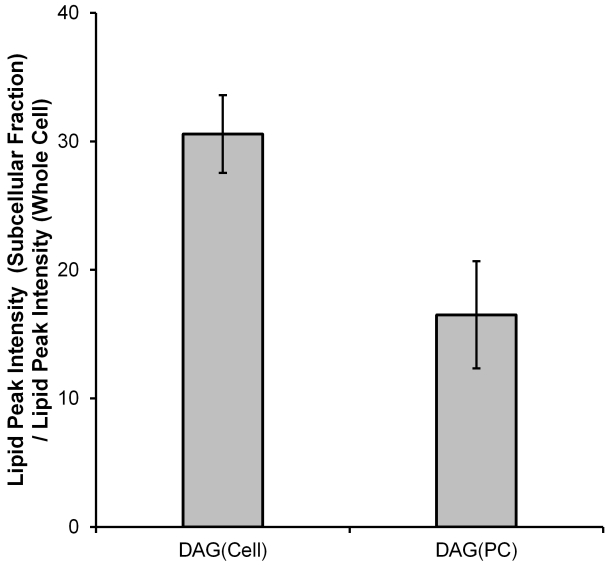


Figure 3.5: The ratio between the lipid peak intensity of a subcellular fraction sample and the whole sample, in *T. pseudonana*. Two cases are presented, the value for DAG(Cell) corresponding to the average value for the major DAG lipids in *T. pseudonana* (DAG(16:1/14:0), DAG(32:4), DAG(32:3), DAG(16:2/16:0), DAG(16:1/16:0) and DAG(18:4/16:0). DAG(PC) corresponds to the average for the PC related (DAG(40:10) and DAG(42:11) which have been shown to be absent/trace in previous total lipid extracts of *T. pseudonana* (Chapter 2). Data represent the average of the ratio for the different lipid species within each category and 1 S.D., determined from analysis of a single sample in each case.

As a semi-quantitative measure for the localisation of different lipid types in the different subcellular locales of *T. pseudonana*, the ratio of intensity for a given lipid peak in the subcellular chloroplast fraction, divided by the whole cell was calculated (Figure 4). We found that the ratio value for the predominant DAG chemotypes (DAG(Cell)), listed in the figure caption, was 30.58 ± 3.02. In contrast, the DAG chemotypes related to the polyunsaturated PC species (DAG(PC): DAG(40:10) and DAG(42:11)) yielded a ratio value of 16.51 ± 4.17. This represents a statistically significant variation with $p = 0.0018$, calculated based upon the ratios for each of the individual lipid species. The DAG(PC) species were therefore, depleted by approximately 0.5 fold in the subcellular fraction.

3.5 Discussion

3.5.1 Untargeted Lipidomic Screening of Phosphrous Stressed *T. pseudonana*

Total lipid extracts from cultured *T. pseudonana* after 72 h of P stressed and P replete growth were analysed by untargeted lipidomic methodology. Ions were ranked according to differential abundance associated with the P+ and P- treatments, and the top 15 were characterised based upon their MS2 spectra.

In positive and negative ion mode, the results were dominated by the presence absence increase of DGCC molecular species, between the P- and P+ cultures respectively. Under P stress, the abundant glycerophospholipid PC is substituted for P free DGCC, the dynamics of which are well established in *T. pseudonana* ((**Van Mooy et al.**, 2009; **Martin et al.**, 2011) and Chapter 2). DGCC species containing 20:5, 22:6 and 16:0 fatty acids increased most, in line with previous observations (Chapter 2).

TAG(16:0/16:0/14:0) also demonstrated a presence absence increase between P- and P+ cultures respectively. TAGs act as storage lipids and have also been studied extensively in *T. pseudonana* and other phytoplankton in order maximise lipid yields for the production of biofuels (**Yu et al.**, 2009; **Hildebrand et al.**, 2012). In the case of *T. pseudonana*, the lipidomic implications of nitrate and silicate nutrient starvation upon the TAGs have been reported (**Yu et al.**, 2009), however, to our knowledge its TAG P stress response has not been characterised.

TAG(16:0/16:0/14:0) in *T. pseudonana* has been reported elsewhere to be a trace species comprising up to 1% of the TAG lipids with no differential response to nitrate or silicate starvation (**Yu et al.**, 2009). Our data indicate a drastic increase in this particular TAG species, an observation that suggests a somewhat different TAG accumulation response under phosphorus stress, than nitrate or silicate stress. This also indicates that further attention to the TAG response of *T. pseudonana* under P stress may be warranted.

In negative ion mode, the sulfolipid species SQDG(34:1) increased approximately 3.5 fold between the P- and P+ cultures. SQDG commonly occurs in the thylakoid membranes

of the chloroplast (**Siegenthaler and Murata**, 2004) and has been reported as a non-phosphorus lipid substitute for PG in *T. pseudonana* (**Van Mooy et al.**, 2006, 2009; **Martin et al.**, 2011). As such an increase in this species is not unexpected under P stress, however, SQDG(34:1) comprises $< 2\%$ of the total SQDG under P+ and P- conditions (Chapter 2).

Of particular interest in negative ion mode was the identification of $(Gly)_2Cer(d18:4/24:0)$ which displayed a presence absence increase between the P- and the P+ cultures respectively. The structural assignment was based on several diagnostic fragments in the MS2 spectra under both positive and negative ionisation, including those corresponding to the long chain base, fatty acid amide and diglycosyl headgroup (Figure 3.2).

This discovery led to the identification of a series of related molecules varying subtly in the degrees of unsaturation of the long chain base or the fatty acid amide (Figure 3.3). All of the diglycosylceramide species increased under P stress, varying in magnitude of change between approximately 2 and 10 fold. A Cer(d18:3/24:0) species was also identified and demonstrated contrasting behaviour, this putative precursor did not vary in cellular abundance subject to P stress. These species represent, to the best of our knowledge, the first sphingolipids characterised in *T. pseudonana*. Chemically identical species and additional related lipids have however, been recently reported in another marine diatom, *Skeletonema costatum* (**Zhao et al.**, 2013). This lends further credence to our identification and highlights these molecules as an area of interest for future research into diatom lipid biochemistry.

Glycosylceramides have been ascribed to several physiological functions including membrane stability, membrane permeability and pathogenesis (**Pata et al.**, 2010). In this context, it is also possible that they are acting as non-phosphorus substitute lipids, akin to DGCC (**Van Mooy et al.**, 2006, 2009; **Martin et al.**, 2011), for another phosphorus containing lipid and perhaps this drives their differential increase under P stress.

Closely related glycosphingolipids have been identified in virally infected *Emiliania huxleyi* and applied as biomarkers for viral infection in the marine environment (**Vardi**

et al., 2009, 2012). In the same manner, we propose the diglycosylceramides that increase in cellular abundance subject to P stress, as candidate biomarkers for the P stress of marine diatoms in the environment. As such, the abundance of these potentially diagnostic biomarkers could be quantified in diatoms isolated from the environment of interest. This data could yield insight into the level of P stress experienced by the phytoplankton directly, which would be complimentary information to the more routine measurement of dissolved or particulate P concentrations in the medium. Targeted quantification of these lipids, in a range of diatom species subjected to P stress, would be achievable by triple quadrupole mass spectrometry and the application of lactosyl-ceramide standards as in (**Zhao et al.**, 2013). Such a lipid biomarker could be useful to follow P related biogeochemical processes in the environment, pending substantial further research into validation.

In order to validate the diglycosylceramides as a proxy for P stress, the relationship must be demonstrated to be consistent in other diatoms and ideally other phytoplankton taxa. Finally, the response must be specific to P stress and not elicited under other conditions. If these assumptions are proven correct, it may then be possible to generate and index based upon these measures and calibrate it directly to P concentrations in the water column, akin to lipid sea surface temperature proxies (**Brassell et al.**, 1986; **Schouten et al.**, 2002).

3.5.2 Fatty Acyl Composition of the Major Individual Glycerolipid Species

Semi-targeted MS2 analyses were used to characterise the fatty acyl composition of each of the major glycerolipid individual lipid species reported in previous work on *T. pseudonana* (Chapter 2). The glycerolipids analysed were comprised of fatty acids between 14:0 and 22:6, in line with previous observations on the total fatty acid pool of *T. pseudonana* (**Volkman et al.**, 1989). We found no corroborating evidence for ions previously reported as PE(29:5), PE(52:12) and PE(54:13) observed to increase in abundance in a presence absence manner subject to P stress (Chapter 2). The true chemical identities of these ions therefore remains unknown.

This characterisation of the fatty acids within each glycerolipid chemotype, compli-

ments previous reports on the glycerolipid dynamics of *T. pseudonana* ((**Yu et al.**, 2009; **Martin et al.**, 2011; **Bromke et al.**, 2013) and Chapter 2) furthering understanding of the major constituents of the lipidome in this important model marine diatom species.

3.5.3 Subcellular Partitioning of Polyunsaturated DAG Lipid Species

Previous work on the response of the *T. pseudonana* glycerolipidome to P stress highlighted a high degree of similarity in the relative compositions of DAG individual lipid species with that of MGDG and DGDG. Furthermore, no correlation with the glycerophospholipid PC was observed (Chapter 2), despite the known role of DAG as an immediate precursor to PC in eukaryotic phospholipid biosynthesis (**Lykidis**, 2007).

In higher plants MGDG and DGDG are known to be abundant in the photosynthetic thylakoid membranes of the chloroplast (**Siegenthaler and Murata**, 2004). DAG is a substrate for two distinct lipid biosynthetic pathways in eukaryotic phytoplankton (**Riekhof et al.**, 2005). One of which is localised to the chloroplast and acts as a precursor to the synthesis of MGDG and DGDG while the other is located within the endoplasmic reticulum and leads to the synthesis of PC and PE (**Riekhof et al.**, 2005).

The observed coupling of DAG with MGDG/DGDG and decoupling with PC/PE in *T. pseudonana* could be explained by two separate pools of DAG: the larger and/or slower turned over DAG pool indicative of MGDG/DGDG synthesis that we observe by characterisation of the total lipid extract and the smaller and/or more rapidly turned over DAG pool indicative of PC/PE synthesis that is conspicuously not observed.

This hypothesis was assessed qualitatively by calculating the ratio of the peak intensity in the subcellular chloroplast fraction to peak intensity in the whole cell fraction for two cases, the predominant DAG chemotypes (DAG(Cell), as reported in Chapter 2) and DAG(40:10)/DAG(42:11) (DAG(PC), which were indicative of PC/PE. We found that DAG(PC) were depleted by approximately 0.5 fold in the prepared subcellular fraction.

This preliminary experiment did not include a quantitative estimate of the subcellular fraction enrichment. It is tentatively concluded that qualitatively there appears to be some partitioning of the DAG pools between the MGDG/DGDG precursor DAG pool within

the chloroplast and PC/PE precursor DAG pool within the rest of the cell. These results represent a proof of concept and form support for further research with full quantification going forward.

3.6 Conclusions

The untargeted lipidomic screening of the model marine diatom *T. pseudonana*, under nutrient replete and phosphorus stressed growth conditions highlighted a number of lipid species, such as several DGCC chemotypes, that increased strongly under P stress in line with previous observations.

A triacylglycerol species, reported elsewhere to be present only at trace levels under P replete, N starved and Si starved conditions was present at increased levels subject to P stress and suggests potentially interesting triacylglycerol dynamics.

Furthermore, a group of diglycosylceramides, not previously detected in *T. pseudonana*, increased in cellular abundance under P stress. These species may prove useful as biomarkers for P stress in diatoms. Such biomarkers give insight into the P stress experienced by the organisms sampled that is complimentary to the routine measurement of dissolved and particulate P concentrations. These biomarkers may have potential applications in the elucidation of P related biogeochemical processes, pending further validation.

Targeted MS2 analyses were used to characterise the fatty acyl compositions of each of the major diglyceride lipid species. This provides new insight into the *T. pseudonana* lipidome filling a conspicuous gap in an otherwise well characterised system.

Finally, a preliminary qualitative study of the subcellular partitioning of precursor DAG lipids supported the hypothesis that DAG lipids were comprised of two subcellular pools: one a putative MGDG/DGDG precursor pool localised to the chloroplast, the other a putative PC/PE precursor localised elsewhere.

3.7 Experimental Procedures

The precision of measurements was accounted for by univariate statistics. Values from biological triplicate measurements (n = 3) were presented as averages with 1 standard deviation error bars where appropriate. Fold changes were tested for statistical significance by unpaired, equal variance T-test with a critical value of < 0.05. Statistical treatments are described in the accompanying figure legends.

3.7.1 UPLC-ESI-AutoMS2 Analysis

All reagents were HPLC/Analytical grade as supplied by Fisher Scientific (Leicestershire, United Kingdom). The UPLC-ESI-AutoMS2 analysis was performed on a Dionex UltiMate 3000 UPLC system coupled to a Bruker maXis 3G quadrupole - time of flight (Q-ToF) mass spectrometer with an electrospray ionisation source.

Lipid samples were dissolved in methanol (200 µL) prior to analysis. A 20 µL injection was taken by autosampler from vials in a cooled sample tray at 5°C. The sample was then chromatographically separated over 30 minutes with a Waters Acquity UPLC BEH C8, 1.7 µm particle, 2.1 x 100mm column. A constant flow rate of 0.3 mL min^{-1} was used resulting in back pressures of between 260 and 460 bar.

Eluent A was water with 0.2% formic acid and 1% 1M ammonium acetate, eluent B was methanol with 0.2% formic acid and 1% 1M ammonium acetate. The column was heated to 50°C and the eluent cooled to 21°C post-column, throughout the analysis. The following multi-step linear gradient was applied, with a constant flow rate of 0.3 mL $minute^{-1}$: 35% eluent B at 0 minutes, increasing to 80% eluent B at 2 minutes, increasing to 95% eluent B at 12 minutes and holding for a further 18 minutes until the end of the run. Eluent B was decreased to 35% over 0.5 minutes post run and the column allowed to equilibrate for 4.5 minutes prior to the next run.

The mass spectrometer was calibrated by direct infusion of sodium formate solution prior to use (10 mM sodium hydroxide + 0.2% formic acid in 1:1 isopropanol/water). The observed mass accuracy was 0.4 ppm, determined from the standard deviation from the

quadratic calibration curve for calibrant ions up to 1000 Da, in positive ion mode. The Q-ToF mass spectrometer yielded a mass resolving power of 21463.70, determined from full width at half maximum (FWHM) of the internal standard peak dilauroylphosphatidylcholine at 622.4470 Da $[M+H]^+$ in positive ion mode.

Full scan MS was acquired in positive and negative mode between 30 and 1500 Da. During each run (positive and negative ion mode require separate analytical runs) ions above the noise level threshold were subjected to data dependant MS2 fragmentation. The threshold was set at 2000 and 1000 counts in positive and negative ion mode respectively. Ions with a mass of between 500 - 1500 and 300 - 1500 Da were subject to MS2 fragmentation in positive and negative ion mode respectively. The most abundant two precursor ions eluting during an MS scan were fragmented and after two MS2 spectra were acquired for a given ion, they were actively excluded from further MS2 for 1 minute. Fragmentation for MS2 was achieved by collision induced dissociation (CID) by impact with Argon gas, with stepped collision energies for precursor ions of increasing mass. Ions of 300-500, 500-800 and 800+ Da were fragmented with collision energies of 25, 40 and 50 eV respectively in positive ion mode and 25, 30 and 40 eV in negative ion mode.

Where minor ions of interest were not selected automatically for MS2 analysis, these were subjected to targeted MS2 fragmentation analysis under the same conditions.

3.7.2 Data Processing

Bruker CompassXport was used to export the raw data prior to processing with the MZMine 2 software package (**Pluskal et al.**, 2010). MZMine was used to generate extracted ion chromatograms and match these chromatograms between different samples. Integrated peak areas were normalised to the total number of cells extracted and adjusted for recovery of the internal standard (dilauroylglycerophosphatidylcholine). Peak assignments were based upon matching to an extensive, accurate mass, structure query language (SQL) lipid database generated in house. The database was populated by permutations of fatty acids (chain length/degree of unsaturation) and common glycerolipids/sphingolipids. The complete LIPID MAPS (version 20130306) structural database (**Sud et al.**, 2007)

and MaConDa mass spectrometry contaminants database (**Weber et al.**, 2012) were also included.

The chemical formulae of database entries were then used to calculate accurate mass m/z values based upon a list of common molecular ion adducts in ESI-MS (**Huang et al.**, 1999). Tentative assignments were made by matching precursor mass ions with the theoretical database to within a mass difference of $<\pm 10$ ppm.

Database assignments were then confirmed by the identification of supporting MS2 fragments in each case. Fragments were assigned within a tolerance of $<\pm 20$ ppm unless otherwise stated.

The MS2 data was processed with Bruker DataAnalysis using the Find AutoMSn function.

The extracted MS2 spectra were then assigned using an in house visual basic macro in Microsoft Excel by matching fragment ions to a database of common and diagnostic fragments and dynamically generated neutral losses based upon the parent ion mass. Matches were made based upon a mass difference of <20 ppm (unless otherwise specified).

3.7.3 Phosphorous Stressed *T. pseudonana* Cultures

The cultured *T. pseudonana* total lipid extracts used to perform the untargeted lipidomic screening experiment were the 72 h samples prepared as described in Chapter 2. Samples were dried down under flowing N_2 and redissolved in methanol prior to analysis.

3.7.4 Subcellular Fractionation of *T. pseudonana*

T. pseudonana was cultured under P replete conditions in 4 L of F/2 + Si media (**Guillard**, 1975) as detailed in Chapter 2. The cells were isolated from the media by filtration during mid/late exponential growth phase, at a cell concentration of 1.90 x 10^6 cells mL^{-1}. A plastid subcellular fraction (chloroplasts and mitochondria) was prepared by a modified method based upon (**Wittpoth et al.**, 1998), as follows:

2 x 10^9 cells were resuspended in 20 mL homogenisation buffer (323 mM sorbitol, 20 mM MOPS, 5 mM -aminocaproic acid and 1 mM benzylamide). The cells were homogenised by

3 cycles of cell disruption with a Constant Systems single shot cell disruptor at 35,000 psi. After rinsing the apparatus, the approximately 100 mL of combined homogenate and buffer rinse was centrifuged at 10,000 x G for 10 minutes and the pellet retained. The pelleted homogenate (including the plastid fraction) was resuspended in 1 mL homogenisation buffer and layered gently on top of a linear percoll gradient. The percoll gradient consisted of 40-60% percoll solution in homogenisation buffer, with 5% steps of 2 mL each. The loaded gradient was centrifuged at 3000 x G for 4 minutes with minimum acceleration/deceleration using a swing bucket rotor. 1 mL fractions were isolated sequentially from the top of the percoll gradient and centrifuged at 24,000 x G for 10 minutes. The pellet was retained and resuspended in phosphate buffered saline prior to lipid extraction by the Bligh-Dyer method (**Bligh and Dyer**, 1959) modified as detailed in Chapter 2.

3.8 Acknowledgements

This work was funded by the University of Southampton - Vice Chancellors Scholarship Award. The purchase of the mass spectrometer was supported by the Wellcome Trust (Grant 057405).

3.9 References

Armbrust, E. V., Berges, J. A., Bowler, C., Green, B. R., Martinez, D., Putnam, N. H., et al. (2004), The genome of the diatom Thalassiosira pseudonana: ecology, evolution, and metabolism., *Science*

Benning, C., Beatty, J. T., Prince, R. C., and Somerville, C. R. (1993), The sulfolipid sulfoquinovosyldiacylglycerol is not required for photosynthetic electron transport in Rhodobacter sphaeroides but enhances growth under phosphate limitation., *Proceedings of the National Academy of Sciences of the United States of America*

Benning, C., Huang, Z. H., and Gage, D. A. (1995), Accumulation of a novel glycolipid and a betaine lipid in cells of Rhodobacter sphaeroides grown under phosphate limitation., *Archives of Biochemistry and Biophysics*

Blanksby, S. J. and Mitchell, T. W. (2010), Advances in Mass Spectrometry for Lipidomics, *Annual Review of Analytical Chemistry*

Bligh, E. G. and Dyer, W. J. (1959), A Rapid Method of Total Lipid Extraction and Purification, *Canadian Journal of Biochemistry and Physiology*

Brassell, S. C., Eglinton, G., Marlowe, I. T., Pflaumann, U., and Sarnthein, M. (1986), Molecular stratigraphy: a new tool for climatic assessment, *Nature*

Bromke, M. A., Giavalisco, P., Willmitzer, L., and Hesse, H. (2013), Metabolic Analysis of Adaptation to Short-Term Changes in Culture Conditions of the Marine Diatom Thalassiosira pseudonana, *PLoS ONE*

Fahy, E., Subramaniam, S., Brown, H. A., Glass, C. K., Merrill, A. H., Murphy, R. C., et al. (2005), A comprehensive classification system for lipids, *Journal of Lipid Research*

Falkowski, P. G., Barber, R. T., and Smetacek, V. (1998), Biogeochemical Controls and Feedbacks on Ocean Primary Production, *Science*

Guillard, R. R. L. (1975), Culture of phytoplankton for feeding marine invertebrates., in W. L. Smith and M. H. Chanley, eds., Culture of Marine Invertebrate Animals (Plenum Press, New York, USA.),

Guschina, I. A. and Harwood, J. L. (2006), Lipids and lipid metabolism in eukaryotic algae, *Progress in Lipid Research*

Han, X. (2009), Lipidomics: Developments and applications, *Journal of Chromatography. B, Analytical Technologies in the Biomedical and Life Sciences*

Hildebrand, M., Davis, A. K., Smith, S. R., Traller, J. C., and Abbriano, R. (2012), The place of diatoms in the biofuels industry, *Biofuels*

Huang, N., Siegel, M. M., Kruppa, G. H., and Laukien, F. H. (1999), Automation of a Fourier transform ion cyclotron resonance mass spectrometer for acquisition, analysis, and e-mailing of high-resolution exact-mass electrospray ionization mass spectral data, *Journal of the American Society for Mass Spectrometry*

Ivanova, P. T., Milne, S. B., Myers, D. S., and Brown, H. A. (2009), Lipidomics: a mass spectrometry based systems level analysis of cellular lipids, *Current Opinion in Chemical Biology*

Lykidis, A. (2007), Comparative genomics and evolution of eukaryotic phospholipid biosynthesis, *Progress in Lipid Research*

Martin, P., Van Mooy, B. A. S., Heithoff, A., and Dyhrman, S. T. (2011), Phosphorus supply drives rapid turnover of membrane phospholipids in the diatom Thalassiosira pseudonana, *The ISME Journal*

Mouritsen, O. G. (2011), Lipidology and lipidomics–quo vadis? A new era for the physical chemistry of lipids, *Physical Chemistry Chemical Physics*

Pata, M. O., Hannun, Y. A., and Ng, C. K.-Y. (2010), Plant sphingolipids: decoding the enigma of the Sphinx, *New Phytologist*

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

Popendorf, Tanaka, T., Pujo-Pay, M., Lagaria, A., Courties, C., Conan, P., et al. (2011a), Gradients in intact polar diacylglycerolipids across the Mediterranean Sea are related to phosphate availability, *Biogeosciences*

Popendorf, K. J., Lomas, M. W., and Van Mooy, B. A. S. (2011b), Microbial sources of intact polar diacylglycerolipids in the Western North Atlantic Ocean, *Organic Geochemistry*

Riekhof, W. R., Sears, B. B., and Benning, C. (2005), Annotation of Genes Involved in Glycerolipid Biosynthesis in Chlamydomonas reinhardtii: Discovery of the Betaine Lipid Synthase BTA1Cr, *Eukaryotic Cell*

Schouten, S., Hopmans, E. C., Schefuß, E., and Sinninghe Damsté, J. S. (2002), Distributional variations in marine crenarchaeotal membrane lipids: a new tool for reconstructing ancient sea water temperatures?, *Earth Planet Sc. Lett.*

Schubotz, F., Wakeham, S. G., Lipp, J. S., Fredricks, H. F., and Hinrichs, K.-U. (2009), Detection of microbial biomass by intact polar membrane lipid analysis in the water column and surface sediments of the Black Sea., *Environmental Microbiology*

Siegenthaler, P.-A. and Murata, N. (2004), Lipids in Photosynthesis: Structure, Function and Genetics (Springer Netherlands)

Sud, M., Fahy, E., Cotter, D., Brown, A., Dennis, E. A., Glass, C. K., et al. (2007), LMSD: LIPID MAPS structure database, *Nucleic Acids Research*

van Meer, G. (2005), Cellular lipidomics, *EMBO J*

Van Mooy, B. A. S., Fredricks, H. F., Pedler, B. E., Dyhrman, S. T., Karl, D. M., Koblížek, M., et al. (2009), Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity, *Nature*

Van Mooy, B. A. S., Rocap, G., Fredricks, H. F., Evans, C. T., and Devol, A. H. (2006), Sulfolipids dramatically decrease phosphorus demand by picocyanobacteria in oligotrophic marine environments, *Proceedings of the National Academy of Sciences*

Vardi, A., Haramaty, L., Van Mooy, B. A. S., Fredricks, H. F., Kimmance, S. A., Larsen, A., et al. (2012), Hostvirus dynamics and subcellular controls of cell fate in a natural coccolithophore population, *Proceedings of the National Academy of Sciences*

Vardi, A., Van Mooy, B. A. S., Fredricks, H. F., Popendorf, K. J., Ossolinski, J. E., Haramaty, L., et al. (2009), Viral glycosphingolipids induce lytic infection and cell death in marine phytoplankton., *Science*

Volkman, J. K., Jeffrey, S. W., Nichols, P. D., Rogers, G. I., and Garland, C. D. (1989), Fatty acid and lipid composition of 10 species of microalgae used in mariculture, *Journal of Experimental Marine Biology and Ecology*

Wakeham, S. G., Hedges, J. I., Lee, C., Peterson, M. L., and Hernes, P. J. (1997), Compositions and transport of lipid biomarkers through the water column and surficial sediments of the equatorial Pacific Ocean, *Deep Sea Research Part II: Topical Studies in Oceanography*

Weber, R. J. M., Li, E., Bruty, J., He, S., and Viant, M. R. (2012), MaConDa: a publicly accessible mass spectrometry contaminants database, *Bioinformatics*

Wenk, M. R. (2005), The emerging field of lipidomics, *Nat. Rev. Drug. Discov.*

Wittpoth, C., Kroth, P. G., Weyrauch, K., Kowallik, K. V., and Strotmann, H. (1998), Functional characterization of isolated plastids from two marine diatoms, *Planta*

Yu, E. T., Zendejas, F. J., Lane, P. D., Gaucher, S., Simmons, B. A., and Lane, T. W. (2009), Triacylglycerol accumulation and profiling in the model diatoms Thalassiosira pseudonana and Phaeodactylum tricornutum (Baccilariophyceae) during starvation, *Journal of Applied Phycology*

Zhao, F., Xu, J., Chen, J., Yan, X., Zhou, C., Li, S., et al. (2013), Structural elucidation of two types of novel glycosphingolipids in three strains of Skeletonema by liquid chromatography coupled with mass spectrometry, *Rapid Communications in Mass Spectrometry*

Chapter 4

Lipid Remodelling by the Diatom *Thalassiosira pseudonana* under Nitrogen Stress – Diglyceride Lipid Dynamics and Relation to Triglyeride Production

Jonathan E. Hunter[1,2], Samantha Earl[3], Joost Brandsma[4], Marcus Dymond[5], Grielof Koster[4], C. Mark Moore[1], Anthony D. Postle[4], George S. Attard[3] and Rachel A. Mills[1].

1. Ocean and Earth Science, University of Southampton, National Oceanography Centre Southampton, European Way, Southampton, SO14 3ZH, United Kingdom
2. Institute for Life Sciences, University of Southampton, SO17 1BJ, United Kingdom
3. Chemistry, University of Southampton, Southampton, SO17 1BJ, United Kingdom
4. Faculty of Medicine, University of Southampton, Southampton General Hospital, Tremona Road, Southampton, SO16 6YD, United Kingdom
5. School of Pharmacy and Biomolecular Sciences, University of Brighton, Brighton, BN2 4GJ, United Kingdom

4.1 Author Contributions

Jonathan E. Hunter designed the experiments, developed methods, carried out culturing, sample preparation, analysis, data interpretation and wrote the manuscript. Samantha Earl assisted in the development of methods and carried out preliminary research. The remaining co-authors assisted with experimental design, interpretation and drafting the manuscript.

4.2 Abstract

Primary production in oligotrophic, low-latitude oceans is limited by nitrate; affecting phytoplankton such as diatoms. Marine diatoms are used as feedstocks for biofuel production, because of their richness in lipids, in particular the major storage lipids triacylglycerides. Nitrogen starvation is often used to increase lipid yield in biofuel production. Lipidomic analyses have previously been conducted to characterise the impact of nitrate limitation on the diglyceride lipidome of the model marine diatom *Thalassiosira Pseudonana*. However, absolute quantification (*i.e.* moles cell^{-1}) and the relative abundance of different chemotypes within a given lipid headgroup type has not to date been reported and therefore little is known about the mechanisms of lipid dynamics within the cell.

We present a detailed quantitative study of the diglyceride lipids of *T. pseudonana*, subject to N limitation, by targeted mass spectrometric lipidomics. In addition, the fatty acid composition of each individual lipid species was explicitly characterised by fragmentation analysis. We find that under N limitation, total glycerophosphatidylcholine (PC) increases, primarily attributable to highly unsaturated PC species bearing 20:5 and 22:6 fatty acids. An increase in total sulfoquinovosyldiacylglycerol (SQDG) subject to N limitation was almost exclusively attributed to a large increase in SQDG(14:0/14:0), the predominant chemotype. Finally, total diacylglycerol (DAG) increased 3 fold under N limitation, comprised of increases in DAG species bearing 16:0, 16:1 and 14:0 fatty acids. This increase in DAG appears to form part of the cells adaption to N limitation that ultimately leads to the accumulation of triacylglycerides. These findings further our understanding of the diatom lipidome and its response to nutrient limitation, with signfiicant implications for biofuel production and adaption of marine diatoms to low N environments.

Keywords: *Thalassiosira pseudonana*, Lipid, Glycerophospholipid, Diatom, Lipidomics, Nitrogen, Stress, Limitation, ESI-MS.

4.3 Introduction

Nitrate (NO_3^-), the most abundant bioavailable form of nitrogen in the oceans, is an important macronutrient for the growth of marine primary producers (**Tyrrell**, 1999). Phytoplankton production in vast regions of oligotrophic low-latitude ocean is proximately limited by depleted nitrate content in surface ocean waters (**Moore et al.**, 2013). Diatoms are a populous group of eukaryotic microalgae that contribute up to 25% of global annual primary production (**Hoek et al.**, 1995) and nitrogen scarcity is therefore an important factor in diatom ecology and biogeochemistry. Understanding the physiological responses and adaptions of marine diatoms to nitrogen scarcity is important in determining the ecological response under nitrate limitation and impact on global biogeochemical cycles including carbon dioxide drawdown.

Marine diatoms have also received attention as potential feedstocks for biofuel production. This interest results from the ecological success of diatoms and the high proportion of their dry weight contributed by lipids, in particular the major storage lipids triacylglycerides (TAGs) (**Levitan et al.**, 2014). These TAGs can be isolated and converted into biofuel via transesterification reactions (**Meher et al.**, 2006). Substantial economic barriers remain however and these are largely a matter of lipid production efficiency (**Hildebrand et al.**, 2012; **Levitan et al.**, 2014). Efforts to maximise lipid yields have often utilised the modification of environmental conditions, such as nitrogen starvation. A survey of the literature found that diatom total lipid content was on average 27.4% of dry cell weight under nutrient replete conditions increasing to 36.1% under N-deficient conditions (**Griffiths and Harrison**, 2009; **Hildebrand et al.**, 2012).

Genetic modification of diatom cells is also possible and given that conventional breeding approaches are not practical due to the characteristics of the diatom sexual cycle, it can be used to further increase diatom lipid production efficiency (**Levitan et al.**, 2014). In order to do so, a characterised genome and knowledge of an organisms lipid metabolism/lipidome is required.

In the case of the centric marine diatom *Thalassisosira Pseudonana* (*T. pseudonana*),

the genome has been sequenced (**Armbrust et al.**, 2004). Furthermore, the lipidome has been characterised at various levels, such as the total fatty acids (**Volkman et al.**, 1989) and TAG specific fatty acids (**Tonon et al.**, 2002). Much attention has been given to the dramatic glycerophospholipid substitution response to phosphorus starvation and the intact glycerolipid dynamics ((**Van Mooy et al.**, 2009; **Martin et al.**, 2011) and Chapter 2). Of particular relevance is the characterisation of the TAG lipid pool of *T. pseudonana* analysed during a nitrogen limited stationary growth phase (**Yu et al.**, 2009). Under nitrogen limitation, *T. pseudonana* reached a maximum TAG content of 14.5% of dry cell weight, primarily comprised of palmitic (16:0), palmitoleic (16:1) and myristic (14:0) fatty acid substituents. Higher molecular weight TAGs, bearing eicosapentaenoic (20:5) and docosahexaenoic (22:6) acids were also observed. With few minor exceptions, the distribution of TAG molecular species was consistent between the nitrogen replete and nitrogen limited samples (**Tonon et al.**, 2002; **Yu et al.**, 2009).

A recent report presents results from a joint metabolic and lipidomic profiling approach of *T. pseudonana* under N starvation (**Bromke et al.**, 2013). Firstly, broad reductions in cellular abundance of nitrogenated primary metabolites were observed. Secondly, lipidomic profiles were reported, between N limited and N replete controls, as determined by an untargeted lipidomic approach (**Giavalisco et al.**, 2011; **Bromke et al.**, 2013). Differential increases were observed across the TAG pool, in line with other literature as previously discussed. Furthermore, diacylglycerol (DAG) species bearing a total of 3-5 fatty acid double bonds were elevated under N limitation. Many phosphoglycerolipids and glycoglycerolipids exhibited differential abundances, however no clear trends were evident (**Bromke et al.**, 2013). These results provided quantification by differential abundance on a per individual lipid species basis. However, absolute quantification (*i.e.* moles cell^{-1}) and between lipid species comparison (relative abundance) was not presented. Furthermore, lipids containing polyunsaturated fatty acids such as 20:5 and 22:6 were not reported, yet previous research shows the importance of these species, in the glycerolipidome ((**Tonon et al.**, 2005; **Yu et al.**, 2009) and Chapter 2). Finally, the fatty acid composition of each individual glycerolipid involved in N limitation has yet to be explicitly characterised (with the exception

of the TAGs (**Yu et al.**, 2009)).

We present a detailed quantitative study of the diglyceride lipids of *T. pseudonana* subject to N limitation, by targeted lipidomics. By utilising direct infusion, electrospray ionisation, triple quadrupole mass spectrometry and appropriate lipid internal standards, absolute quantification (moles cell^{-1}) of the diglyceride lipids was achieved. In addition, the fatty acid composition of each individual lipid species was characterised by fragmentation analysis with ultra-performance liquid chromatography coupled to an accurate mass quadrupole time of flight mass spectrometer. We describe trends in the response of the diglyceride lipidome to nitrogen limitation and relate these observations to the well-developed literature on triglyceride lipids. These results further our knowledge of the diatom lipidome, its response to N limitation and TAG production. This knowledge bears implications for the adaption of marine diatoms to N limited conditions throughout the worlds oceans and can inform further optimisation of TAG production efficiency for biofuel applications by genetic modification or other means.

4.4 Results

4.4.1 Culture Growth Parameters

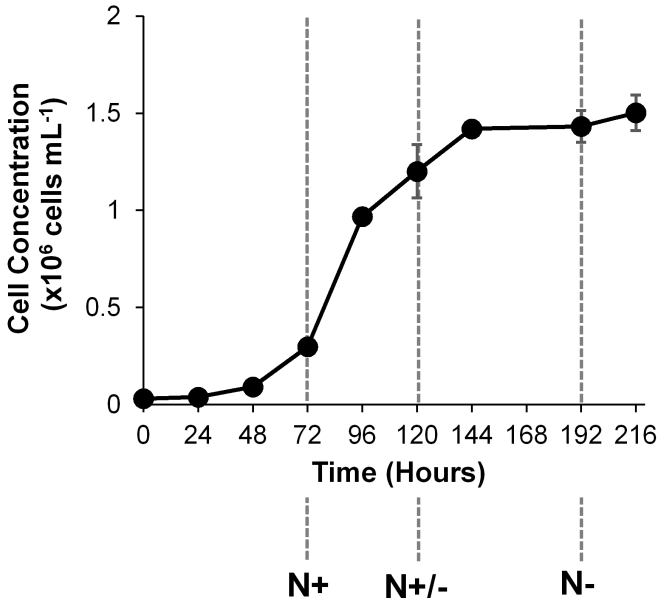


Figure 4.1: Cell concentration growth curve, through time, indicating sampling time for nitrogen replete (N+), nitrogen stressed (N+/-) and nitrogen limited (N-) total lipid extracts. Data are the mean of n = 3 biological replicates, with error bars of 1 standard deviation. Statistical significance is indicated in accompanying text, *p<0.05 **p<0.005 by two-tailed, paired equal-variance T-test.

Growth of the cultures through time is shown in Figure 4.1. The cultures were grown on low N media (50 μM), with phosphate, orthosilicate and micronutrients in excess, such that N would initially be replete but be rapidly depleted. The first sample was taken during exponential growth at 72 h such that N was replete (N+). The second sample was taken at 120 h as the growth rate began to slow with the interface to stationary phase and the initiation of N stress, designated (N+/-). The final sample was taken at 192 h during the stationary phase after growth was limited by N availability for approximately 48 h (N-). The cultures were inoculated at 3.10 x 10^4 $\pm$ 2.54 x 10^3 cells mL^{-1} and grew exponentially until 96 h. Samples were taken at 72 h for lipids derived from exponentially growing, N+

cells. A percentage viability of 97.07 ± 3.00 % at 72 h indicated culture health.

After 96 h growth rate slowed with the interface to the stationary phase and samples were taken at 120 h for N+/- lipids. Cell concentration in the cultures at 120 h was 1.20 x 10^6 ± 1.37 x 10^5 cells mL^{-1} with a percentage viability of 97.13 ± 1.25 %. Culture growth plateaued after 144 h indicating the commencement of a N- stationary phase. N- lipid samples were isolated from the cultures at 192 h, with a cell concentration of 1.43 x 10^6 ± 8.12 x 10^4 cells mL^{-1} and percentage viability of 97.90 ± 0.61 %

4.4.2 Total Lipid Class Response to N Limitation

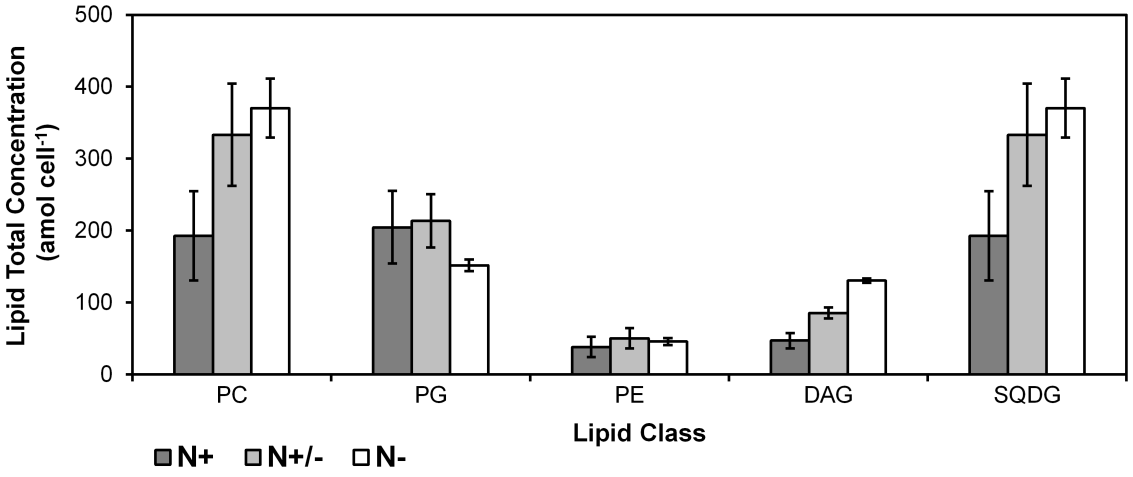


Figure 4.2: Lipid total concentration within each class. Data are the mean of n = 3 biological replicates, with error bars of 1 standard deviation. Lipid class definitions are in the text below. Statistical significance is indicated in accompanying text, *p<0.05 **p<0.005 by two-tailed, paired equal-variance T-test.

The total lipid concentration within the polar lipid classes is shown in Figure 4.2. Total glycerophosphatidylcholine (PC) per cell showed a 1.92 ± 0.65* fold increase under N- conditions, relative to N+. Glycerophosphatidylglycerol (PG) and glycerophosphatidylethanolamine (PE) per cell were statistically similar between N- and N+ samples. Total diacylglycerol (DAG) showed the most dramatic increase of 2.77 ± 0.62** fold between N+ and N- samples. Sulfoquinovosyldiacylglycerol (SQDG) increased 1.92 ± 0.65*

fold in the N- samples.

4.4.3 Individual Lipid Species Response to N Limitation

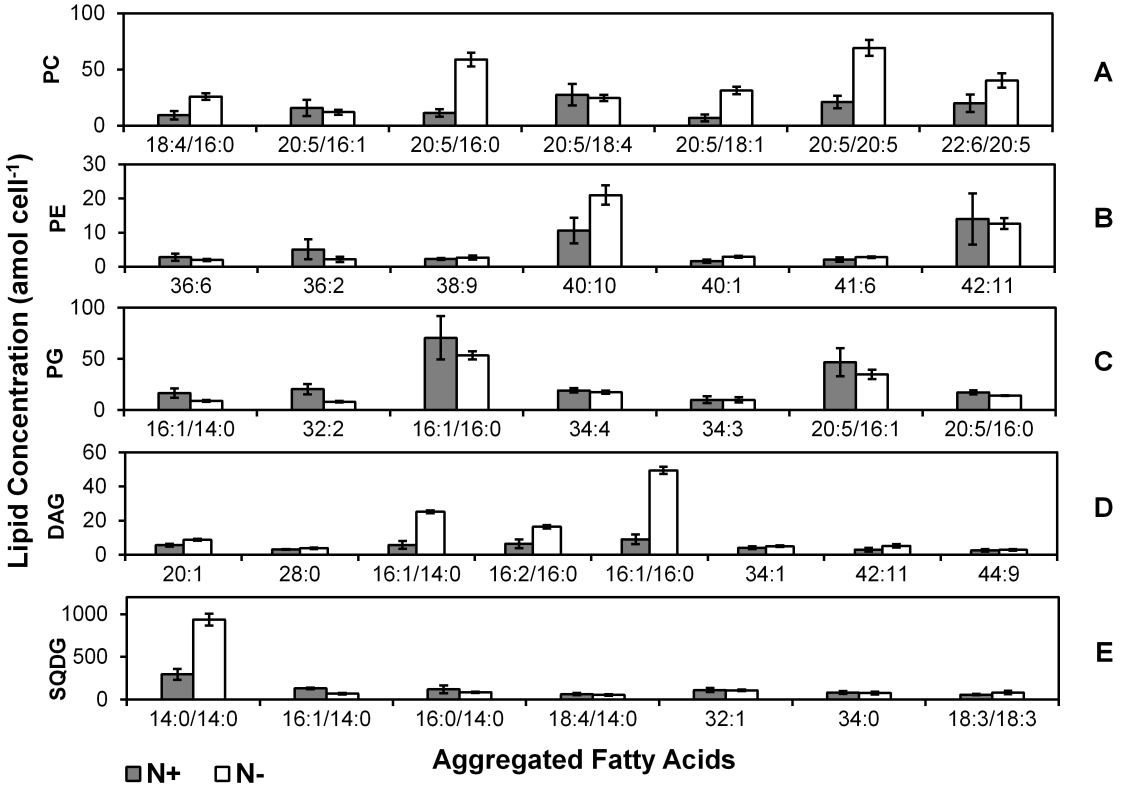


Figure 4.3: Concentration per cell of individual lipid species within each of the classes quantified: PC (A), PE (B), PG (C), DAG (D), SQDG (E). Fatty acid structures are presented as an aggregate of the two fatty acids or the primary fatty acid combination is specified where available (non-regiospecific). Lipid species contributing less than 5% within their class in both the N+ and N- samples are excluded. Data are the mean of n = 3 biological replicates, with error bars of 1 standard deviation. Statistical significance is indicated in accompanying text, *$p<0.05$ **$p<0.005$ by two-tailed, paired equal-variance T-test.

The concentration per cell of individual lipid species with each lipid class are shown in Figure 4.3. Individual lipid species contributing more than 5% within their lipid class and displaying a statistically significant differential abundance between N+ and N- samples are

presented in the following section:

A number of significant increases in quantity were observed within the PC head group class (Figure 4.3A), subject to N limitation. Molecular species PC(20:5/16:0), PC(20:5/18:1), PC(20:5/20:5), PC(18:4/16:0) and PC(22:6/20:5) increased 5.26 ± 1.69**, 4.52 ± 2.05**, 3.30 ± 0.95**, 2.80 ± 1.16** and 2.01 ± 0.84* fold respectively, between the N- and N+ samples.

The major PE lipid species remained statistically similar in abundance between N- and N+ samples, with two exceptions (Figure 4.3B). PE(40:10) and PE(40:1) showed increases of 1.98 ± 0.75* and 1.80 ± 0.54* fold respectively, under N- conditions.

Of the PG lipid species contributing more than 5% of the total PG, only two showed statistically significant variation in abundance under N limitation (Figure 4.3C). PG(32:2) and PG(20:5/16:0) decreased by 0.40 ± 0.11* and 0.81 ± 0.090* fold respectively, between the N- and N+ samples.

Reflecting the increasing trend under N- at the total class level observed in the previous section, a number of DAG lipid species increased in concentration per cell under N limitation (Figure 4.3D). DAG(16:1/16:0), DAG(16:1/14:0), DAG(16:2/16:0) and DAG(20:1) increased 5.44 ± 1.67**, 4.35 ± 1.77**, 2.57 ± 1.06** and 1.57 ± 0.26** fold respectively, between N- and N+.

The class level increase in SQDG observed was attributed exclusively to a 3.18 ± 0.74** fold increase in the predominate SQDG(14:0/14:0) species (Figure 4.3E), between the N- and N+ samples. In contrast, SQDG(16:1/14:0) displayed a per cell decrease of 0.54 ± 0.081** fold, subject to N limitation.

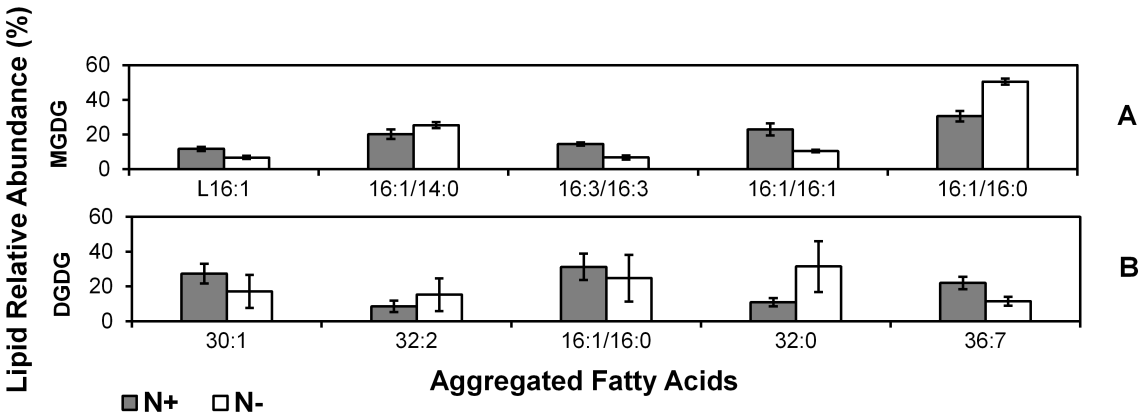


Figure 4.4: Relative abundance of individual lipid species as a percentage of the whole lipid class: MGDG (A), DGDG (B). Fatty acid structures are presented as an aggregate of the two fatty acids or the primary fatty acid combination is specified where available (non-regiospecific). Lipid species contributing less than 5% within their class in both the N+ and N- samples are excluded. Data are the mean of n = 3 biological replicates, with error bars of 1 standard deviation. Statistical significance is indicated in accompanying text, *$p<0.05$ **$p<0.005$ by two-tailed, paired equal-variance T-test.

The glyceroglycolipids monogalactosyldiaclyglycerol (MGDG) and digalactosyldiacylglycerol (DGDG) were quantified in terms of their relative contribution to the whole lipid class (Figure 4.4).

MGDG(16:1/16:0) increased in its relative abundance by 1.65 ± 0.17** fold (Figure 4.4A), between the N- and N+ samples, balanced by decreases in MGDG(16:1/16:1), MGDG(16:3/16:3) and LMGDG(16:1) relative abundance of 0.46 ± 0.075**, 0.47 ± 0.087** and 0.58 ± 0.094** fold respectively.

DGDG(36:7) displayed a 0.52 ± 0.14** fold reduction in relative abundance in the N- samples, relative to N+, the only statistically significant variation observed in the DGDG molecular species (Figure 4.4B).

4.5 Discussion

4.5.1 Culture Growth Parameters

Culture growth parameters (cell concentration and viability through time) indicate healthy, exponentially growing cultures under nutrient replete conditions prior to the initiation of an N-limited stationary phase after 144 h. The cultures were sampled for lipid analysis at 72, 120 and 192 h representing N replete (N+), N stressed (N+/-) and N limited (N-) samples respectively.

4.5.2 Total Lipid Class Response to N Limitation

Total glycerophospholipid PC and sulfolipid SQDG increased in cellular abundance subject to N limitation. Glycerophospholipids PG and PE did not respond to N limitation at the total class level. Total DAG responded most strongly to N limitation with a $\sim$ 3 fold increase in cellular abundance. These observations are discussed in the context of the individual lipid species dynamics in the following section.

4.5.3 Individual Lipid Species Response to N Limitation

The observed increase in total PC was comprised of five lipid species that significantly increased in abundance in the N- samples. Four of these five species contained the highly unsaturated eicosapentaenoic acid (FA(20:5)) and one contained docosahexaenoic acid (FA(22:6)). This agrees with previous observations on increased FA(22:6) production in the *T. pseudonana* polar metabolome under N limitation (**Bromke et al.**, 2013), FA(20:5) however was not quantified in that study. As such, we find that *T. pseudonana* became enriched in highly unsaturated PC lipids when subjected to N limitation.

In contrast, the remaining glycerophospholipid classes PE and PG did not vary as strongly as PC. No significant variation in total PE or PG quantity per cell was evident. Furthermore, no major individual lipid species varied more than a factor of two under N limitation, with the notable exception of a decrease in the comparatively saturated PG(32:2). PG(20:5/16:0) exhibited a minor $\sim$ 0.8 fold decrease. The variation in PG lipid

species is in line with previous observations (**Bromke et al.**, 2013), the PE lipids analysed in that study are distinct from those observed in our own. Overall, PE and PG lipids do not appear to respond substantially to N limitation.

The most dramatic response of the polar lipidome to N limitation concerned the DAGs. An approximately 3 fold increase in total DAG was comprised primarily of substantial increases in DAG species containing palmitic (16:0), palmitoleic (16:1) and myristic (14:0) fatty acids.

It is well known that eukaryotic microalgae, such as *T. pseudonana*, increase production and cellular abundance of storage lipids when subject to nutrient stress (**Yu et al.**, 2009). In culture experiments similar to those presented in this report, triacylglyceride (TAG) storage lipids were observed to increase in cellular abundance under N limitation, whilst retaining an approximately constant composition of individual TAG lipid species (**Yu et al.**, 2009). The authors report that the majority ($\sim$ 60%) of the observed TAG lipid species were comprised of 16:0, 16:1 and 14:0 fatty acids (**Yu et al.**, 2009). DAG serves as a substrate for DAG acyltransferase during the synthesis of TAGs (**Lykidis**, 2007).

It is highly likely therefore that the observed increase in cellular DAG, bearing TAG like fatty acid compositions, forms part of the stress response that drives the production of storage lipids under N limitation in *T. pseudonana*. These observations further our understanding of the TAG biosynthetic pathway, which is critical in the production of biofuels from phytoplankton (**Meher et al.**, 2006; **Levitan et al.**, 2014). We speculate that this knowledge of the metabolic pathway leading to TAG could inform genetic modification or other methods in order to maximise the production of TAGs and hence biofuel yield (**Levitan et al.**, 2014). This outcome is substantial as the economics of lipid biofuel production from phytoplankton are determined largely by lipid production efficiency (**Hildebrand et al.**, 2012; **Levitan et al.**, 2014).

Increases in a number of DAG molecular species subject to N limitation were reported previously by (**Bromke et al.**, 2013). The increases observed however were DAGs with total unsaturation numbers of between three and five. The fatty acids involved were,

therefore, of an unclear relationship with the increasing TAG lipids. This discrepancy is one of a number of broad differences in the observed polar lipidomes between their observations and our own and is likely attributable to differing extraction conditions (**Bromke et al.**, 2013).

We observed an approximately two fold decrease in SQDG(16:1/14:0) subject to N limitation, in line with previous observations (**Bromke et al.**, 2013). The predominant SQDG species SQDG(14:0/14:0) has not been previously characterised in *T. pseudonana* under N limitation (**Bromke et al.**, 2013). We found that SQDG(14:0/14:0) increased $\sim$ 3 fold in cellular abundance, driving the $\sim$ 2 fold increase observed in total SQDG. SQDG is typically associated with the photosynthetic thylakoid membranes (**Benning**, 1998; **Frentzen**, 2004). This shift to an even more SQDG(14:0/14:0) dominated composition shows that N limitation affects the composition of the thylakoid membranes of *T. pseudonana* via SQDG lipids.

The glyceroglycolipids MGDG and DGDG were quantified in relative terms only, resulting in a modest relative increase in MGDG(16:1/16:0) and relative decreases in MGDG(16:1/16:1), MGDG(16:3/16:3) and LMGDG(16:1). Overall, this appears to represent an enrichment of a 16:0 containing species and depletion of 16:1 (and 16:3) containing MGDG species. These observations agree with reports of the behaviour of 16:0 and 16:1 to nitrogen stress at the total fatty acid level (**Tonon et al.**, 2002). An approximately $\sim$ 2 fold reduction in the relative abundance of DGDG(36:7) was also observed, the only significant variation in DGDG composition. Whilst the DGDG pool appears stable to nitrogen limitation in comparison to MGDG, the poor sensitivity of this method to DGDG measurement should be noted.

4.6 Conclusions

In summary, we present results from a targeted lipidomic analysis of *T. pseudonana* under N limited stationary phase growth compared with N replete exponential growth. We find that under N limitation, total PC increases that were primarily attributable to highly unsaturated PC species bearing 20:5 and 22:6 fatty acids, resulted in a relative enrichment of these species. The pools of PE and PG lipids remained largely unchanged by N limitation.

An increase in total SQDG subject to N limitation was almost exclusively attributed to a large increase in SQDG(14:0/14:0), the predominant SQDG chemotype. Within the MGDG pool, we observed a relative enrichment of MGDG species containing 16:0 fatty acids and relative depletion of 16:1 fatty acid containing species.

Finally, total DAG increased 3 fold under N limitation. This increase was comprised of increases in DAG species bearing 16:0, 16:1 and 14:0 fatty acids. Given the biosynthetic relationship and literature reports on the production and chemical composition of TAG under N limitation, this increase in DAG appears to form part of the cells adaption to N limitation that ultimately leads to the accumulation of TAG. These findings further our understanding of the diatom lipidome and its response to nutrient limitation and adaption to low N environments.

Furthermore, this understanding of the TAG metabolic pathway under N limitation could inform genetic modification or other means to increase lipid yield, improving the efficiency and economics of biofuel production.

4.7 Experimental Procedures

4.7.1 Culturing

Axenic culture of *T. pseudonana* (1085/12 also designated CCMP1335/3H) was obtained from the Culture Collection of Algae and Protozoa, Scottish Association for Marine Science, U.K. All culture manipulations were performed under a sterile, laminar flow environment. Borosilicate culture vessels were washed with methanol/chloroform before use. Solvents were HPLC grade (Fisher Scientific).

F/2+Si growth media (**Guillard**, 1975), based on artificial seawater was used (**Kester et al.**, 1967). SiO_4^{4-} concentration in the media was doubled. NO_3^- concentration in seed culture media was reduced slightly to 400 μM to reduce carryover. NO_3^- concentration in the low N media was reduced to 50 μM. All reagents were analytical or biological grade. Seed culture ($\sim$ 100 mL) was grown to mid-log phase concentration of 1.83 x 10^6 cells mL^{-1} in N replete media over 5 days incubation at 18°C; 12:12h light/dark cycle; 123μmol quanta m^{-2} s^{-1} illumination; 70 rpm gentle orbital agitation.

The seed culture (6.57 mL) cells were isolated from the media by filtration (Millipore Steritop, 0.22 μm pore size). Cells were washed on the filter with 20 mL of low N media then resuspended in 30 mL low N media. The 30 mL of resuspended cells were split 3 x 10 mL and made up to 200 mL total in low N media to give biological triplicate cultures. These experimental cultures were incubated as above and monitored daily by coulter counter for culture growth. Cultures were sampled after 72, 120 and 192 hours for size distribution/cell count; viability and lipid extracts.

4.7.2 Cell Size Distribution/Counting

An aliquot (950 μL) of culture was mixed with freshly prepared paraformaldehyde solution (170 μL, 34% w/v, dH_2O) and stored at 4°C for <24 h before analysis. Cell size distributions were generated with a Beckmann Coulter Multisizer 3 Coulter Counter. A 70μm aperture and 3% NaCl electrolyte were used and the samples diluted to ensure <10% aperture coincidence concentration. The Coulter Counter was calibrated with 5.023μm

polystyrene latex standard beads prior to use (Beckmann Coulter via Meritics Ltd., Dunstable, U.K.). Size distributions were used to generate cell concentration values, between the limits of 3 and 9 μm particle diameter.

4.7.3 Viability Assay

Experimental culture (50μL) was incubated with SYTOX-Green dye (Invitrogen Life Technologies, Paisley, U.K.) at a concentration of 0.5μM for 5 minutes in the dark. 18 μL of this solution was then imaged with a Cellometer Vision Duo (Nexcelcom Bioscience via. Peqlab, Sarisbury Green, U.K; X100-F101 Optics; SD100 Slides). All cells were counted manually under brightfield mode and stained, non-viable cells under fluorescence mode (470/535 nm excitation/emission).

4.7.4 Lipid Extraction and Lipidomic Analyses

Lipid samples (20mL) from the experimental cultures were isolated by syringe filtration over pre-combusted (450°C, 12h) GF/F filters. The filtrate was discarded and the filters stored at -78°C until extraction. Total lipid extracts were prepared using a Bligh-Dyer extraction procedure (**Bligh and Dyer**, 1959) as detailed in Chapter 2.

Targeted mass spectrometric analysis was performed on a Waters Micromass Quattro Ultima triple quadrupole instrument and the data processed as outlined in Chapter 2. High resolution MS/MS fragmentation analysis to assign fatty acids was conducted on a Dionex UltiMate 3000 UHPLC coupled to a Bruker maXis II Q-TOF mass spectrometer as described in Chapter 3.

4.8 Acknowledgements

This work was funded by the University of Southampton - Vice Chancellors Scholarship Award. Mass spectrometric infrastructure was supported by the Wellcome Trust (Grant 057405).

4.9 References

Armbrust, E. V., Berges, J. A., Bowler, C., Green, B. R., Martinez, D., Putnam, N. H., et al. (2004), The genome of the diatom Thalassiosira pseudonana: ecology, evolution, and metabolism., *Science*

Benning, C. (1998), Biosynthesis and Function of the Sulfolipid Sulfoquinovosyl Diacylglycerol, *Annual Review of Plant Physiology and Plant Molecular Biology*

Bligh, E. G. and Dyer, W. J. (1959), A Rapid Method of Total Lipid Extraction and Purification, *Canadian Journal of Biochemistry and Physiology*

Bromke, M. A., Giavalisco, P., Willmitzer, L., and Hesse, H. (2013), Metabolic Analysis of Adaptation to Short-Term Changes in Culture Conditions of the Marine Diatom Thalassiosira pseudonana, *PLoS ONE*

Frentzen, M. (2004), Phosphatidylglycerol and sulfoquinovosyldiacylglycerol: anionic membrane lipids and phosphate regulation, *Current Opinion in Plant Biology*

Giavalisco, P., Li, Y., Matthes, A., Eckhardt, A., Hubberten, H.-M., Hesse, H., et al. (2011), Elemental formula annotation of polar and lipophilic metabolites using 13C, 15N and 34S isotope labelling, in combination with high-resolution mass spectrometry, *The Plant Journal*

Griffiths, M. and Harrison, S. (2009), Lipid productivity as a key characteristic for choosing algal species for biodiesel production, *Journal of Applied Phycology*

Guillard, R. R. L. (1975), Culture of phytoplankton for feeding marine invertebrates., in W. L. Smith and M. H. Chanley, eds., Culture of Marine Invertebrate Animals (Plenum Press, New York, USA.),

Hildebrand, M., Davis, A. K., Smith, S. R., Traller, J. C., and Abbriano, R. (2012), The place of diatoms in the biofuels industry, *Biofuels*

Hoek, C., Mann, D. G., and Jahns, H. M. (1995), Algae, (Cambridge University Press)

Kester, D. R., Duedall, I. W., Connors, D. N., and Pytkowicz, R. M. (1967), Preparation of Artificial Seawater, *Limnology and Oceanography*

Levitan, O., Dinamarca, J., Hochman, G., and Falkowski, P. G. (2014), Diatoms: a fossil fuel of the future, *Trends in Biotechnology*

Lykidis, A. (2007), Comparative genomics and evolution of eukaryotic phospholipid biosynthesis, *Progress in Lipid Research*

Martin, P., Van Mooy, B. A. S., Heithoff, A., and Dyhrman, S. T. (2011), Phosphorus supply drives rapid turnover of membrane phospholipids in the diatom Thalassiosira pseudonana, *The ISME Journal*

Meher, L. C., Vidya Sagar, D., and Naik, S. N. (2006), Technical aspects of biodiesel production by transesterificationa review, *Renewable and Sustainable Energy Reviews*

Moore, E. K., Hopmans, E. C., Rijpstra, W. I. C., Villanueva, L., Dedysh, S. N., Kulichevskaya, I. S., et al. (2013), Novel Mono-, Di-, and Trimethylornithine Membrane Lipids in Northern Wetland Planctomycetes, *Applied and Environmental Microbiology*

Tonon, T., Harvey, D., Larson, T. R., and Graham, I. A. (2002), Long chain polyunsaturated fatty acid production and partitioning to triacylglycerols in four microalgae, *Phytochemistry*

Tonon, T., Sayanova, O., Michaelson, L. V., Qing, R., Harvey, D., Larson, T. R., et al. (2005), Fatty acid desaturases from the microalga Thalassiosira pseudonana, *The FEBS journal*

Tyrrell, T. (1999), The relative influences of nitrogen and phosphorus on oceanic primary production, *Nature*

Van Mooy, B. A. S., Fredricks, H. F., Pedler, B. E., Dyhrman, S. T., Karl, D. M., Koblížek, M., et al. (2009), Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity, *Nature*

Volkman, J. K., Jeffrey, S. W., Nichols, P. D., Rogers, G. I., and Garland, C. D. (1989), Fatty acid and lipid composition of 10 species of microalgae used in mariculture, *Journal of Experimental Marine Biology and Ecology*

Yu, E. T., Zendejas, F. J., Lane, P. D., Gaucher, S., Simmons, B. A., and Lane, T. W. (2009), Triacylglycerol accumulation and profiling in the model diatoms Thalassiosira pseudonana and Phaeodactylum tricornutum (Baccilariophyceae) during starvation, *Journal of Applied Phycology*

Chapter 5

Targeted and Untargeted Lipidomics of *Emiliania huxleyi* Viral Infection and Life Cycle Phases Highlights Molecular Biomarkers of Infection, Susceptibility, and Ploidy

Jonathan E. Hunter[1,2], Miguel J. Frada[3], Helen F. Fredricks[4], Assaf Vardi[3] and Benjamin A. S. Van Mooy[4].

1. Ocean and Earth Science, University of Southampton, National Oceanography Centre Southampton, European Way, Southampton, SO14 3ZH, United Kingdom
2. Institute for Life Sciences, University of Southampton, SO17 1BJ, United Kingdom
3. Department of Plant and Environmental Sciences, Weizmann Institute of Science, Rehovot 76100, Israel
4. Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, 02543, United States of America.

5.1 Author Contributions

Jonathan E. Hunter carried out sample preparation, analysis, data interpretation and wrote the manuscript. Miguel J. Frada designed the experiment and carried out the culturing. The remaining co-authors assisted with experimental design, interpretation and drafting the manuscript.

5.2 Abstract

Marine viruses that infect phytoplankton strongly influence the ecology and evolution of their hosts. *Emiliania huxleyi* is characterized by a biphasic life cycle composed of a diploid (2N) and haploid (1N) phase; diploid cells are susceptible to infection by specific coccolithoviruses, yet haploid cells are resistant. Glycosphingolipids (GSLs) play a role during infection, but their molecular distribution in haploid cells is unknown. We present mass spectrometric analyses of lipids from cultures of uninfected diploid, infected diploid, and uninfected haploid *E. huxleyi*. Known viral GSLs were present in the infected diploid cultures as expected, but surprisingly, trace amounts of viral GSLs were also detected in the uninfected haploid cells. Sialic-acid GSLs have been linked to viral susceptibility in diploid cells, but were found to be absent in the haploid cultures, suggesting a mechanism of haploid resistance to infection. Additional untargeted high-resolution mass spectrometry data processed via multivariate analysis unveiled a number of novel biomarkers of infected, non-infected, and haploid cells. These data expand our understanding on the dynamics of lipid metabolism during *E. huxleyi* host/virus interactions and highlight potential novel biomarkers for infection, susceptibility, and ploidy.

Keywords: Emiliania Huxleyi, Lipidomics, Coccolithovirus, Glycosphingolipid, Glycerolipids, Haploid.

5.3 Introduction

Emiliania huxleyi (Lohmann) is the numerically dominant coccolithophore in the modern oceans and an important component of phytoplankton assemblages, inhabiting all but extreme polar oceans. Moreover, it forms large, dense blooms in high-latitude coastal and shelf ecosystems that exert a critical impact upon the global carbon cycle and the earths climate (**Paasche**, 2001; **Tyrrell and Merico**, 2004; **Westbroek et al.**, 1993). In later stages these blooms become visible to satellites as large scale cell death leads to the mass shedding of highly scattering calcium carbonate coccoliths, that typically coat the surface of *E. huxleyi* cells (**Holligan et al.**, 1983; **Vardi et al.**, 2012; **Lehahn et al.**, 2014).

Marine viruses are the most abundant biological agents in the oceans (**Fuhrman**, 1999; **Suttle**, 2007). Specific, giant, lytic double-stranded DNA *E. huxleyi* viruses (EhV), belonging to the *phycodnaviridae* family that infect microalgae (**Van Etten et al.**, 2002) are heavily implicated in the decay of *E. huxleyi* blooms (**Bratbak et al.**, 1993; **Brussaard et al.**, 1996; **Vardi et al.**, 2012).

Viruses that induce host cell lysis are thought to release particulate carbon and other nutrients into the water column (the viral shunt), thus circumventing the export of particulate organic matter to the deep ocean by way of the biological pump (**Fuhrman**, 1999; **Suttle**, 2007; **Jover et al.**, 2014). The efficacy of the biological pump has direct implications upon atmospheric carbon dioxide (**Suttle**, 2007). Conversely, viral infection is also known to induce increased production of transparent exopolymer particles (TEP) in *E. huxleyi*, that accelerate the formation of sinking particulates, enhancing the biological pump and removing virus particles from the upper water column (**Vardi et al.**, 2012).

At the cellular level, as a large dsDNA virus with high metabolic demand for the building blocks of DNA, lipids and protein synthesis, EhV triggers a rapid remodelling of the host metabolism (**Rosenwasser et al.**, 2014; **Schatz et al.**, 2014). In particular, recent studies have highlighted the crucial role of membrane lipids in the progression and regulation of EhV infection (**Vardi et al.**, 2009, 2012; **Rosenwasser et al.**, 2014). Evidence from genome and microscopic investigation suggests that EhV86 utilises an animal-like,

membrane dependent infection strategy. EhV entry, by membrane fusion or endocytosis and the acquisition of host membrane lipids via budding (**Mackinder et al.**, 2009; **Schatz et al.**, 2014), seem to be localised to membrane lipid raft regions (**Rose et al.**, 2014). Furthermore, the EhV genome contains a cluster of genes composing a nearly complete sphingolipid biosynthetic pathway analogous to the host pathway (**Monier et al.**, 2009; **Wilson**, 2005).

Glycosphingolipids (GSLs) bearing a sphingoid base derived from palmitoyl-CoA host glycopshingolipids (hGSL) - are abundant in uninfected *E. huxleyi*. Under lytic infection however, viral glycosphingolipids (vGSLs) derived from myristoyl-CoA are synthesized *de novo*. vGSLs are known to play a role in the regulation of cell death in infected cells and are enriched in the membrane of newly formed virions (**Vardi et al.**, 2009; **Fulton et al.**, 2014; **Rosenwasser et al.**, 2014). These vGSLs were detected in coccolithophore populations in the North Atlantic, which highlights their potential as biomarkers for viral infection in the oceans (**Vardi et al.**, 2009, 2012).

A glycosphingolipid with a sialic acid modified glycosyl headgroup (sGSL) was recently described to have a direct relationship with susceptibility to infection in diploid (2N) cells. Across 11 strains of *E. huxleyi* (**Fulton et al.**, 2014), sGSL was only detected at greater than trace levels in susceptible host strains, but not in viral-resistant host strains. Given the presence of a sialidase gene in the EhV genome (**Wilson**, 2005), it has been speculated that sGSL is a target for hydrolysis by EhV sialidases and/or a ligand for attachment by EhV lectin proteins during infection (**Fulton et al.**, 2014). This mechanism is of an analogous fashion to a range of human viral pathogens including influenza (**Stray et al.**, 2000). Additionally, betaine-like glycerolipids (BLL) were also recently described whose fatty acid composition appears highly indicative of the progression of infection. It was reported that uninfected *E. huxleyi* BLL composition was almost exclusively C16:0/C22:6 and C18:1/C22:6. Under viral infection the total BLL composition shifted to contain 50% C22:6/22:6 (**Fulton et al.**, 2014).

A basic feature of *E. huxleyi* is the possession of a biphasic, haplodiplontic and heteromorphic life cycle comprising a diploid (2N) coccolith-bearing phase that is nonmotile

and involved in the formation of blooms and a contrasting haploid (1N) form bearing flagella and nonmineralized organic body-scales (**Green et al.**, 1996; **Houdan et al.**, 2004). Recent transcriptomic analyses have revealed a dramatic differentiation between 1N and 2N cells, with less than 50% of transcripts estimated to be shared between the two phases, unravelling a deep degree of physiological segregation (**von Dassow et al.**, 2009; **Rokitta et al.**, 2011). Interestingly, whilst the 2N form is generally susceptible to EhV infection, the 1N form appears completely resistant to EhV. Furthermore, when 2N *E. huxleyi* are subject to EhV infection, a transition toward the resistant 1N flagellated form is induced, likely allowing for continuity of *E. huxleyi* following viral bloom termination (**Frada et al.**, 2008, 2012).

Given the central role that membrane lipids play in the progression and regulation of EhV infection and the resistance to EhV infection exhibited by 1N *E. huxleyi*, a number of questions arise that are as yet unanswered. Early targeted analyses have shown similar compositions with respect to the major structural lipids, storage lipids and pigments between 1N and 2N cells in the case of a single strain of *E. huxleyi* (**Bell and Pond**, 1996). Minor or novel lipid classes however, are as yet uncharacterised in the 1N cell lipidome and it is unknown whether the susceptibility marker sGSL is absent from EhV resistant 1N cells. By characterising these lipids and the lipidome as a whole, we can gain insight into the mechanism of 1N *E. huxleyi* resistance to EhV infections and potentially highlight biomarkers of each of the life cycle phases.

We present herein, a detailed characterisation of the lipidomes of cultured *E. huxleyi*: an uninfected 2N strain (RCC 1216); the 2N strain under infection with coccolithovirus (RCC 1216 + EhV201); and an uninfected 1N strain (RCC 1217). Total lipid extracts derived from these cultures collected over 120 h post-infection were characterised by mass spectrometry. We used two approaches, utilising both targeted analysis for quantification of known GSL/glycerolipid species and untargeted analyses for screening for unknown lipids.

5.4 Results

5.4.1 Host Cell and Viral Dynamics and Relative Abundances

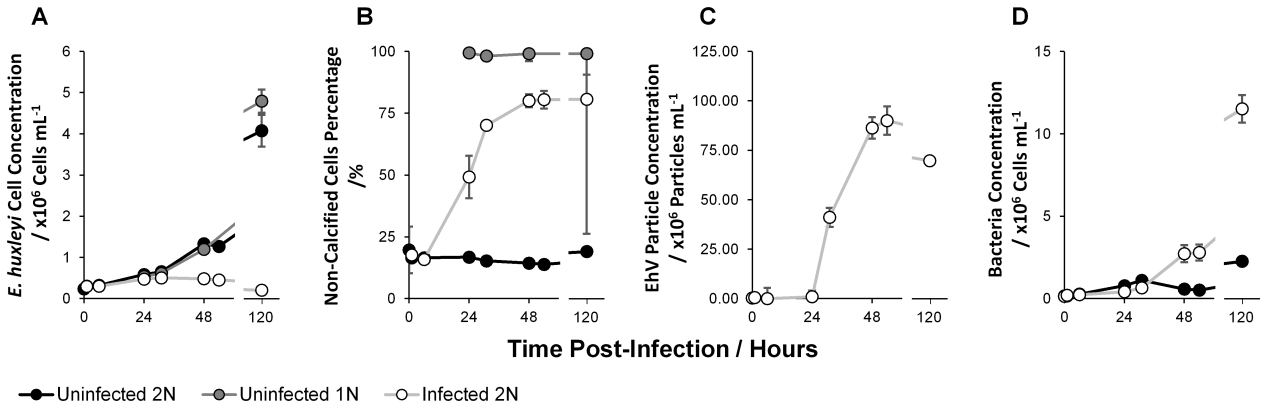


Figure 5.1: Host cell and viral dynamics during infection: Total Host Cell Concentration (A); Non-calcified Cell Percentage of Total (Host Cell, B); EhV Particle Concentration (C); Bacterial Cell Concentration (D). Data represent the mean average of two biological replicates with error bars of one standard deviation.

Uninfected 2N and uninfected 1N cultures grew at a comparable average rate of 0.023 h^{-1}, reaching a population density of $4.07 \times 10^6 \pm 0.39 \times 10^6$ cells mL^{-1} and $4.79 \times 10^6 \pm 0.28 \times 10^6$ cells mL^{-1} respectively within 120 hours (Figure 5.1A). In contrast, the infected 2N culture population peaked at $5.07 \times 10^5 \pm 1.15 \times 10^4$ cells mL^{-1} at 31 hours and rapidly declined thereafter concomitant with the emergence of EhV particles in the medium. EhV concentration peaked at 54h at $9.17 \times 10^7 \pm 0.71 \times 10^7$ virions mL^{-1} in the infected 2N cultures (Figure 5.1C). The decline phase was accompanied by an increase in the percentage of non-calcified (low scatter) cells from 17.70 ± 0.73 % at 0 h to 80.05 ± 2.63 % at 48 h. Bacteria concentration (Figure 5.1D), whilst remaining low in the uninfected 2N control, increased at a rate of 0.037 h^{-1} in the infected 2N cultures after the onset of infection reaching a maximum of $1.15 \times 10^7 \pm 8.39 \times 10^5$ cells mL^{-1} at 120 h.

5.4.2 Glycerolipid Targeted Lipidomics

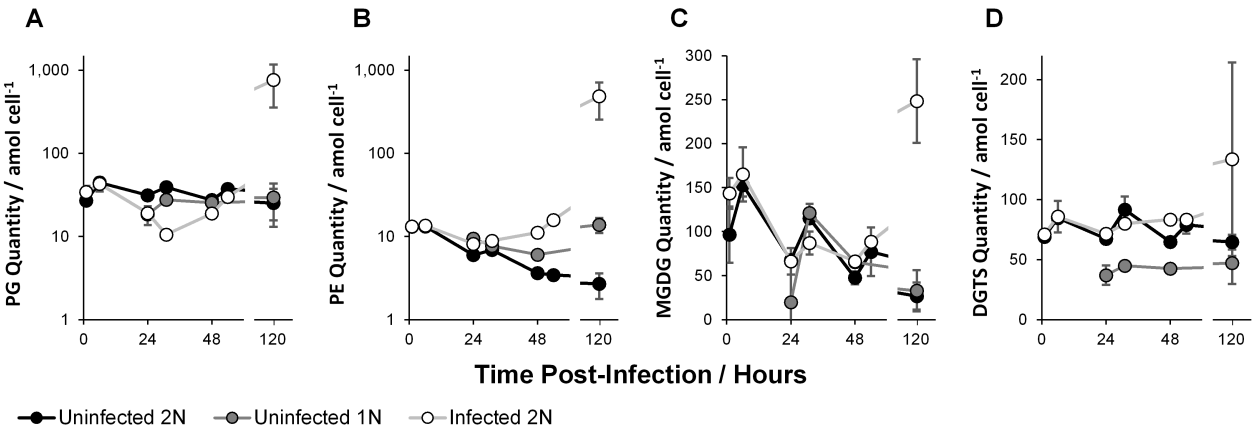


Figure 5.2: Polar glycerolipid quantity per *E. huxleyi* cell: Glycerophosphatidylglycerol (PG, A); Glycerophosphatidylethanolamine (PE, B); Monogalactosyldiacylglycerol (MGDG, C) and Diacylglyceryltrimethylhomoserine (DGTS, D). Data represent the mean average of two biological replicates with error bars of one standard deviation. Statistically significant variations between cultures are described in the following section.

The polar glycerolipids glycerophosphatidylcholine (PC), diacylglyceryl-3-O-carboxy-(hydroxymethyl)-choline (DGCC), digalactosyldiacylglycerol (DGDG), sulfoquinovosyldiacylglycerol (SQDG), and the sulphur containing phospholipid phosphatidyl-S,S-dimethylpropanethiol (PDPT) (**Fulton et al.**, 2014) did not vary between cultures (Supplementary Figure 7.12). However, several other classes of polar glycerolipid showed interesting differences.

Phosphatidylglycerol (PG) quantity per cell (Figure 5.2) in the uninfected 2N and uninfected 1N cultures was statistically similar through time. PG per cell in the infected 2N cultures initially declined by 0.27 ± 0.023* fold at 31 h relative to the uninfected 2N control. A large increase in infected 2N PG per cell of 30.29 ± 21.80 fold was observed at 120 h, although relative to the uninfected 2N control cultures the increase was not statistically significant ($p = 0.12$), due to a large variation between the biological replicates at this time point.

Phosphatidylethanolamine (PE) quantity per cell (Figure 5.2B) showed a decreasing

trend through time in the uninfected 1N and uninfected 2N cultures. PE was slightly elevated in the uninfected 1N cultures compared to the uninfected 2N control, this was statistically significant only at 24 h, where PE per cell in the uninfected 1N was 1.57 ± 0.16* fold higher than in uninfected 2N. PE quantity per cell in the infected 2N cultures diverged from the uninfected 2N control after 24 h, progressively increasing to 4.58 ± 0.18** fold greater at 54 h. In common with PG, a large increase of infected 2N PE per cell at 120 h, of 180.10 ± 105.19 fold, was not statistically significant ($p = 0.10$), due to variability in the replicate samples at this time point.

Two non-phosphorous polar glycerolipids showed interesting and significant dynamics. Monogalactosyldiacylglycerol (MGDG) quantity per cell (Figure 5.2C) was statistically similar in all of the cultures before 52 h, but showed a decreasing trend with time. However, at 120 h MGDG in the infected 2N cultures increased by 9.31 ± 5.68* fold, relative to the uninfected 2N cultures. Diacylglyceryltrimethylhomoserine (DGTS) quantity per cell (Figure 5.2D) appeared to be approximately half as abundant in the uninfected 1N compared with the uninfected 2N control cultures at 0.55 ± 0.12* fold less at 24 h and 0.66 ± 0.01** fold less at 48 h. DGTS per cell in the infected 2N cultures did not vary statistically significantly from the uninfected 2N control, with the exception of a 1.29 ± 0.02** fold increase at 48 h. Similar to PG, DGTS levels in the uninfected 2N and uninfected 1N cultures remained relatively consistent through time.

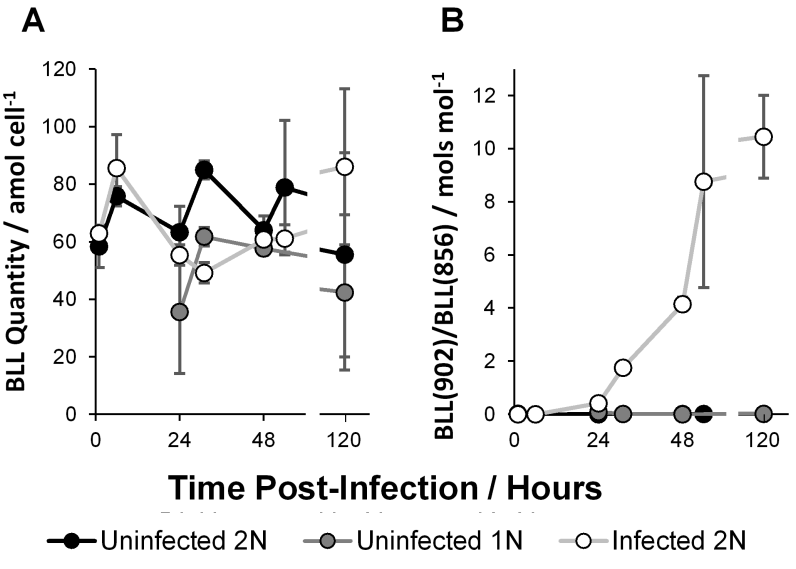


Figure 5.3: Betaine-Like Lipid (BLL), total quantity per cell (A) and ratio of BLL(902; 22:6/22:6) quantity to BLL(856; 18:1/22:6) quantity (B). Data represent the mean average of two biological replicates with error bars of one standard deviation. Statistically significant variations between cultures are described in the following section.

Total Betaine-Like Lipid (BLL; (**Fulton et al.**, 2014)) quantity per cell (Figure 5.3A) varied considerably in all cultures, ranging from a minimum of 35.55 ± 21.35 amol cell^{-1} at 24 h in the uninfected 1N cultures to a maximum of 86.07 ± 27.21 amol cell^{-1} at 120 h in the infected 2N cultures. Only at 31 h were significant differences observed, where the infected 2N cultures were 0.58 ± 0.05 fold less and the uninfected 1N cultures were 0.73 ± 0.05* fold less than in the uninfected 2N control. No clear trend was evident in the total BLL quantity per cell, in response to time or between the different cultures. However, the ratio between two BLL molecular species, BLL(22:6/22:6) and BLL(18:1/22:6) showed marked differences (Figure 3B). In the uninfected 2N and uninfected 1N cultures, the ratio was < 0.1 throughout, representing the relative absence of BLL(22:6/22:6) compared to BLL(18:1/22:6) when uninfected. The ratio of these species rose rapidly under infection reaching 10.46 ± 1.56 at 120 h after inoculation of the infected 2N cultures.

5.4.3 Glycosphingolipid Targeted Lipidomics

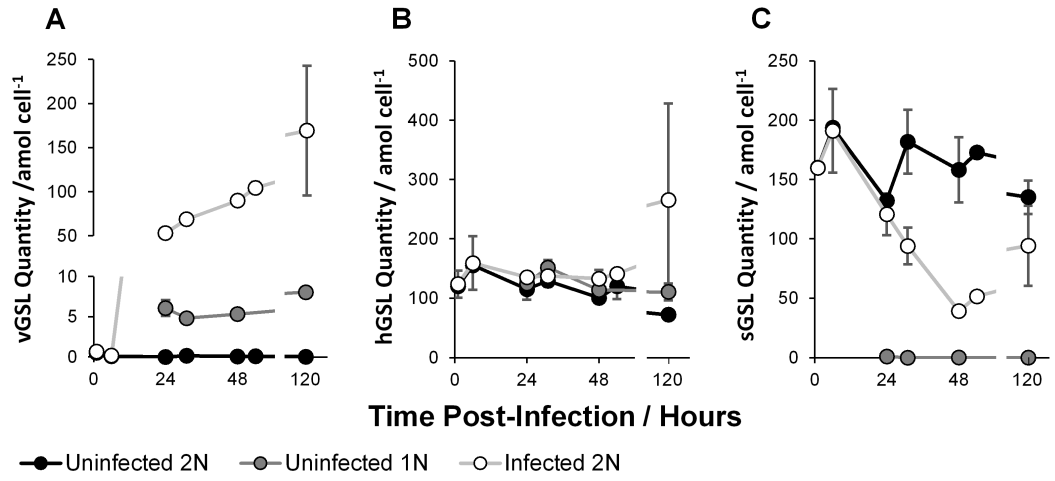


Figure 5.4: Glycosphingolipid quantity per *E. huxleyi* cell: Viral glycosphingolipid (vGSL, A); Host glycosphingolipid (hGSL, B); Sialic-acid glycosphingolipid (sGSL, C). Data represent the mean average of two biological replicates with error bars of one standard deviation. Statistically significant variations between cultures are described in the following section.

Concentrations of two classes of glycosphingolipids were highly dynamic through the course of infection. The infection marker viral glycosphingolipid (vGSL, Figure 5.4A) was abundant only in the infected 2N cultures. Infected 2N vGSL concentration rose consistently through time from absent at 0 h to 169.32 ± 73.77 amol cell^{-1} at 120 h, concomitant with the increase in EhV concentrations. Interestingly, lower levels of vGSL (4-8 amol cell^{-1}) were also detected in the uninfected 1N cultures, reaching a maximum of 7.63 ± 0.13 amol cell^{-1} at 120 h.

The susceptibility marker sialic glycosphingolipid (sGSL, Figure 5.4C) was abundant in the uninfected 2N control, with no clear time dependence, between a range of 132.32 ± 4.09 amol cell^{-1} at 24 h and 193.76 ± 3.20 amol cell^{-1} at 6 h. The concentrations of sGSL in the infected 2N cultures displayed an approximately decreasing trend from 191.07 ± 35.35 amol cell^{-1} at 6 h to 51.59 ± 0.57 amol cell^{-1} at 54 h. Importantly, sGSL was absent from the uninfected 1N cultures. In contrast to vGSL and sGSL, a third class of glycosphingolipid, the intrinsic host glycosphingolipid (hGSL, Figure 5.4B) presented

no statistically significant variation between cultures, and no trend was evident with the progression of time.

5.4.4 Untargeted Lipidomics and Biomarker Selection

Untargeted lipidomic data were collected from the incubations at 48 h and subsequent Partial Least Squares Discriminant Analysis (PLS-DA) revealed a number of potential biomarkers for *E. huxleyi* life stages and viral infection. The top 5 loadings that were indicative for each culture type, bearing significant differences by univariate statistics ($p=<0.05$ by single factor ANOVA), were assigned by database matching (Figure 5.5). All ion fragmentation MS2 data was used to confirm database hits (Supplementary Figure 7.14 and 7.15).

A - Positive Ions

Selection	Rank	M/Z	Retention Time (Min)	Identity	Adduct	P	Normalised Abundance Uninfected 2N	Uninfected 1N	Infected 2N
Uninfected 2N	1	872.6824	15.8	sGSL + H_2	(M+H)+	*	0.81	0.00	0.19
	2	868.6510	14.4	sGSL - H_2	(M+H)+	*	0.79	0.00	0.21
	3	870.6666	15.2	sGSL	(M+H)+	*	0.78	0.00	0.22
	4	892.6484	15.3	sGSL	(M+Na)+	*	0.72	0.02	0.26
	5	752.5226	12.3	PC(34:5)	(M+H)+	*	0.60	0.15	0.25
Uninfected 1N	1	549.5356	15.3	WE(31:0)	(M+2AcN+H)+	*	0.00	1.00	0.00
	2	508.5091	15.3	WE(31:0)	(M+AcN+H)+	**	0.00	1.00	0.00
	3	802.6768	17.3	GSL(t40:0)	(M+H)+	**	0.04	0.67	0.29
	4	800.6614	17.0	GSL(t40:1)	(M+H)+	**	0.05	0.65	0.30
	5	738.5155	11.9	MGDG(32:5)	(M+NH4)+	**	0.22	0.70	0.09
Infected 2N	1	840.7079	18.2	TAG(50:6)	(M+NH4)+	*	0.00	0.00	1.00
	2	858.7546	19.0	TAG(51:4)	(M+NH4)+	*	0.00	0.00	1.00
	3	799.6788	19.4	TAG(46:1)	(M+Na)+	*	0.00	0.00	1.00
	4	885.7650	18.7	TAG(50:4)	(M+AcN+NH4)+	*	0.00	0.00	1.00
	5	856.7391	18.6	TAG(51:5)	(M+NH4)+	*	0.00	0.00	1.00

B - Negative Ions

Uninfected 2N	1	870.6694	15.8	sGSL + H_2	(M-H)-	*	0.82	0.00	0.18
	2	868.6532	15.2	sGSL	(M-H)-	*	0.79	0.00	0.21
	3	854.5800	12.6	BLL(18:1/22:6)	(M-H)-	*	0.61	0.27	0.13
	4	827.4963	10.9	MGDG(36:9)	(M+HAc-H)-	**	0.60	0.23	0.17
	5	949.5362	14.9	SQDG(40:8)	(M+HAc-H)-	**	0.58	0.26	0.16
Uninfected 1N	1	623.3450	6.2	LMGDG(22:6)	(M+HAc-H)-	*	0.06	0.87	0.08
	2	696.6158	17.3	Cer(d18:1/22:0(OH))	(M+HAc-H)-	**	0.03	0.70	0.27
	3	694.6004	17.1	Cer(d18:1/22:1(OH))	(M+HAc-H)-	**	0.06	0.61	0.33
	4	941.5497	11.2	DGDG(32:5)	(M+HAc-H)-	*	0.19	0.71	0.11
	5	833.4686	12.0	DGDG(30:6)	(M-H2O-H)-	*	0.22	0.69	0.09
Infected 2N	1	838.6196	16.3	vGSL	(M+Cl)-	**	0.00	0.00	1.00
	2	802.6404	16.3	vGSL	(M-H)-	**	0.00	0.00	1.00
	3	900.5646	10.9	BLL(22:6/22:6)	(M-H)-	**	0.01	0.01	0.98
	4	816.6564	16.7	vGSL + CH_2	(M-H)-	**	0.00	0.03	0.97
	5	531.5149	18.0	WE(36:2)	(M-H)-	**	0.06	0.07	0.86

Figure 5.5: Top 5 loadings (assigned species only) for each culture type in the PLS-DA model: Positive ions (A) and Negative ions (B). Normalised abundance represents the average abundance of an ion in the specified culture type divided by the total abundance of that ion in all three cultures. M/Z represents the mass to charge ratio of a given molecular ion. AcN and HAc represent acetonitrile and acetate adducts respectively. *$p<0.05$, **$p<0.005$ by single factor ANOVA.

The potential biomarkers for the 2N control cultures were sGSLs, bearing various adducts and plus or minus one double bond resulting in the highest positive loadings on PC2. In positive ion mode (Figure 5.5A), tentatively assigned sGSL+H_2, sGSL-H_2, and sGSL (with two different adducts) gave normalized abundances of 0.81, 0.79, 0.78, and 0.72 respectively, where the sGSL of calculated m/z 870.6665 $[M+H]^+$ is regarded as the archetypal sGSL (Fulton et al., 2014). A 752.5226 m/z species, with a database hit to PC(34:5) showed a normalized abundance of 0.60. In negative ion mode (Figure 5.5B), sGSL+H_2, sGSL and BLL(18:1/22:6) showed a normalized abundance of 0.82, 0.79, and 0.61 respectively. Species of m/z 827.4963 and 949.5362 with hits to MGDG(36:9) and SQDG(40:8) respectively showed normalized abundances of 0.60 and 0.58.

The top potential biomarker for uninfected 1N cells, in positive ion mode, was a wax ester/alkanoate WE(31:0) with two and one acetonitrile adducts, which was detected solely in the 1N cultures yielding normalized abundances of 1.00 and 1.00 respectively. Two GSLs, GSL(t40:0) and GSL(t40:1), distinct to previously documented GSLs in the *E. huxleyi* host/virus system, had normalized abundances of 0.67 and 0.65 respectively. Finally, MGDG(32:5) was ranked fifth with a normalized abundance of 0.70. In negative ion mode, a lyso-MGDG (bearing a single fatty acid rather than a pair), LMGDG(22:6) was most strongly indicative of the uninfected 1N cultures with a normalized abundance of 0.87. Ceramide species, chemically similar to sGSL but lacking any headgroup and bearing a hydroxy fatty amide, Cer(d18:1/22:0(OH)) and Cer(d18:1/22:1(OH)) in the uninfected 1N cultures had normalised abundances of 0.70 and 0.61 respectively. Lastly, digalactosyldiacylglycerols and DGDG(32:5) and DGDG(30:6) were enriched in the uninfected 1N cultures with normalised abundances of 0.71 and 0.69 respectively.

The potential biomarkers for the infected 2N cells were highly diagnostic, with 7 out of the top 10 yielding normalized abundances of 1.00. In positive ion mode, a series of triacylglycerols (TAGs) gave the highest negative loadings on PC1 and were ranked highest. TAG(50:6); TAG(51:4); TAG(46:1); TAG(48:2) and TAG(51:5) each had normalised abundances of 1.00. In negative ion mode, previously described biomarkers vGSL (with a chloride adduct); vGSL; BLL(22:6/22:6) and vGSL + CH_2 had normalised abundances of

1.00, 1.00, 0.98, 0.97, and 0.86 respectively, where vGSL of calculated m/z 802.6414 [M-H]$^-$ is regarded as the archetypal vGSL (**Vardi et al.**, 2009). A fatty acid/wax ester like FA/WE(36:2) species was ranked fifth in negative ion mode with a normalised abundance of 0.86.

Further to the assigned species (Figure 5.5) there were many molecular species of interest that could not be assigned by accurate mass database matching (Supplementary Figure 7.16)

5.5 Discussion

5.5.1 Host Cell and Viral Dynamics and Relative Abundances

Uninfected 2N and uninfected 1N *E. huxleyi* populations increased exponentially through time with a similar growth rate and maximum population, but the infected 2N cultures peaked at 31 h and declined thereafter, consistent with typical lytic viral infection by EhV (**Bratbak et al.**, 1993; **Wilson et al.**, 2002) (Figure 5.1A). This decline was concomitant with a rapid increase in EhV particles and non-calcified, low side scatter *E. huxleyi* cell concentration in the growth media, indicating both viral burst from infected cells and the demise of 2N calcified cells (Figure 5.1B and C). These growth dynamics reflect identical trends to those previously described (**Frada et al.**, 2008, 2012). High scatter cells were predominate in the uninfected 2N control, representing the coccolith bearing 2N phase. Conversely, cells in the uninfected 1N cultures were mostly low scatter, indicative of the non-calcifying 1N phase.

The cultures used in this study were non-axenic, and a rapid increase in bacterial numbers in the infected 2N cultures was likely fuelled by the release of organic carbon during cell lysis. However, in terms of overall biomass, the level of contamination from bacteria is comparatively minor: considering the cell concentrations of *E. huxlyei* and bacteria, and cellular carbon quotas of 2 pmol C $cell^{-1}$ and 1 fmol C $cell^{-1}$, respectively (**Gundersen et al.**, 2002; **Borchard and Engel**, 2012), *E. huxleyi* dominate biomass by approximately a factor of 20. Based on the cell counts of the uninfected 2N and uninfected 1N cultures, we would expect this factor to be orders of magnitude higher. Thus, except in the case of a few specific lipid molecules (discussed below) we would expect the lipidomes of our samples to be overwhelmingly dominated by contributions from *E. huxleyi*.

5.5.2 Glycerolipid Targeted Lipidomics

For the most part, polar glycerolipids per cell were largely invariant between cultures and with time, with only a few notable exceptions. Despite the relatively small contributions of bacterial biomass, the observed increase in PG concentration in the infected 2N cultures

(Figure 5.2A) was most likely due to an increase in the bacterial population (Figure 5.1D). This result is not entirely unexpected since PG generally composes about half of the glycerolipids in marine bacteria, while in *E. huxleyi* PG only represents 0.2% of the lipidome (**Popendorf et al.**, 2011b; **Carini et al.**, 2015).

In the infected 2N cultures, PE quantity also increased compared to the uninfected 2N control after 24 h (Figure 5.2B). Like PG, PE is also a major component of bacterial membranes but is scarce in *E. huxleyi* (**Van Mooy et al.**, 2009; **Popendorf et al.**, 2011b; **Carini et al.**, 2015), and thus a contribution from bacteria is possible (**Fulton et al.**, 2014). Alternatively, this increase in PE may be derived from the host cells autophagy machinery (**Schatz et al.**, 2014). Autophagy is a highly conserved eukaryotic mechanism for the degradation of damaged organelles and unwanted macromolecules (**Shemi et al.**, 2015). It has been demonstrated that autophagy is induced during the lytic phase of EhV infection and plays a role in the propagation of the virus (**Schatz et al.**, 2014). A well-known hallmark of autophagy is PE lipidation of the Atg8 protein, as also reported in *E. huxleyi* under EhV infection (**Schatz et al.**, 2014). Thus, the observed increase in PE under EhV infection may be attributable to the induction of autophagy via lipidation of the Atg8 protein. Interestingly, total PE appeared slightly more abundant per cell in the uninfected 1N cultures, relative to the uninfected 2N control.

MGDG quantity per cell was similar between uninfected 2N and uninfected 1N cultures, while the infected 2N cultures showed a large increase in MGDG quantity per cell at 120 h post inoculation (Figure 5.2C). MGDG is often associated with the thylakoid membranes of the chloroplast (**Sakurai et al.**, 2006) and hence variation may indicate a consequence of infection upon the host photosynthetic machinery. In contrast, a decrease in MGDG under EhV infection of *E. huxleyi* has been shown elsewhere (**Fulton et al.**, 2014). Furthermore, a recent study on another haptophyte *Haptolina ericina* infected with a dsDNA virus observed a decrease in glycolipid abundance under infection (**Ray et al.**, 2014). This decrease was suggested to derive from cell disruption during infection, resulting in a loss of chloroplasts during sample preparation. The haptophyte *Phaeocystis pouchetii* did not display a major loss of cellular MGDG when infected with a dsDNA virus (**Ray et al.**,

2014). MGDG is one of the most abundant lipids in *E. huxleyi* (**Fulton et al.**, 2014), but this molecule is scarce in bacteria because they have a clear preference for synthesizing phospholipids over glycolipids when phosphate is abundant, as in the K/2 medium used here (**Popendorf et al.**, 2011b; **Carini et al.**, 2015); thus, it is unlikely that the MGDG observed in the infected 2N cultures is of bacterial origin. Clearly at this stage we have an incomplete understanding of the behaviour of MGDG under viral infection in haptophytes such as *E. huxleyi*.

Total DGTS abundance per cell in the uninfected 1N cultures was consistently around half of its abundance in the uninfected 2N control (Figure 5.2D). Betaine lipids such as DGTS are known to substitute for membrane phospholipids (primarily PC) under phosphorus stress in many eukaryotic phytoplankton. As such the BL/PC ratio has been considered as a measure for phosphorus limitation in the environment (**Van Mooy et al.**, 2009; **Martin et al.**, 2011). Reduced DGTS abundance from 1N cells within an environmental population would impact such a measure. Furthermore, when coupled with a slight increase in PE in the 1N case, outlined earlier, this may shift the glycerolipid stoichiometry to an N : P ratio lower than the 2N counterpart (**Van Mooy et al.**, 2009; **Carini et al.**, 2015).

Whilst total BLL quantity per cell was generally similar (Figure 5.3A), BLL(22:6/22:6) was present only at trace level in the uninfected 2N cultures and increased dramatically in abundance in the infected 2N cultures. BLL(18:1/22:6) was present in the uninfected 2N and uninfected 1N cultures and we found that the ratio between BLL(22:6/22:6)/BLL(18:1/22:6) was a strong indicator of infection in these experiments (Figure 5.3B). We suggest that in light of these and previous observations (**Fulton et al.**, 2014), that BLL(22:6/22:6) and ratios thereof may be a useful biomarker of EhV infection to compliment vGSL. Application as a biomarker would require further investigation to verify BLL(22:6/22:6) presence/absence under the infected/host states is conserved across different *E. huxleyi* and viral strains and specific to *E. huxleyi*. The presence/absence relationship has been to date demonstrated in *E. huxleyi* 1216 (2N), 1217 (1N) and 1216 with EhV201 in this study and *E. huxleyi* 374 (2N) with EhV86 (**Fulton et al.**, 2014).

Notably, the polar glycerolipids PC, DGCC, DGDG, SQDG and PDPT were effectively identical in all three cultures through time (Supplementary Figure 7.12). This observation, in conjunction with the many similarities in PG, PE and MGDG indicates that the overall polar glycerolipid content is not radically affected by life cycle or even viral infection, except potentially at the very termination of the lytic phase; only DGTS and BLL (22:6/22:6) show notable earlier variations. **Bell and Pond** (1996) observed that the distributions of fatty acids within polar glycerolipids were also fairly similar in uninfected 2N and 1N cells. Thus, overall these molecules seem to play a largely passive role in life cycle differentiation and infection. Furthermore, since the cellular lipid content of the major glycerolipids is mostly similar in all three cultures and cellular lipid content is a first-order approximation of cellular biomass, these data validate our assumption that the different types of cells do not vary greatly in size (*i.e.* biomass content) and support our decision to normalize all lipid data to cell abundance.

5.5.3 Glycosphingolipid Targeted Lipidomics

In contrast to the polar glycerolipids, the glycosphingolipids were highly dynamic, showing marked differences between culture treatments and through time. The infection marker vGSL was absent from the uninfected 2N cultures and abundant in the infected 2N cultures, rising through time with the progression of infection (Figure 5.4A), in agreement with previous observations (**Vardi et al.**, 2009, 2012). Interestingly, low levels of vGSL-like molecules were detected in the uninfected 1N cultures. These vGSL-like species produce the same diagnostic glycosyl headgroup fragments and co-elute with previously characterised vGSLs. Untargeted lipidomic data, discussed below, suggest however that these molecules may not be exactly the same as the vGSLs in the infected diploid cultures originally recognized by **Vardi et al.** (2009). Future developments in targeted lipidomic methods hold the promise of being able to distinguish true vGSL from infected diploid cells and this novel vGSL-like molecule from uninfected haploid cells.

Host intrinsic hGSL did not vary in quantity per cell between the uninfected 2N control and the infected 2N cultures, in line with previous studies (**Vardi et al.**, 2009, 2012; **Fulton**

et al., 2014). Furthermore, we find that hGSL cellular abundance is also statistically similar in the uninfected 1N cultures (Figure 5.4B).

The marker lipid sGSL was recently described as indicative of susceptibility to viral infection in *E. huxleyi* (**Fulton et al.**, 2014) and we found that sGSL was abundant in the virus-susceptible 2N strain used in this study, RCC1216. Levels of sGSL were somewhat reduced with the progression of infection in the infected 2N cultures at intermediate time, before returning to a level statistically similar to that of the uninfected control (Figure 5.4C). This contrasts with a previous study with a different *E. huxleyi* strain (**Fulton et al.**, 2014) showing an increase in sGSL cellular abundance under infection. The same report however, included mesocosm field experiments that showed decreases in sGSL abundance per cell subject to EhV infection (**Fulton et al.**, 2014); these results were interpreted as preferential infection of sGSL rich cells and relative growth of sGSL poor virus resistant cells. In light of our finding that sGSL is absent in 1N cells, the interpretation of **Fulton et al.** (2014) is entirely consistent with the observations of **Frada et al.** (2012) from the same mesocosms showing that 1N cells became proportionally more abundant, an affirmation of the Cheshire Cat hypothesis (**Frada et al.**, 2008). Although the dynamics of sGSL subject to EhV infection appear variable between 2N systems (**Fulton et al.**, 2014), the discovery here that sGSL is absent in 1N *E. huxleyi* cells could make sGSL an even more important biomarker for viral infection dynamics and associated shifts in ploidy.

Since 1N cells are resistant to viral infection (**Frada et al.**, 2008) and sGSL is linked with susceptibility in several strains of 2N *E. huxleyi* (**Fulton et al.**, 2014), the absence of sGSL from 1N cells further implicates sGSL as playing a key role in viral infection (**Fulton et al.**, 2014). Sialic acid moieties similar to that of sGSL are implicated as ligands for viral attachment in the cyanobacterium *Prochlorococcus* (**Avrani et al.**, 2011; **Fulton et al.**, 2014). It has been speculated therefore, that viral attachment and entry, facilitated by sGSL, may be the means by which this resistance is achieved (**Fulton et al.**, 2014). Furthermore, it has been speculated that sGSL may have intrinsic function in the production or externalisation of the coccoliths that characterise the exterior of the calcifying 2N phase (**Fulton et al.**, 2014). It is noteworthy therefore, that the 1N phase

that does not produce sGSL also does not produce coccoliths, further suggesting a link between sGSL and calcification.

5.5.4 Untargeted Lipidomics and Biomarker Selection

By using PLS-DA to analyse the untargeted lipidomic data, we have highlighted molecular species diagnostic of each 1N, 2N, and infected 2N cells (Figure 5.5). These species were assigned identities based upon matches between accurate m/z and an extensive database of lipids. The included assignments are supported by secondary MS2 data (Supplementary Figures 7.14 and 7.15). Due to the nature of the all ion fragmentation MS2 method applied in the interest of unbiased methodology, incorrect assignments of co-eluting species may occur in some cases. As such the assigned lipid identities in the unbiased analyses are presented as tentative. The differential abundances of each species give a quantitative representation for the discovery of potential biomarkers.

For the uninfected 2N control, the known susceptibility biomarker sGSL (**Fulton et al.**, 2014) and related species were the top ranked assigned hits, in both positive and negative ion mode. This observation is in agreement with the targeted data, where sGSL was absent from the uninfected 1N cultures, and reduced in abundance in the infected 2N case. In negative ion mode, BLL(18:1/22:6) was ranked 3rd for the uninfected 2N cultures. BLL(18:1/22:6) is a known marker of *E. huxleyi*, whose concentration has been shown to drop under viral infection (**Fulton et al.**, 2014). At the present time, sGSL and BLL (18:1/22:6) have not been reported in any other species than *E. huxleyi*. By contrast, PC(34:5), MGDG(36:9), and SQDG(40:8) which were ranked 5th in positive, 4th and 5th in negative ion mode respectively, are common polar membrane glycerolipids in eukaryotic microalgae (**Popendorf et al.**, 2011b; **Brandsma et al.**, 2012). These observations underscore the challenge of developing these biomarkers for use in the field: a molecule must be both diagnostic for *E. huxleyi* viral infection and ploidy and absent in all other marine microbes. The former criterion is addressed with the data presented here, but the latter will only be met after the lipidomes of many additional marine taxa are examined using the techniques similar to those employed here. In the meantime, the biomarkers we

identified by untargeted lipidomics should be applied with due discretion.

In the uninfected 1N cultures, the top hits in positive ion mode appeared to be a form of 31:0 wax ester or alkenoate. It is possible that changes in this molecule could reflect differences in the amounts of long chain alkenones in 1N cells. Alkenones are widely used in palaeoceanography temperature reconstructions (**Prahl and Wakeham**, 1987; **Volkman et al.**, 1980). Direct comparisons between this 31:0 wax ester/alkenoate and alkenones remain to be conducted.

The 3rd and 4th ranked species in 1N cultures in positive ion mode were matched to a GSL(t40:0) and GSL(t40:1) species. These GSLs appear to share some structural similarities with archetypal forms of hGSL (GSL(d41:5(OH)) and vGSL (GSL(t39:0(OH)) identified by (**Vardi et al.**, 2009, 2012). These haploid specific GSLs may be the source of the observed vGSL-like species in the targeted uninfected haploid analyses, discussed previously. In light of the aforementioned targeted GSL data, these haploid GSLs certainly warrant further characterization. In negative ion mode, Cer(d18:1/22:0(OH)) and Cer(d18:1/22:1(OH)) hydroxyceramide species were ranked 2nd and 3rd for the uninfected 1N cultures. These ceramides are effectively the same chemical composition as GSL minus the glycosyl headgroup. Furthermore, it is likely that the GSL(t40:1) highlighted is in fact the Cer(d18:1/22:0(OH)) ceramide moiety with a glycosyl headgroup, although this could not be confirmed from the MS2 data.

Interestingly, upregulation of enzymes responsible for the production of ceramides from the hydrolysis of the beta-glycosidic linkages of GSLs is a characteristic response of 2N cells to viral infection (**Rosenwasser et al.**, 2014). To our knowledge, the GSL(t40:0) and GSL(t40:1) species have not been previously reported in any marine microbial organism, although they are closely related chemically to the other GSLs. Thus, these GSL species represent, pending further validation and characterisation, candidate biomarkers for 1N *E. huxleyi*. 1N *E. huxleyi* are not easily identifiable in field samples by classic microscopy techniques (**von Dassow et al.**, 2009) and lipid biomarkers have particular utility in bulk environmental samples, such as filtered particulate organic material collected from the water column or sinking particulate trap matter, where other methods may not be

applicable.

Finally, a number of glyceroglycolipids appeared indicative of the uninfected 1N cultures. MGDG(32:5) was ranked 5th in positive ion mode and LMGDG(22:6), DGDG(32:5) and DGDG(30:6) were ranked 1st, 4th and 5th in negative ion mode. Glyceroglycolipids are typically associated with the thylakoid membranes of the chloroplast (**Petroutsos et al.**, 2014). Our analysis highlights differences in these photosynthetic membrane lipids and alludes to some implication of ploidy upon the photosynthetic apparatus. However, these are fairly common molecules that are unlikely to be of much utility as biomarkers for *E. huxleyi* viral infection or ploidy (**Popendorf et al.**, 2011b; **Brandsma et al.**, 2012).

For the infected 2N cultures, the PLS-DA model highlights a series of TAGs found under positive ionization as indicative. TAG biosynthesis is commonly upregulated in phytoplankton in response to a variety of stressors, and TAGs are found primarily in lipid bodies. TAGs are unlikely to be attributable to heterotrophic bacterial contaminants as they are not generally produced in significant quantities by bacteria (**Alvarez and Steinbüchel**, 2002). It has been proposed that EhV capsid assembly occurs via a mechanism similar to that of the Hepatitis C virus (**Herker and Ott**, 2011; **Fulton et al.**, 2014), where lipids are incorporated from the interface of lipid bodies and the endoplasmic reticulum. In uninfected 2N *E. huxleyi*, these lipid bodies are composed mainly of alkenones, alkenoates, and small quantities of TAGs and other lipids (**Eltgroth et al.**, 2005). Research on human hepatocyte cells shows that TAGs are critical to Hepatitis C capsid assembly (**Liefhebber et al.**, 2014). Thus, we speculate that the observed increase in a number of TAG species in *E. huxleyi* under EhV infection may result from the virus upregulating TAG biosynthesis in order to assist capsid assembly.

In negative ion mode, vGSL and BLL(22:6/22:6) represent the top four assigned species indicative of the infected cultures, in line with previous observations (**Vardi et al.**, 2012; **Fulton et al.**, 2014) and the targeted data discussed previously, while a 36:2 wax ester/alkenoate species is ranked fifth. None of these molecules are likely be derived from contaminating bacteria in the cultures, but instead are of decidedly coccolithophore origin (**Eltgroth et al.**, 2005; **Fulton et al.**, 2014; **Ray et al.**, 2014). A molecule with

the exact same elemental formula, 36:2 methyl alkenoate, was identified in *E. huxleyi* lipid bodies (**Eltgroth et al.**, 2005). Closely related alkenones are abundant in EhV virions and alkenone content in infected *E. huxleyi* has previously been demonstrated to be increased (**Fulton et al.**, 2014). As mentioned previously, alkenoates and alkenones are thought to occur primarily in lipid bodies (**Eltgroth et al.**, 2005). We suggest therefore, that this 36:2 wax ester/alkenoate species we observed was likely localised to lipid bodies. If true, this result, in conjunction with the TAG data mentioned above, further implicates the role of lipid bodies in viral assembly (**Fulton et al.**, 2014).

In addition to the assigned species discussed previously, many novel biomarkers could not be assigned from our database and/or corroborated by all ion fragmentation MS2 (Supplementary Figure 7.16). Many of these species demonstrate absence/presence behaviour between cultures and have potential as biomarkers of EhV infection or ploidy in *E. huxleyi*. The assigned and unassigned candidate biomarkers discovered by these unbiased analyses are tantalising. A further biomarker validation study, utilising targeted MS2/MS3 fragmentation in conjunction with high resolution mass spectrometry has the potential to assign and confirm the identities of the additional molecular species.

5.6 Conclusions

We have presented critical new data on the glycerolipids and glycosphingolipids of 1N *E. huxleyi* by the application of a targeted lipidomics approach. Comparison of these data to that of infected and uninfected 2N *E. huxleyi* has provided new insights on ploidy and viral infection. Firstly, we note the detection of trace levels of the EhV type viral glycosphingolipid (vGSL) in the uninfected 1N samples. The implications of this observation are as yet unclear but should be considered in the application of vGSL as a biomarker of EhV infection. Secondly, we find that the sialic glycosphingolipid (sGSL), a proposed marker of susceptibility to EhV infection in *E. huxleyi*, was absent from the uninfected 1N cell. This absence provides further evidence for the role of sGSL in EhV infection and may confer 1N *E. huxleyi* its documented resistance to EhV.

In addition, we have highlighted promising lipid biomarker candidates for each of the uninfected 2N, 1N and infected 2N cases by way of untargeted lipidomics. Differentially enriched biomarker candidates have been tentatively identified, such as novel glycosphingolipids, hydroxyceramides and a wax ester/alkanoate that are highly indicative of the 1N phase. Following further biomarker validation and structural studies these lipids may yield powerful lipid biomarkers for the determination of 1N *E. huxleyi*, identification of which is not feasible by classic microscopy techniques.

These findings contribute to our understanding of the critical role of lipids in *E. huxleyi*/EhV interactions. Moreover, our findings further the potential of lipid based biomarkers as indicators of the progression of infection and life cycle in *E. huxleyi*. Extrapolation from simplified cell culture models to complex environmental systems must be approached with due caution and the discussed biomarkers require further validation in culture and in situ. However, such biomarkers have the potential to yield great insight into the processes that dictate the characteristics of *E. huxleyi* blooms, translating to substantial implications for global carbon cycling and climate.

5.7 Materials and Methods

5.7.1 Culturing Procedures

The calcifying, 2N *Emiliania huxleyi* strain RCC1216 and the non-calcified, flagellated 1N *E. huxleyi* RCC1217 (isolated from RCC1216 following a partial phase change (2N to 1N), were used for this study (**Houdan et al.**, 2005). Cells were cultured in K/2 medium (**Keller et al.**, 1987) and incubated at 18°C with a 16 : 8 h, light : dark illumination cycle. Light intensity was provided at 100 μmol photons m^{-2} s^{-1} with cool white LED lights. All experiments were performed in duplicate. The virus used for this study is the lytic *E. huxleyi* virus EhV201 (**Schroeder et al.**, 2002) used at an initial multiplicity of infection (MOI) of 0.2 viral particles $cell^{-1}$. Samples of 25 mL were collected daily over 120 h post-infection by gentle vacuum-filtration onto pre-combusted GFF filters (Whatmann) and stored at -80°C until further analysis.

5.7.2 Targeted Lipid Analysis

Total lipid extracts were prepared from the cell isolates by a modified Bligh and Dyer extraction (**Bligh and Dyer**, 1959; **Popendorf et al.**, 2013), with addition of the internal standard DNP-PE(16:0/16:0) (2,4-dinitrophenyl modified PE). The prepared total lipid extracts were then subjected to targeted lipid analysis by normal phase high performance liquid chromatography tandem mass spectrometry (HPLC-MS/MS) on an Agilent 1200 HPLC coupled to a Thermo Scientific TSQ Vantage triple quadrupole MS. Chromatography and mass spectrometry conditions were as described by **Popendorf et al.** (2013).

Lipid classes were identified by retention time and characteristic MS2 fragmentation as in **Popendorf et al.** (2013); **Fulton et al.** (2014) and quantified based upon peak area within a given MS2 mass chromatogram. Quantification was achieved relative to external standard response factor calibrations. These calibrations were generated immediately before each analytical run, from a mixture of standards in a dilution series. PC, PG, PE and DNP-PE (Avanti Polar Lipids, Alabaster, AL, USA); MGDG, DGDG (Matreya LLC,

Pleasant Gap, PA, USA) and SQDG (Lipid Products, South Nutfield, UK) standards were used to generate calibrations for the glycerolipids. DGCC was quantified based on a purified extract from cultured *Thalassiosira Pseudonana* (**Popendorf et al.**, 2013). DGTS was quantified based upon the DGCC calibration subject to a scaling factor as described by **Popendorf et al.** (2011a). The internal standard mixture was run after every seven samples as a control upon instrument variability. Quantities were corrected based upon the quality control run prior to a given sample. BLL and PDPT were quantified directly from the DNP-PE internal standard, as a semi-quantitative solution in the absence of an available internal standard. Quantification of hGSL, vGSL and sGSL was based upon the response factor calibration of soy glucocerebroside extract (Avanti Polar Lipids, Alabaster, AL, USA) (**Fulton et al.**, 2014).

While uninfected 1N cells have been reported to have a smaller diameter than calcified uninfected 2N cells, this difference is thought to be due primarily to the presence of the calcified skeletons as opposed to a large difference in cytoplasmic volume (**Mausz and Pohnert**, 2015). Furthermore, uncalcified 2N cells were found to be of comparable diameter to the 1N cells (**Klaveness**, 1972; **Mausz and Pohnert**, 2015). Therefore, we assume that the biovolume of 1N and 2N cells was similar, and the measured lipid quantities were normalised to cell concentration for ease of interpretation.

All data represent the mean average of two biological replicates with error bars of one standard deviation. Statistical significance for the targeted analyses was determined by two-tailed, paired equal-variance T-test. A Bonferroni correction was applied to the significance threshold to account for dual comparisons (uninfected 2N/uninfected 1N, uninfected 2N/infected 2N), thus a p value of <0.025 was considered statistically significant. Any such significant variation is described in the results section. P values were indicated as $*p<0.025$; $**p<0.0025$ where appropriate.

5.7.3 Untargeted Lipid Analysis

Total lipid extracts from the 48 h samples were also analysed by untargeted reverse phase HPLC-MS methodology on an Agilent 1200 HPLC coupled to a Thermo Scientific Exactive

orbitrap mass spectrometer. Chromatography and mass spectrometry conditions were as described by **Hummel et al.** (2011); with the exception of a Waters XBridge C8 column (5uM packing, 150 x 2.1mm). During each sample run, the mass spectrometer continuously cycled between full positive, full negative, positive all ion fragmentation and negative all ion fragmentation modes, generating spectra with high mass resolution.

The data was processed with the Thermo Scientific Sieve software package using the component extraction algorithm for chromatographic alignment, peak detection and integration. Identified peak areas were then normalised to the DNP-PE internal standard and number of cells isolated. These unbiased analyses yield relative quantification; hence the abundance of a given molecular species is only comparable with the abundance of the same molecular species in other samples. The extracted data was filtered to remove molecular species where the deviation in abundance between replicates exceeded a factor of 10.

Peak areas were mean centred, level scaled to their means (**van den Berg et al.**, 2006) and used to build a partial least square discriminant analysis model, using the Classification Toolbox for MATLAB (**Ballabio and Consonni**, 2013). The PLS-DA models (Positive ions/Negative ions) were built using the following parameters: two components, Bayes assignment, six cross validation (CV) groups in contiguous blocks (*i.e.* Leave-One-Out validation as six samples were used). The positive model described 94% of the variance and had an error and CV error rate of 0. The negative model described 89% of the variance and had an error and CV error rate of 0. Ions from the mass spectrometry data were ranked upon their contribution (loading) toward the model score of samples of a given culture type (Supplementary Figure 7.13).

The top five ions ranked as indicative of each culture type in positive and negative ion mode were identified by matching to an extensive, accurate mass, structure query language (SQL) lipid database. The database was populated by permutations of fatty acids (chain length/degree of unsaturation) and common glycerolipids/sphingolipids. The complete LIPID MAPS (version 20130306) structural database (**Sud et al.**, 2007) and MaConDa mass spectrometry contaminants database (**Weber et al.**, 2012) were also included. The chemical formulae of database entries were then used to calculate accurate mass m/z values

based upon a list of common molecular ion adducts in ESI-MS (**Huang et al.**, 1999). Database hits were within 2.5 ppm of the measured m/z. Supporting MS2 fragmentation and diagnostic retention time information was recorded for the assignments presented in the Results and is included in Supplementary Figures 7.14 and 7.15.

5.8 Acknowledgements

The authors would like to thank J. Tagliafere and J. Ossolinski for assistance in the laboratory and J. Collins, J. Fulton and B. Edwards for helpful discussions about data analysis. This work was funded by the University of Southampton - Vice Chancellors Scholarship Award; Graduate School of the National Oceanography Centre WHOI Exchange Award (J.E.H.) and the Gordon and Betty Moore Foundation through Grant GBMF3301 (B.A.S.V.M. and A.V.).

5.9 References

Alvarez, H. and Steinbüchel, A. (2002), Triacylglycerols in prokaryotic microorganisms, *Applied Microbiology and Biotechnology*

Avrani, S., Wurtzel, O., Sharon, I., Sorek, R., and Lindell, D. (2011), Genomic island variability facilitates Prochlorococcus-virus coexistence, *Nature*

Ballabio, D. and Consonni, V. (2013), Classification tools in chemistry. Part 1: linear models. PLS-DA, *Analytical Methods*

Bell, M. V. and Pond, D. (1996), Lipid composition during growth of motile and coccolith forms of Emiliania huxleyi, *Phytochemistry*

Bligh, E. G. and Dyer, W. J. (1959), A Rapid Method of Total Lipid Extraction and Purification, *Canadian Journal of Biochemistry and Physiology*

Borchard, C. and Engel, A. (2012), Organic matter exudation by Emiliania huxleyi under simulated future ocean conditions, *Biogeosciences*

Brandsma, J., Hopmans, E. C., Philippart, C. J. M., Veldhuis, M. J. W., Schouten, S., and Sinninghe Damsté, J. S. (2012), Low temporal variation in the intact polar lipid composition of North Sea coastal marine water reveals limited chemotaxonomic value, *Biogeosciences*

Bratbak, G., Egge, J. K., and Heldal, M. (1993), Viral mortality of the marine alga Emiliania huxleyi (Haptophyceae) and termination of algal blooms, *Marine Ecology Progress Series*

Brussaard, C., Gast, G., van Duyl, F., and Riegman, R. (1996), Impact of phytoplankton bloom magnitude on a pelagic microbial food web, *Marine Ecology Progress Series*

Carini, P., Van Mooy, B. A. S., Thrash, J. C., White, A., Zhao, Y., Campbell, E. O., et al. (2015), SAR11 lipid renovation in response to phosphate starvation, *Proceedings of the National Academy of Sciences*

Eltgroth, M. L., Watwood, R. L., and Wolfe, G. V. (2005), Production and Cellular Localization of Neutral Long-Chain Lipids in the Haptophyte Algae Isochrysis Galbana and Emiliania Huxleyi., *Journal of Phycology*

Frada, M., Probert, I., Allen, M. J., Wilson, W. H., and de Vargas, C. (2008), The "Cheshire Cat" escape strategy of the coccolithophore Emiliania huxleyi in response to viral infection, *Proceedings of the National Academy of Sciences*

Frada, M. J., Bidle, K. D., Probert, I., and de Vargas, C. (2012), In situ survey of life cycle phases of the coccolithophore Emiliania huxleyi (Haptophyta), *Environmental Microbiology*

Fuhrman, J. A. (1999), Marine viruses and their biogeochemical and ecological effects., *Nature*

Fulton, J. M., Fredricks, H. F., Bidle, K. D., Vardi, A., Kendrick, B. J., DiTullio, G. R., et al. (2014), Novel molecular determinants of viral susceptibility and resistance in the lipidome of Emiliania huxleyi, *Environmental Microbiology*

Green, J. C., Course, P. A., and Tarran, G. A. (1996), The life-cycle of Emiliania huxleyi: A brief review and a study of relative ploidy levels analysed by flow cytometry, *Journal of Marine Systems*

Gundersen, K., Heldal, M., Norland, S., Purdie, D. A., and Knap, A. H. (2002), Elemental C, N, and P cell content of individual bacteria collected at the Bermuda Atlantic Time-series Study (BATS) site, *Limnology and Oceanography*

Herker, E. and Ott, M. (2011), Unique ties between hepatitis C virus replication and intracellular lipids, *Trends in Endocrinology & Metabolism*

Holligan, P. M., Viollier, M., Harbour, D. S., Camus, P., and Champagne-Philippe, M. (1983), Satellite and ship studies of coccolithophore production along a continental shelf edge, *Nature*

Houdan, A., Billard, C., Marie, D., Not, F., Sáez, A. G., Young, J. R., et al. (2004), Holococcolithophore - heterococcolithophore (Haptophyta) life cycles: Flow cytometric analysis of relative ploidy levels, *Systematics and Biodiversity*

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

Huang, N., Siegel, M. M., Kruppa, G. H., and Laukien, F. H. (1999), Automation of a Fourier transform ion cyclotron resonance mass spectrometer for acquisition, analysis, and e-mailing of high-resolution exact-mass electrospray ionization mass spectral data, *Journal of the American Society for Mass Spectrometry*

Hummel, J., Segu, S., Li, Y., Irgang, S., Jueppner, J., and Giavalisco, P. (2011), Ultra Performance Liquid Chromatography and High Resolution Mass Spectrometry for the Analysis of Plant Lipids, *Frontiers in Plant Science*

Jover, L. F., Effler, T. C., Buchan, A., Wilhelm, S. W., and Weitz, J. S. (2014), The elemental composition of virus particles: implications for marine biogeochemical cycles, *Nat Rev Micro*

Keller, M. D., Selvin, R. C., Claus, W., and Guillard, R. R. L. (1987), Media for the Culture of Oceanic Ultraphytoplankton, *Journal of Phycology*

Klaveness, D. (1972), Coccolithus huxleyi (Lohm.) Kamptn II. The flagellate cell, aberrant cell types, vegetative propagation and life cycles, *British Phycological Journal*

Lehahn, Y., Koren, I., Schatz, D., Frada, M., Sheyn, U., Boss, E., et al. (2014), Decoupling Physical from Biological Processes to Assess the Impact of Viruses on a Mesoscale Algal Bloom, *Current Biology*

Liefhebber, J. M. P., Hague, C. V., Zhang, Q., Wakelam, M. J. O., and McLauchlan, J. (2014), Modulation of Triglyceride and Cholesterol Ester Synthesis Impairs Assembly of Infectious Hepatitis C Virus, *Journal of Biological Chemistry*

Mackinder, L. C. M., Worthy, C. A., Biggi, G., Hall, M., Ryan, K. P., Varsani, A., et al. (2009), A unicellular algal virus, Emiliania huxleyi virus 86, exploits an animal-like infection strategy, *Journal of General Virology*

Martin, P., Van Mooy, B. A. S., Heithoff, A., and Dyhrman, S. T. (2011), Phosphorus supply drives rapid turnover of membrane phospholipids in the diatom Thalassiosira pseudonana, *The ISME Journal*

Mausz, M. A. and Pohnert, G. (2015), Phenotypic diversity of diploid and haploid Emiliania huxleyi cells and of cells in different growth phases revealed by comparative metabolomics, *Journal of Plant Physiology*

Monier, A., Pagarete, A., de Vargas, C., Allen, M. J., Read, B., Claverie, J.-M., et al. (2009), Horizontal gene transfer of an entire metabolic pathway between a eukaryotic alga and its DNA virus, *Genome Research*

Paasche, E. (2001), A review of the coccolithophorid Emiliania huxleyi (Prymnesiophyceae), with particular reference to growth, coccolith formation, and calcification-photosynthesis interactions, *Phycologia*

Petroutsos, D., Amiar, S., Abida, H., Dolch, L.-J., Bastien, O., Rébeillé, F., et al. (2014), Evolution of galactoglycerolipid biosynthetic pathways From cyanobacteria to primary plastids and from primary to secondary plastids, *Progress in Lipid Research*

Popendorf, Tanaka, T., Pujo-Pay, M., Lagaria, A., Courties, C., Conan, P., et al. (2011a), Gradients in intact polar diacylglycerolipids across the Mediterranean Sea are related to phosphate availability, *Biogeosciences*

Popendorf, K. J., Fredricks, H. F., and Van Mooy, B. A. (2013), Molecular Ion-Independent Quantification of Polar Glycerolipid Classes in Marine Plankton Using Triple Quadrupole MS, *Lipids*

Popendorf, K. J., Lomas, M. W., and Van Mooy, B. A. S. (2011b), Microbial sources of intact polar diacylglycerolipids in the Western North Atlantic Ocean, *Organic Geochemistry*

Prahl, F. G. and Wakeham, S. G. (1987), Calibration of unsaturation patterns in long-chain ketone compositions for palaeotemperature assessment, *Nature*

Ray, J. L., Haramaty, L., Thyrhaug, R., Fredricks, H. F., Van Mooy, B. A. S., Larsen, A., et al. (2014), Virus infection of Haptolina ericina and Phaeocystis pouchetii implicates evolutionary conservation of programmed cell death induction in marine haptophyte-virus interactions, *Journal of Plankton Research*

Rokitta, S. D., de Nooijer, L. J., Trimborn, S., de Vargas, C., Rost, B., and John, U. (2011), Transcriptome Analyses Reveal Differential Gene Expression Patterns Between the Life-Cycle Stages of Emiliania Huxleyi (Haptophyta) and Reflect Specialization to Different Ecological Niches., *Journal of Phycology*

Rose, S. L., Fulton, J. M., Brown, C. M., Natale, F., Van Mooy, B. A. S., and Bidle, K. D. (2014), Isolation and characterization of lipid rafts in Emiliania huxleyi: a role for membrane microdomains in hostvirus interactions, *Environmental Microbiology*

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, *The Plant Cell*

Sakurai, I., Shen, J.-R., Leng, J., Ohashi, S., Kobayashi, M., and Wada, H. (2006), Lipids in Oxygen-Evolving Photosystem II Complexes of Cyanobacteria and Higher Plants, *Journal of Biochemistry*

Schatz, D., Shemi, A., Rosenwasser, S., Sabanay, H., Wolf, S. G., Ben-Dor, S., et al. (2014), Hijacking of an autophagy-like process is critical for the life cycle of a DNA virus infecting oceanic algal blooms, *New Phytologist*

Schroeder, D. C., Oke, J., Malin, G., and Wilson, W. H. (2002), Coccolithovirus (Phycodnaviridae): Characterisation of a new large dsDNA algal virus that infects Emiliana huxleyi, *Archives of Virology*

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

Stray, S. J., Cummings, R. D., and Air, G. M. (2000), Influenza virus infection of desialylated cells, *Glycobiology*

Sud, M., Fahy, E., Cotter, D., Brown, A., Dennis, E. A., Glass, C. K., et al. (2007), LMSD: LIPID MAPS structure database, *Nucleic Acids Research*

Suttle, C. A. (2007), Marine viruses - major players in the global ecosystem, *Nat. Rev. Micro.*

Tyrrell, T. and Merico, A. (2004), Emiliania huxleyi: bloom observations and the conditions that induce them, in H. R. Thierstein and J. R. Young, eds., Coccolithophores SE - 4 (Springer Berlin Heidelberg)

van den Berg, R. A., Hoefsloot, H. C. J., Westerhuis, J. A., Smilde, A. K., and van der Werf, M. J. (2006), Centering, scaling, and transformations: improving the biological information content of metabolomics data., *BMC Genomics*

Van Etten, J. L., Graves, M. V., Müller, D. G., Boland, W., and Delaroque, N. (2002), Phycodnaviridae large DNA algal viruses, *Archives of Virology*

Van Mooy, B. A. S., Fredricks, H. F., Pedler, B. E., Dyhrman, S. T., Karl, D. M., Koblížek, M., et al. (2009), Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity, *Nature*

Vardi, A., Haramaty, L., Van Mooy, B. A. S., Fredricks, H. F., Kimmance, S. A., Larsen, A., et al. (2012), Hostvirus dynamics and subcellular controls of cell fate in a natural coccolithophore population, *Proceedings of the National Academy of Sciences*

Vardi, A., Van Mooy, B. A. S., Fredricks, H. F., Popendorf, K. J., Ossolinski, J. E., Haramaty, L., et al. (2009), Viral glycosphingolipids induce lytic infection and cell death in marine phytoplankton., *Science*

Volkman, J. K., Eglinton, G., Corner, E. D. S., and Sargent, J. R. (1980), Novel unsaturated straight-chain C37-C39 methyl and ethyl ketones in marine sediments and a coccolithophore Emiliania huxleyi, *Phys. Chem. Earth*

von Dassow, P., Ogata, H., Probert, I., Wincker, P., Da Silva, C., Audic, S., et al. (2009), Transcriptome analysis of functional differentiation between haploid and diploid cells of Emiliania huxleyi, a globally significant photosynthetic calcifying cell, *Genome Biology*

Weber, R. J. M., Li, E., Bruty, J., He, S., and Viant, M. R. (2012), MaConDa: a publicly accessible mass spectrometry contaminants database, *Bioinformatics*

Westbroek, P., Brown, C. W., van Bleijswijk, J., Brownlee, C., Brummer, G. J., Conte, M., et al. (1993), A model system approach to biological climate forcing. The example of Emiliania huxleyi, *Global and Planetary Change*

Wilson, W. H. (2005), Complete Genome Sequence and Lytic Phase Transcription Profile of a Coccolithovirus, *Science*

Wilson, W. H., Tarran, G. A., Schroeder, D., Cox, M., Oke, J., and Malin, G. (2002), Isolation of viruses responsible for the demise of an Emiliania huxleyi bloom in the English Channel, *Journal of the Marine Biological Association of the UK*

Chapter 6

Summary and Concluding Remarks

6.1 Lipid Remodelling by the Diatom *Thalassiosira pseudonana* under Phosphorus Stress – New Insights into Phospholipid Substitution

Chapter 2 presents findings consistent with previous reports on the response of the polar lipids of *T. pseudonana* to P stress, at the level of lipid class. Specifically, the phospholipids PC, PG and PE were observed to decrease in quantity per cell with increasing incubation time, while the betaine lipid substitute DGCC increased.

Beyond the previous knowledge, it was observed that P stress resulted in a net cessation of glycerophospholipid synthesis, with a minor degree of breakdown. The majority of the original glycerophospholipid remained intact and was diluted through culture growth by cell division. As a result, most of the P bound in the glycerophospholipid headgroups remained as such and was unavailable for other cellular processes as hypothesised elsewhere (**Martin et al.**, 2011). The broken down glycerophospholipid was equivalent to only a small proportion of the biosynthesised DGCC, suggesting phospholipid breakdown cannot form a major source of recycled diglyceride for DGCC biosynthesis.

Further insight was gained by investigation of the individual molecular species within each lipid headgroup class. During P-replete growth, significant variability was observed within the per cell quantities of PC molecular species, with the progression of time. The complexity and degree of these temporal changes highlight the plasticity of the lipidome, often overlooked in biomarker studies. During P-stress, the quantity per cell of individual molecular species within the PC, PG and PE classes decreased in line with the total class data. The magnitude of this decrease varied between individual PC species, notably the highly polyunsaturated PC(40:10) and PC(42:11) showed the smallest decreases. In contrast, several minor PE species, some of which appear to bear very long chain (C >22:6) polyunsaturated fatty acids, increased under P stress. Subject to further characterisation, these may have utility as molecular biomarkers.

Variation in the relative abundance of molecular species of the glyceroglycolipids MGDG and DGDG were observed. Therefore, despite previous observations on the stability in the total MGDG and DGDG quantity per cell to P stress, we find that there are implications upon the glyceroglycolipids.

Finally, the similarity of the lipid molecular species composition between lipid classes was assessed. The compositions of PC and DGCC were found to be similar, providing further evidence for the role of DGCC as a substitute for PC. The molecular composition of DAG, a common precursor to many of the polar lipid classes investigated, was distinct from PC and DGCC, being strongly correlated instead to the glyceroglycolipids MGDG and DGDG. Thus, the observed DAG appears to be utilised primarily by monogalactosyl-diacylglycerol synthase as a precursor to MGDG and DGDG.

These observations lead to the formation of the hypothesis that there are two (or more) separate pools of DAG. Firstly, the larger and/or slower turned over DAG pool indicative of MGDG/DGDG synthesis, observed by characterisation of the total lipid extract. Secondly, the smaller and/or more rapidly turned over DAG pool indicative of PC/PE synthesis that is conspicuously not observed. This hypothesis was investigated in the following chapter.

6.2 Untargeted Lipidomic Characterisation of the Marine Diatom *Thalassiosira Pseudonana* Subject to Phosphorus Stress - Increase in Diglycosylceramides under Low Phosphorus Conditions

Chapter 3 presents the untargeted lipidomic screening of the model marine diatom *T. pseudonana*, under nutrient replete and phosphorus stressed growth conditions. This highlighted a number of lipid species, such as several DGCC chemotypes, that increased strongly under P stress in line with previous observations.

A triacylglycerol species, reported elsewhere to be present only at trace levels under P replete, N starved and Si starved conditions was present at increased levels subject to P stress and suggests potentially interesting triacylglycerol dynamics.

Furthermore, a group of diglycosylceramides, not previously detected in *T. pseudonana*, increased in cellular abundance under P stress. These species may prove useful as biomarkers for P stress in diatoms. Such biomarkers give insight into the P stress experienced by the organisms sampled that is complimentary to the routine measurement of dissolved and particulate P concentrations. These biomarkers may have potential applications in the elucidation of P related biogeochemical processes, pending further validation.

Targeted MS2 analyses were used to characterise the fatty acyl compositions of each of the major diglyceride lipid species. This provides new insight into the *T. pseudonana* lipidome filling a conspicuous gap in an otherwise well characterised system.

Finally, a preliminary qualitative study of the subcellular partitioning of precursor DAG lipids supported the hypothesis that DAG lipids were comprised of two subcellular pools: one a putative MGDG/DGDG precursor pool localised to the chloroplast, the other a putative PC/PE precursor localised elsewhere.

6.3 Lipid Remodelling by the Diatom *Thalassiosira pseudonana* under Nitrogen Stress – Diglyceride Lipid Dynamics and Relation to Triglyeride Production

Chapter 4 presents results from a targeted lipidomic analysis of *T. pseudonana* under N limited stationary phase growth compared with N replete exponential growth. We find that under N limitation, total PC increases that were primarily attributable to highly unsaturated PC species bearing 20:5 and 22:6 fatty acids, resulted in a relative enrichment of these species. The pools of PE and PG lipids remained largely unchanged by N limitation.

An increase in total SQDG subject to N limitation was almost exclusively attributed to a large increase in SQDG(14:0/14:0), the predominant SQDG chemotype. Within the MGDG pool, we observed a relative enrichment of MGDG species containing 16:0 fatty acids and relative depletion of 16:1 fatty acid containing species.

Finally, total DAG increased 3 fold under N limitation. This increase was comprised of increases in DAG species bearing 16:0, 16:1 and 14:0 fatty acids. Given the biosynthetic relationship and literature reports on the production and chemical composition of TAG under N limitation, this increase in DAG appears to form part of the cells adaption to N limitation that ultimately leads to the accumulation of TAG. These findings further our understanding of the diatom lipidome and its response to nutrient limitation and adaption to low N environments.

Furthermore, this understanding of the TAG metabolic pathway under N limitation could inform genetic modification or other means to increase lipid yield, improving the efficiency and economics of biofuel production.

6.4 Targeted and Untargeted Lipidomics of *Emiliania huxleyi* Viral Infection and Life Cycle Phases Highlights Molecular Biomarkers of Infection, Susceptibility, and Ploidy

Chapter 5 presents new insights into the glycerolipids and glycosphingolipids of 1N *E. huxleyi* by the application of a targeted lipidomics approach. By comparison of these data to that of infected and uninfected 2N *E. huxleyi* we have made a number of interesting observations: Firstly, we note the detection of trace levels of the EhV type viral glycosphingolipid (vGSL) in the uninfected 1N samples. The implications of this observation are as yet unclear but should be considered in the application of vGSL as a biomarker of EhV infection. Secondly, we find that the sialic glycosphingolipid (sGSL), a proposed marker of susceptibility to EhV infection in *E. huxleyi*, was absent from the uninfected 1N cell. This absence provides further evidence for the role of sGSL in EhV infection and may confer 1N *E. huxleyi* its documented resistance to EhV.

In addition, we have highlighted promising lipid biomarker candidates for each of the uninfected 2N, 1N and infected 2N cases by way of untargeted lipidomics. Differentially enriched biomarker candidates have been tentatively identified, such as novel glycosphingolipids, hydroxyceramides and a wax ester/alkanoate that are highly indicative of the 1N phase. Following further biomarker validation and structural studies these lipids may yield powerful lipid biomarkers for the determination of 1N *E. huxleyi*, identification of which is not feasible by classic microscopy techniques.

These findings contribute to our understanding of the critical role of lipids in *E. huxleyi*/EhV interactions. Moreover, our findings further the potential of lipid based biomarkers as indicators of the progression of infection and life cycle in *E. huxleyi*. Such biomarkers will yield insight into the processes that dictate the characteristics of *E. huxleyi* blooms, translating to substantial implications upon global carbon cycling and climate.

6.5 The Power of Mass Spectrometric Lipidomics in Environmental Science

The application of mass spectrometric lipidomics is of considerable benefit to the field of environmental science. An overview of the study of lipids in the marine environment is presented in Chapter 1. Mass spectrometric lipidomics (in this case constrained to electrospray ionisation methods with and without chromatographic separation) allow unique insights into cellular composition and metabolism. The investigations presented within this thesis provide many examples:

Firstly, lipidomic techniques may yield targeted, highly sensitive and quantitative analyses of known lipids or lipid classes. Few analytical techniques can match the excellent sensitivity and specicifity of these methods (**Blanksby and Mitchell**, 2010). Targeted lipidomic analyses were applied in Chapter 2 and Chapter 4 to investigate the response of the major glycerolipid classes to P and N stress respectively, in a model marine diatom. The major shifts detected in glycerolipid abundance at the total class and individual chemotype levels, bear substantial biogeochemical and ecological implications. These implications are rooted in a firm biological basis and metabolic context and demonstrate how the lipidomic investigation of a subset of lipids may yield insight into environmental processes.

Secondly, untargeted lipidomic screening of phytoplankton cultures allows for the detection of hundreds to thousands of individual lipid species in an unbiased manner. This demonstrates the breadth of analytes that can be detected in complex biological matrices by these methods (**Blanksby and Mitchell**, 2010). Furthermore, a given analyte may be structurally annotated based upon its fragment ions in MS2 mode. Each analyte may then be quantified in relative terms between samples giving a very broad insight into an organism's lipidic response to its conditions. This approach is particularly useful for the discovery of biomarkers that are diagnostic of organisms or processes in the environment. In Chapter 3 and Chapter 5 untargeted lipidomic screening was used to highlight novel candidate biomarkers of P stress in a model marine diatom and viral infection, susceptibility and ploidy in a model haptophyte, respectively.

In summary, the research presented within each chapter of this thesis provides clear evidence in support of the analytical power of mass spectrometric lipidomics in the study of phytoplankton lipids both in culture and in the environment. This power arises from its one of a kind combination of excellent sensitivity, ability to annotate molecular ion structure through fragmentation, and capability to deal with highly complex biological matrices.

6.6 Synthesis

When the presented studies are considered in conjunction, a number of broad themes become apparent. Firstly, variation in the composition and/or abundance of the plastid glycolipids MGDG was observed in each of the studies, highlighting the plasticity of MGDG subject to stress in general.

Like MGDG, TAG lipids were also increased in abundance in each of the presented studies. This observation is not surprising given the well documented accumulation of TAG lipids under a wide variety of stress conditions (**Cagliari et al.**, 2011).

Several novel candidate biomarkers have been highlighted during the course of this research. Taken with the large theoretical diversity of lipid species (**van Meer**, 2005) these results suggest a wealth of biomarkers for important processes may remain undiscovered. This reinforces the need for untargeted screening experiments such as those presented in Chapter 3 and Chapter 5.

Broader still, these findings highlight the high plasticity of the lipidome to environmental conditions and the importance of accounting for this variability accordingly when measuring biomarkers in situ. In its entirety, the findings presented within this thesis contribute strongly to our understanding of the response of the phytoplankton lipidome to some of the major environmental stressors encountered in the oceans. These results bear numerous implications upon phytoplankton biology, biogeochemistry and the production of biodiesels.

6.7 Future Research

The research presented within this thesis has yielded several outcomes with the potential to be applied for the benefit of the environmental sciences, following further translational research and development.

The first research avenue in direct continuation would be the chemical characterisation and validation of the candidate biomarkers highlighted, with a view to their application in situ. This would require targeted fragmentation and quantitative analyses by mass spectrometry in order to verify the identity and the relationship of each lipid species with the desired parameter. If evidence in support of specificity of the biomarker in response to a given process can be attained the biomarker may be applicable to extant, live biomass (**Volkman**, 2006). Furthermore, if the biomarker or a specific derivative is found to be stable to degredation processes during and after sedimentation, then it may be applicable as a palaeoceanographic proxy (**Rosell-Melé and McClymont**, 2007). The requirements for biomarkers are discussed in detail in Chapter 1.

Regarding the diglycosylceramide candidate biomarkers for P stress in *T. pseudonana*, presented in Chapter 3 - these may be diagnostic of P stress to diatoms or other phytoplankton in the environment. In order to validate the diglycosylceramides as a proxy for P stress, the relationship must be demonstrated by the investigation of laboratory cultures to be consistent in other diatoms and ideally other phytoplankton taxa. Finally, the response must be proven to be specific to P stress and not elicited under other conditions. If these assumptions are proven correct, it may then be possible to generate and index based upon these measures and calibrate it directly to P concentrations in the water column, akin to lipid sea surface temperature proxies (**Brassell et al.**, 1986; **Schouten et al.**, 2002).

Regarding the glycosphingolipid candidate biomarkers for haploid *E. huxleyi*, presented in Chapter 5 - these appear to be indicative of the haploid cells only. Thus, these glycosphingolipid species represent, pending further validation and characterisation, candidate biomarkers for 1N *E. huxleyi*. 1N *E. huxleyi* are not easily identifiable in field samples by classic microscopy techniques (**von Dassow et al.**, 2009) and lipid biomarkers have par-

ticular utility in bulk environmental samples, such as filtered particulate organic material collected from the water column or sinking particulate trap matter, where other methods may not be applicable.

As in the case of the diglycosylceramides above, the haploid glycosphingolipids must be validated via translational research. This must demonstrate if the response is conserved between related haptophyte strains and species by the examination of laboratory cultures. Furthermore, culture investigations are required to demonstrate the specificity of the analyte to haploid haptophytes, as opposed to other phytoplankton present in the environment. Once these conditions are satisfied the biomarker may be applied with due caution to the environmental study of haploid *E. huxleyi*. This outcome is of significance as haploid *E. huxleyi* are well known to play a key role in the dynamics of viral bloom termination (**Frada et al.**, 2008, 2012), but this has not been well studied in natural blooms.

Beyond the development and application of the discovered biomarkers - preliminary investigation of subcellular fractions of glycerolipids yielded promising results (Chapter 3). Further investigation into the compartmentalisation of lipids in eukaryotic phytoplankton and the response of the plastid lipidome to P stress has the potential to reveal new layers to our understanding of the phytoplankton adaptation to P stress.

More broadly, one major obstacle and one conspicuous gap in our knowledge are apparent from this research. As eluded to in the introduction, the availability of authentic lipid standards for absolute quantification remains a limitation for lipid research, especially in the field of phytoplankton/plant lipids.

Finally, the biosynthesis of the environmentally important and abundant betaine lipid DGCC remains unknown and its elucidation, perhaps by way of isotopic labelling or transcriptomic/proteomic analysis, would constitute a major advance.

6.8 References

Blanksby, S. J. and Mitchell, T. W. (2010), Advances in Mass Spectrometry for Lipidomics, *Annual Review of Analytical Chemistry*

Brassell, S. C., Eglinton, G., Marlowe, I. T., Pflaumann, U., and Sarnthein, M. (1986), Molecular stratigraphy: a new tool for climatic assessment, *Nature*

Cagliari, A., Margis, R., dos Santos Maraschin, F., Turchetto-Zolet, A. C., Loss, G., and Margis-Pinheiro, M. (2011), Biosynthesis of Triacylglycerols (TAGs) in plants and algae, *International Journal of Plant Biology*

Frada, M., Probert, I., Allen, M. J., Wilson, W. H., and de Vargas, C. (2008), The "Cheshire Cat" escape strategy of the coccolithophore Emiliania huxleyi in response to viral infection, *Proceedings of the National Academy of Sciences*

Frada, M. J., Bidle, K. D., Probert, I., and de Vargas, C. (2012), In situ survey of life cycle phases of the coccolithophore Emiliania huxleyi (Haptophyta), *Environmental Microbiology*

Martin, P., Van Mooy, B. A. S., Heithoff, A., and Dyhrman, S. T. (2011), Phosphorus supply drives rapid turnover of membrane phospholipids in the diatom Thalassiosira pseudonana, *The ISME Journal*

Rosell-Melé, A. and McClymont, E. L. (2007), Chapter Eleven Biomarkers as Paleoceanographic Proxies, in Developments in Marine Geology, volume 1 (Elsevier)

Schouten, S., Hopmans, E. C., Schefuß, E., and Sinninghe Damsté, J. S. (2002), Distributional variations in marine crenarchaeotal membrane lipids: a new tool for reconstructing ancient sea water temperatures?, *Earth Planet Sc. Lett.*

van Meer, G. (2005), Cellular lipidomics, *EMBO J*

Volkman, J. (2006), Lipid Markers for Marine Organic Matter (Springer Berlin / Heidelberg)

von Dassow, P., Ogata, H., Probert, I., Wincker, P., Da Silva, C., Audic, S., et al. (2009), Transcriptome analysis of functional differentiation between haploid and diploid cells of Emiliania huxleyi, a globally significant photosynthetic calcifying cell, *Genome Biology*

Chapter 7

Appendices

7.1 Appendix 1 - Supplementary Information for Chapter 2

Lipid Remodelling by the Diatom *Thalassiosira pseudonana* under Phosphorus Stress – New Insights into Phospholipid Substitution

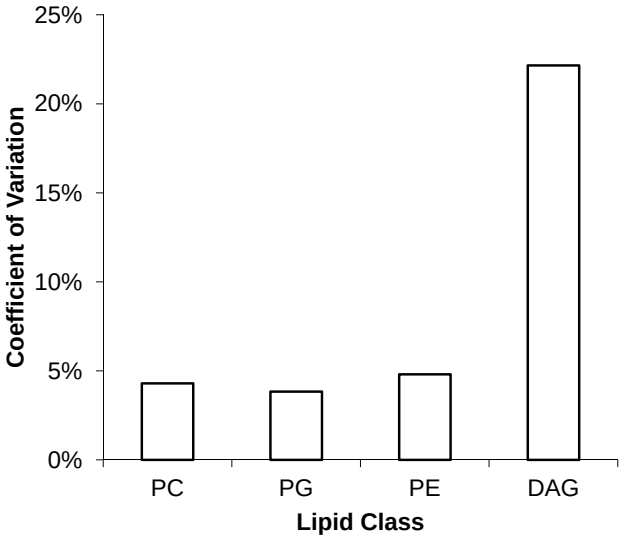


Figure 7.1: Assessment of analytical variability of targeted lipidomic analysis by direct infusion ESI-MS/MS. Quality control mixtures of standards were run daily over the six days in which the analyses were carried out. The mixtures were comprised of 5 nmol of each of the 'standards' used for quantification as outlined in Section 2.7.5. The mixtures also contained 5 nmol of the following 'analytes': PC(18:1/18:1); PG(14:0/14:0); PE18:1/18:1); and DAG(16:0/18:1). The 'analyte' standards were acquired from Avanti Polar Lipids (Alabaster, U.S.A.) and manipulated as described in Chapter 2. The mixtures were analysed as per the samples and each quantity of the analyte determined based upon its relative intensity to the standard peak of the same lipid headgroup class. The coefficient of variation was then determined based upon the range of calculated quantities, for a given analyte, throughout the duration of the analysis.

All subsequent data represent the mean of n=3 with error bars of 1 standard deviation. Lipid data shown throughout the time course in phosphorus replete (P+) and phosphorus stressed (P-) cultures.

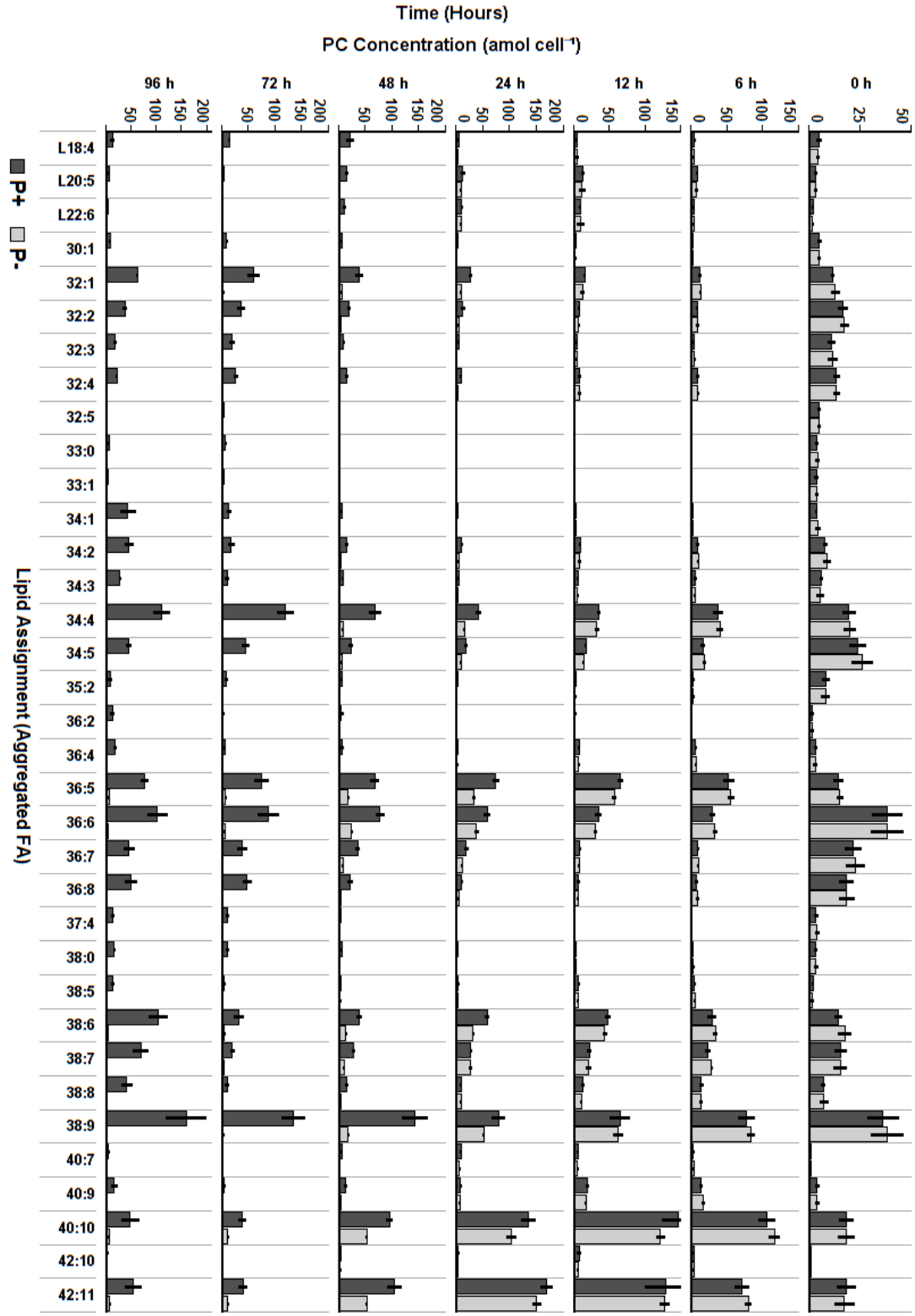


Figure 7.2: Full lipid dataset for Chapter 2, absolute quantification of glycerophosphatidylcholine (PC) species.

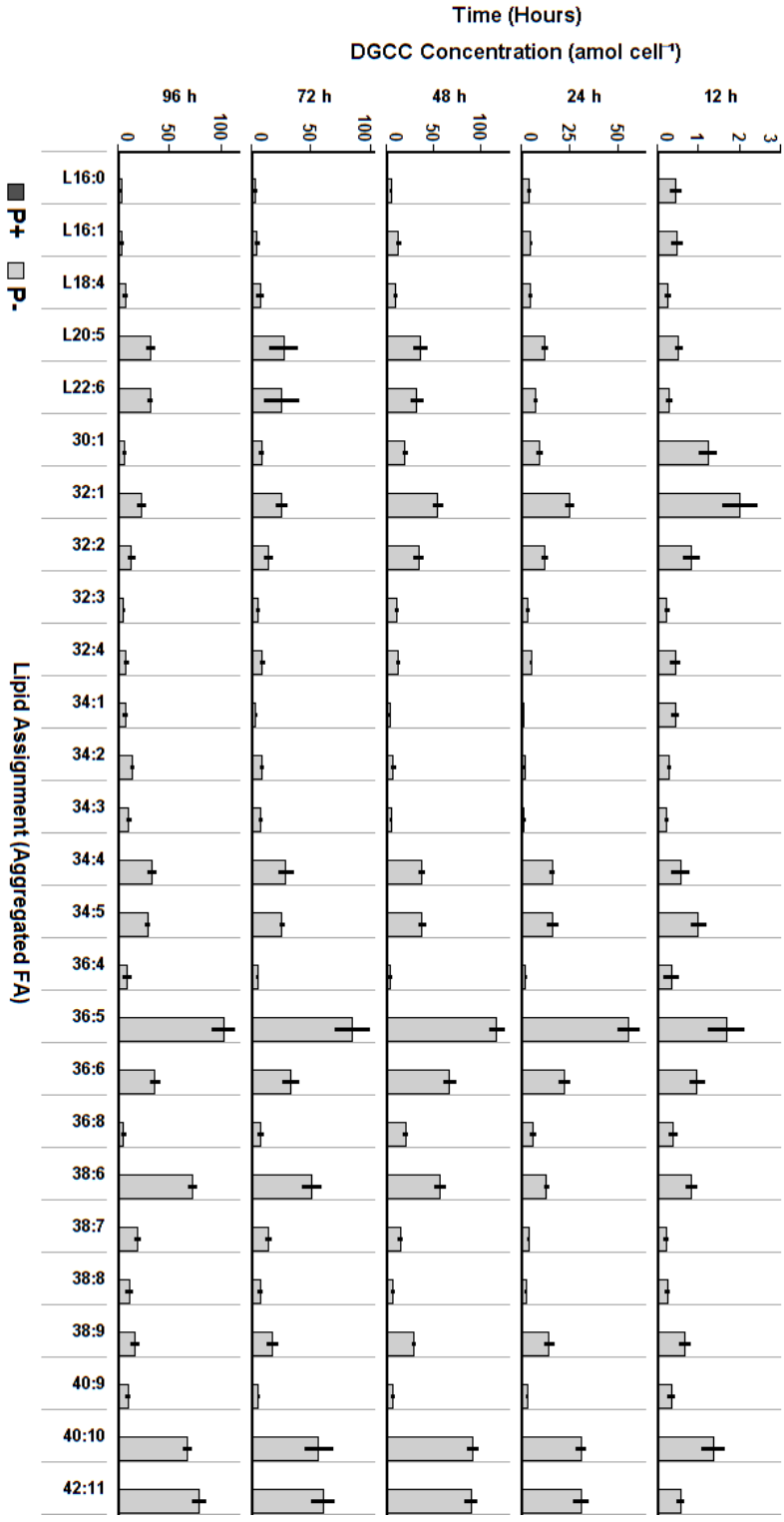


Figure 7.3: Full lipid dataset for Chapter 2, absolute quantification of diacylglyceryl-3-O-carboxy-(hydroxymethyl)-choline (DGCC) species.

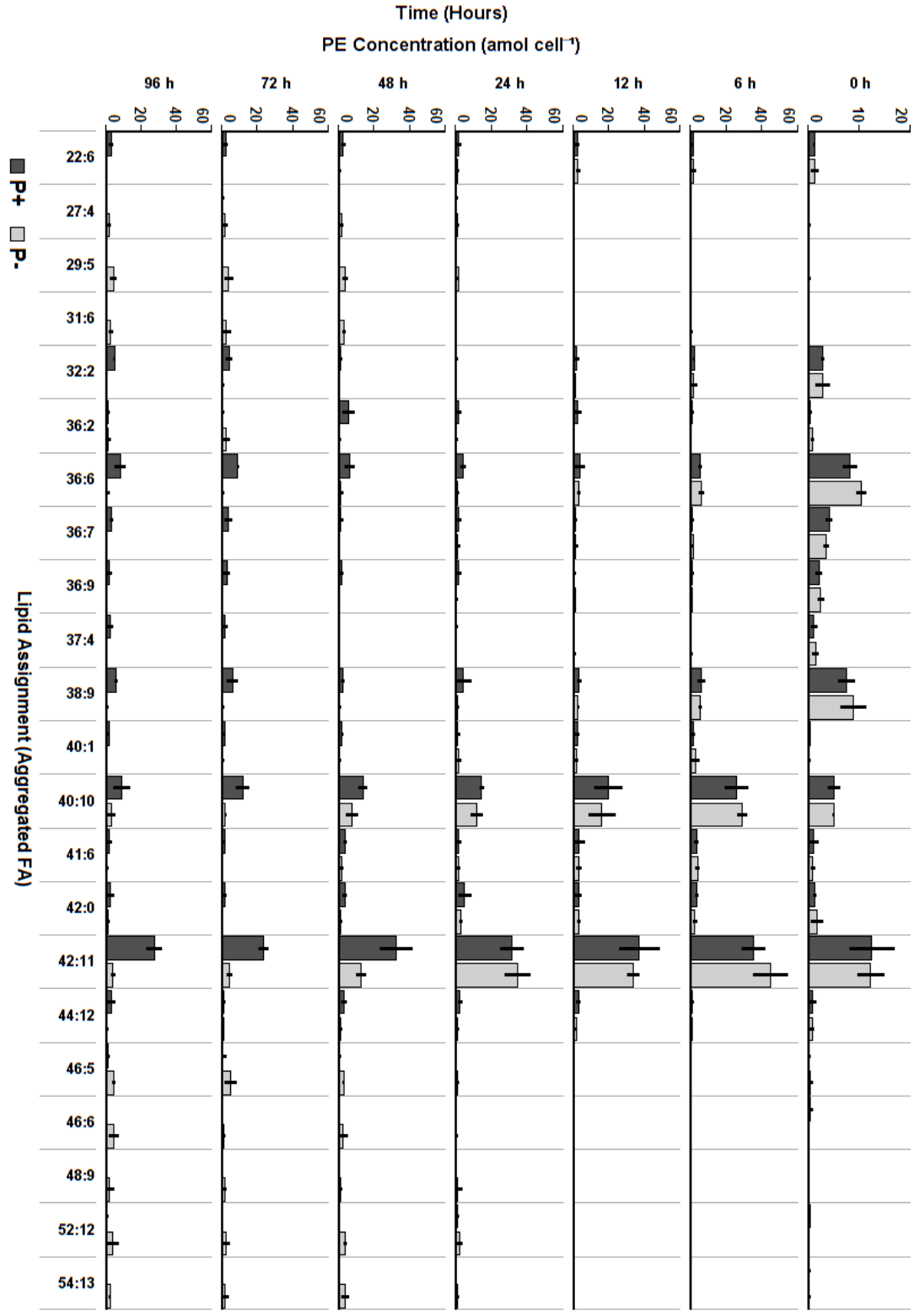


Figure 7.4: Full lipid dataset for Chapter 2, absolute quantification of glycerophosphatidylethanolamine (PE) species.

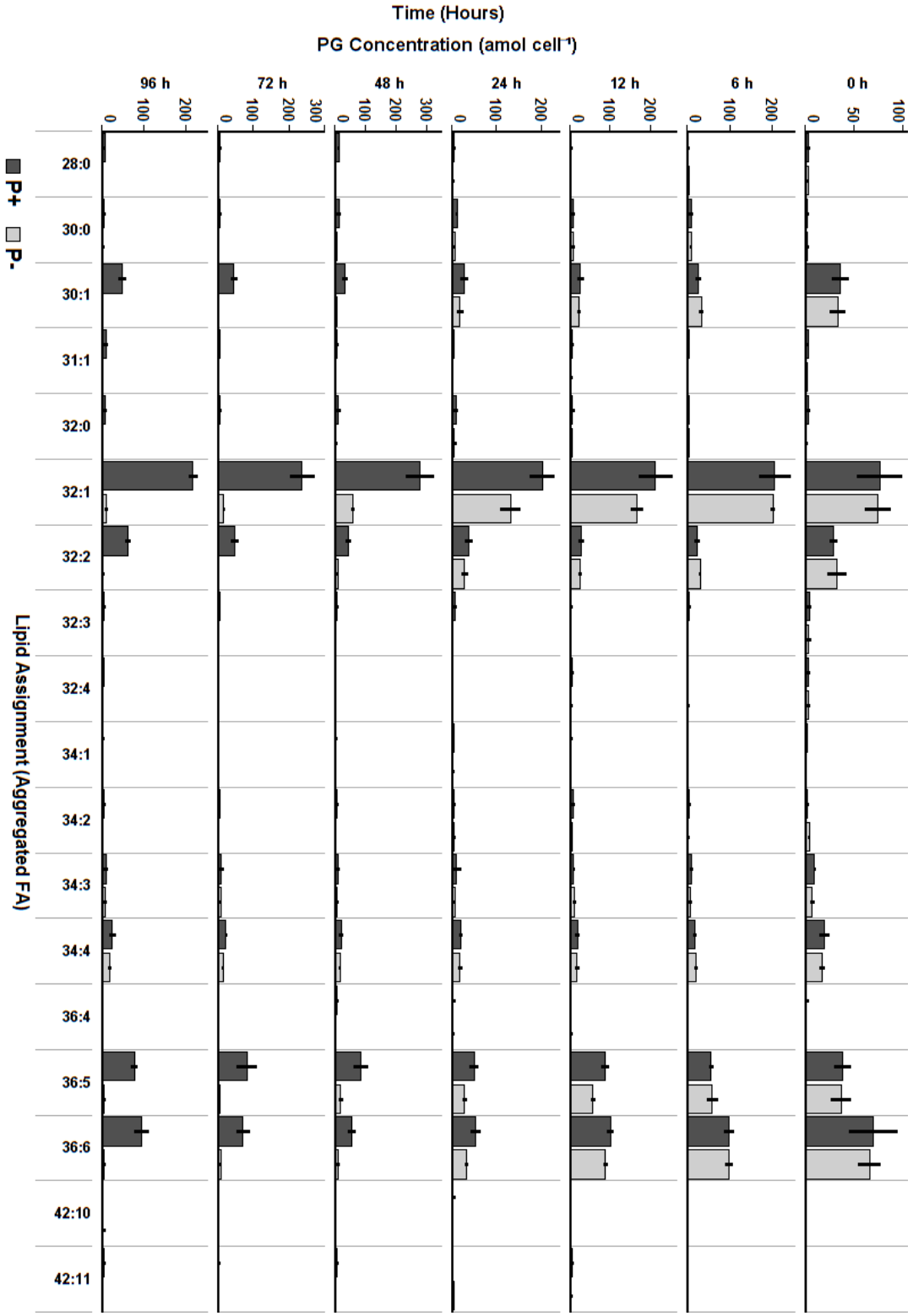


Figure 7.5: Full lipid dataset for Chapter 2, absolute quantification of glycerophosphatidylglycerol (PG) species.

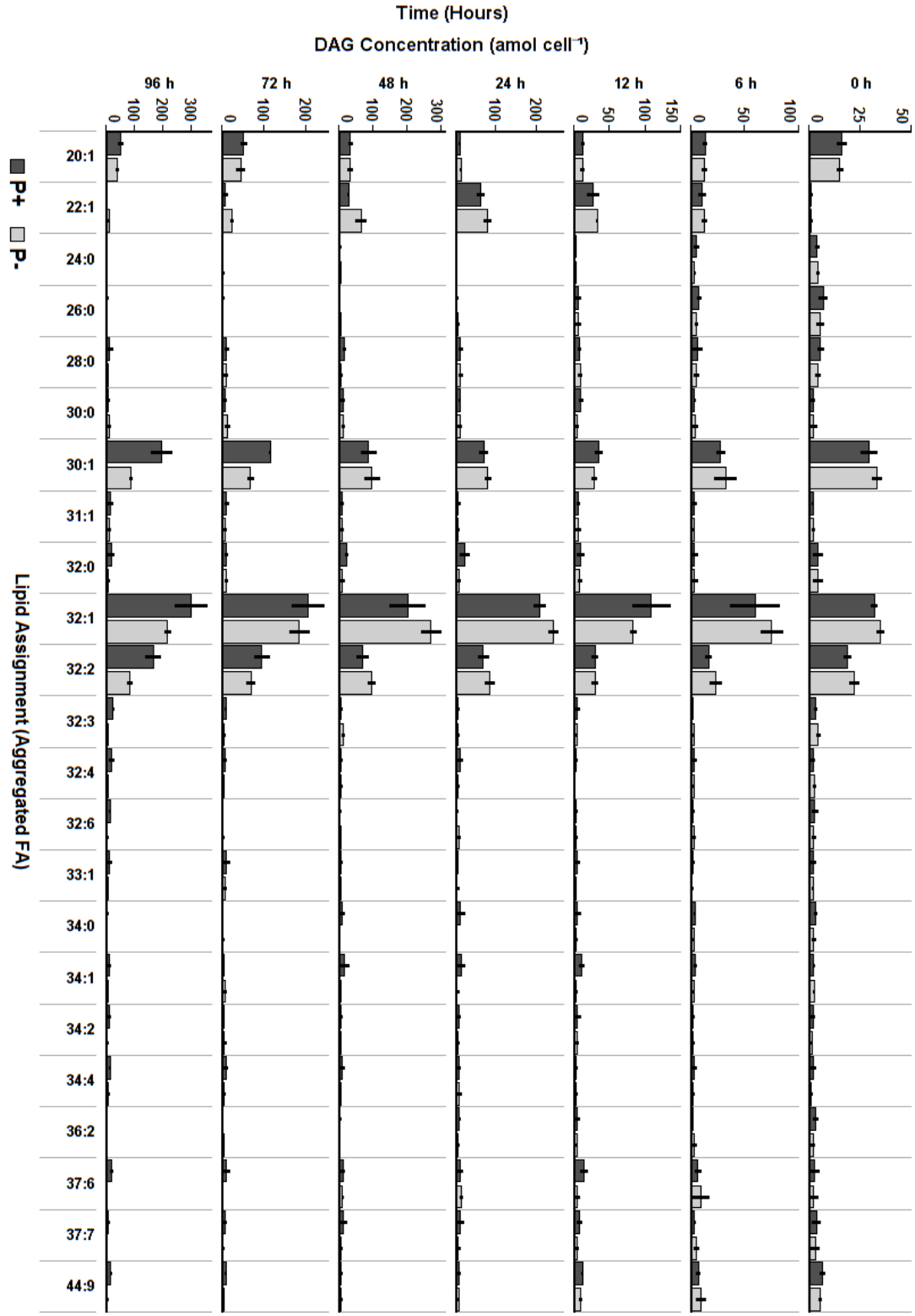


Figure 7.6: Full lipid dataset for Chapter 2, absolute quantification of diacylglycerol (DAG) species.

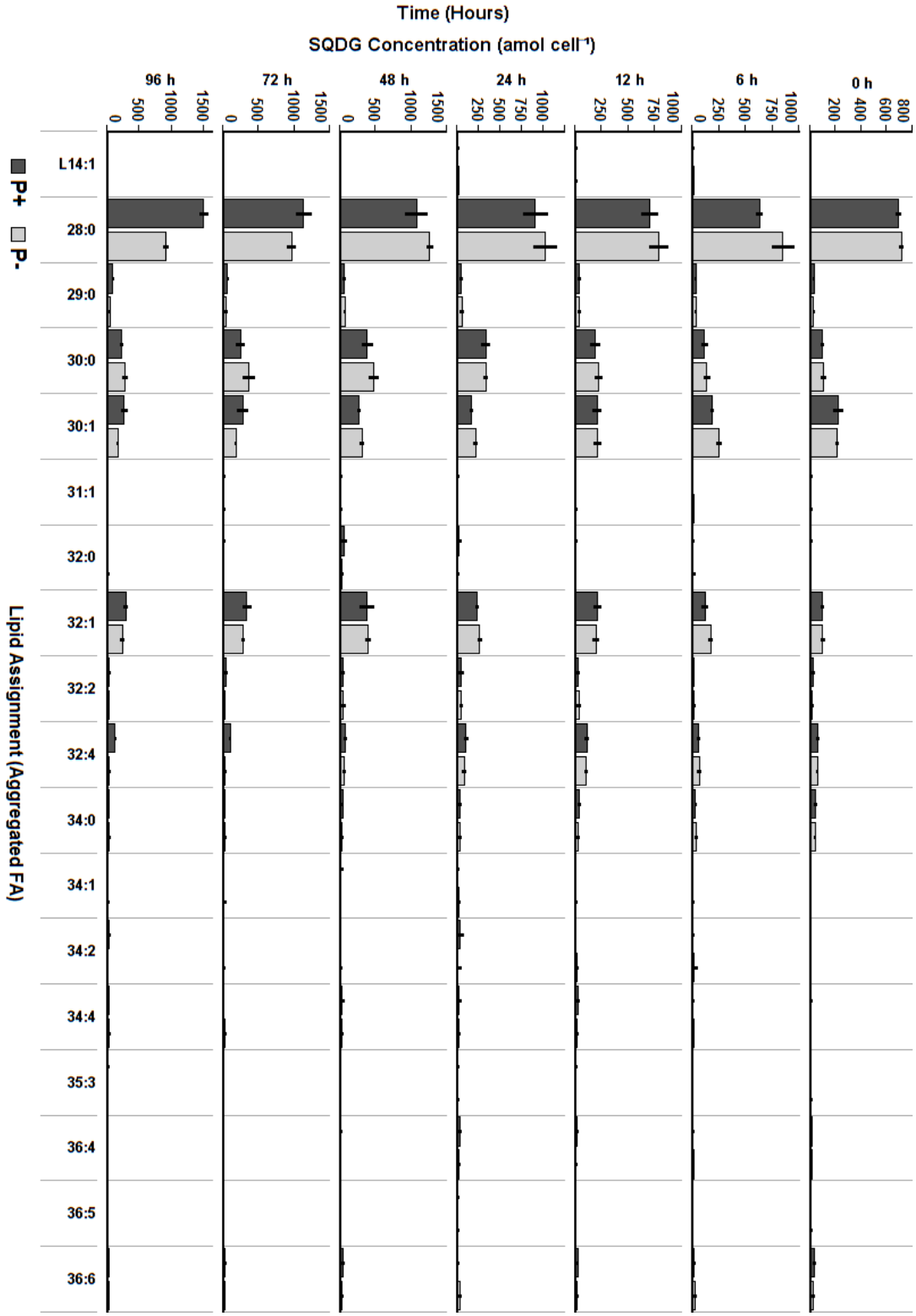


Figure 7.7: Full lipid dataset for Chapter 2, absolute quantification of sulfoquinovosyldiacylglycerol (SQDG) species.

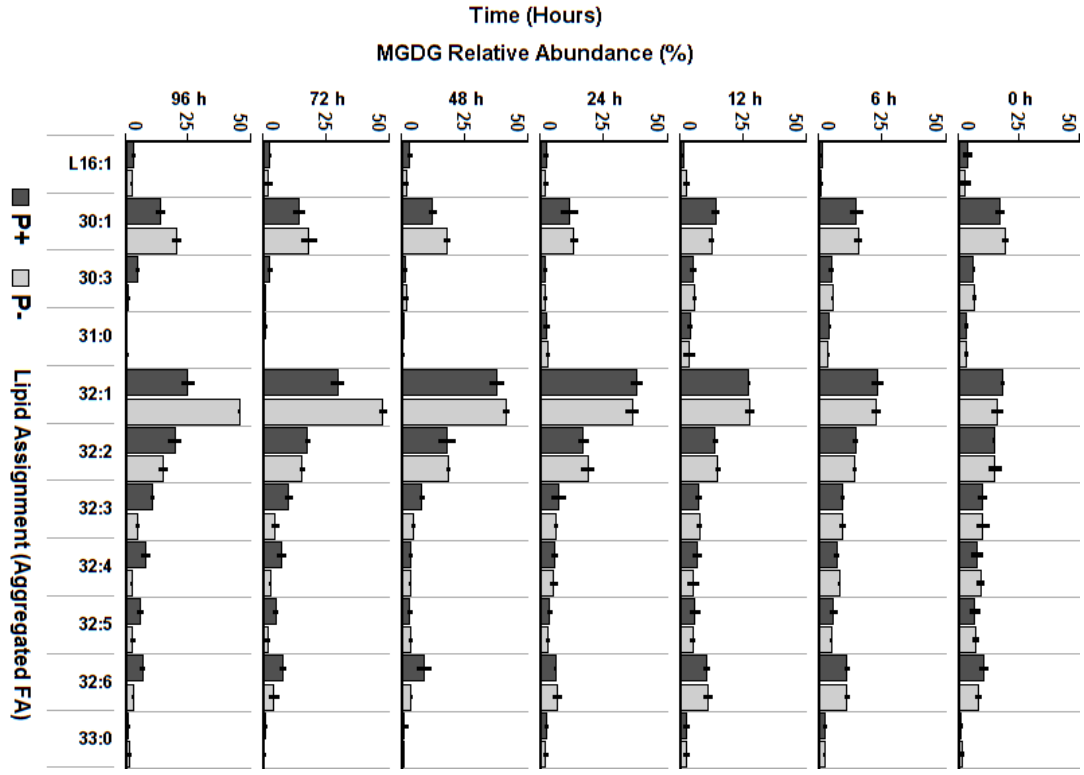


Figure 7.8: Full lipid dataset for Chapter 2, relative abundance of monogalactosyldiacylglycerol (MGDG) species.

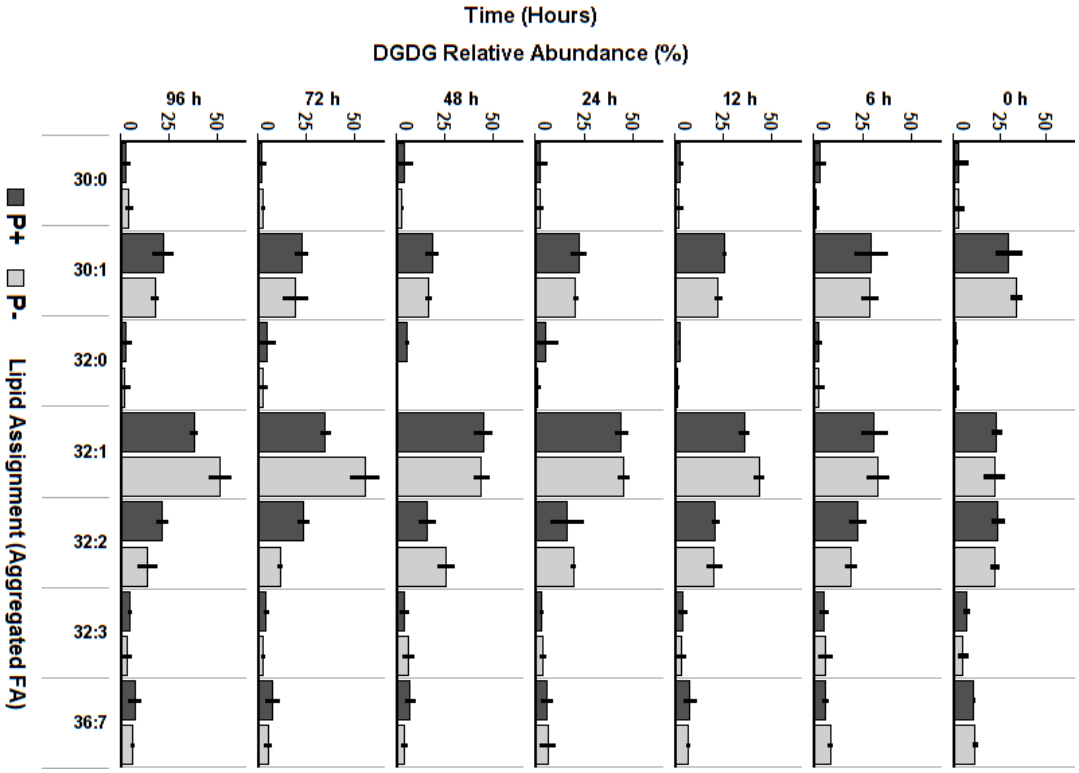


Figure 7.9: Full lipid dataset for Chapter 2, relative abundance of digalactosyldiacylglycerol (DGDG) species.

7.2 Appendix 2 - Supplementary Information for Chapter 4

Lipid Remodelling by the Diatom *Thalassiosira pseudonana* under Nitrogen Stress – Diglyceride Lipid Dynamics and Relation to Triglyeride Production

All data represent the mean of n=3 with error bars of 1 standard deviation. Lipid data shown for nitrogen replete (N+), nitrogen stressed (interface to stationary phase, N+/-) and nitrogen limited (N-) samples.

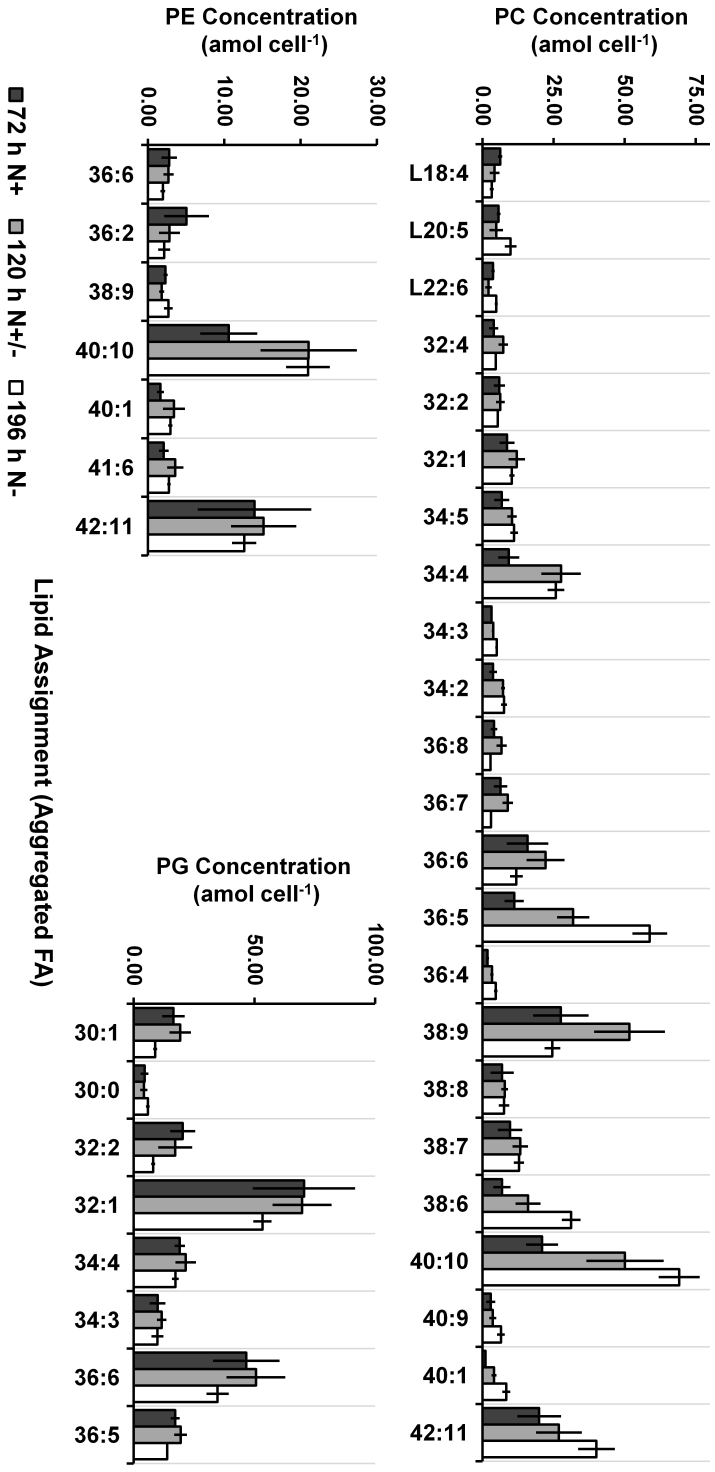


Figure 7.10: Full lipid dataset for Chapter 4, absolute quantification of glycerophosphatidylcholine (PC), glycerophosphatidylglycerol (PG) and glycerophosphatidylethanolamine (PE) species.

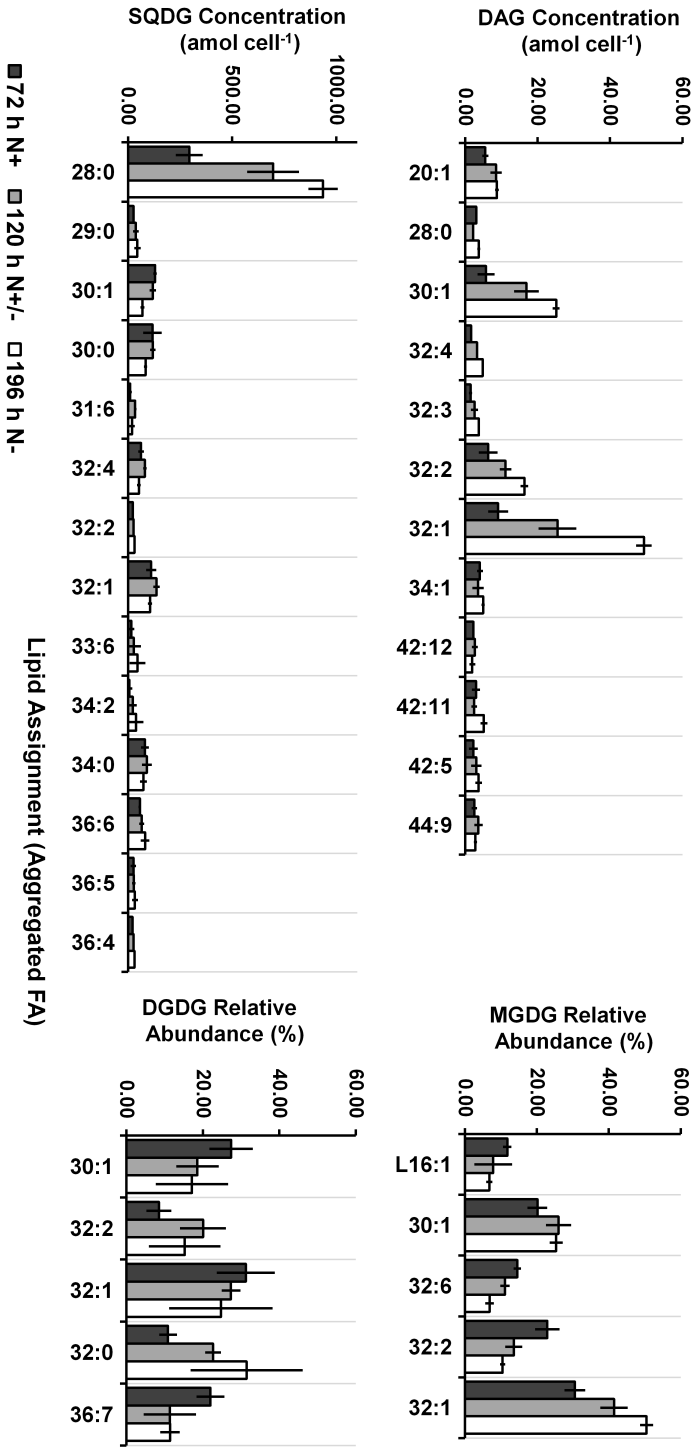


Figure 7.11: Full lipid dataset for Chapter 4, absolute quantification of sulfoquinovosyldiacylglycerol (SQDG) and diacylglycerol (DAG) species, relative abundance of monogalactosyldiacylglycerol (MGDG) and digalactosyldiacylglycerol (DGDG) species.

7.3 Appendix 3 - Supplementary Information for Chapter 5

Targeted and Untargeted Lipidomics of *Emiliania huxleyi* Viral Infection and Life Cycle Phases Highlights Molecular Biomarkers of Infection, Susceptibility, and Ploidy

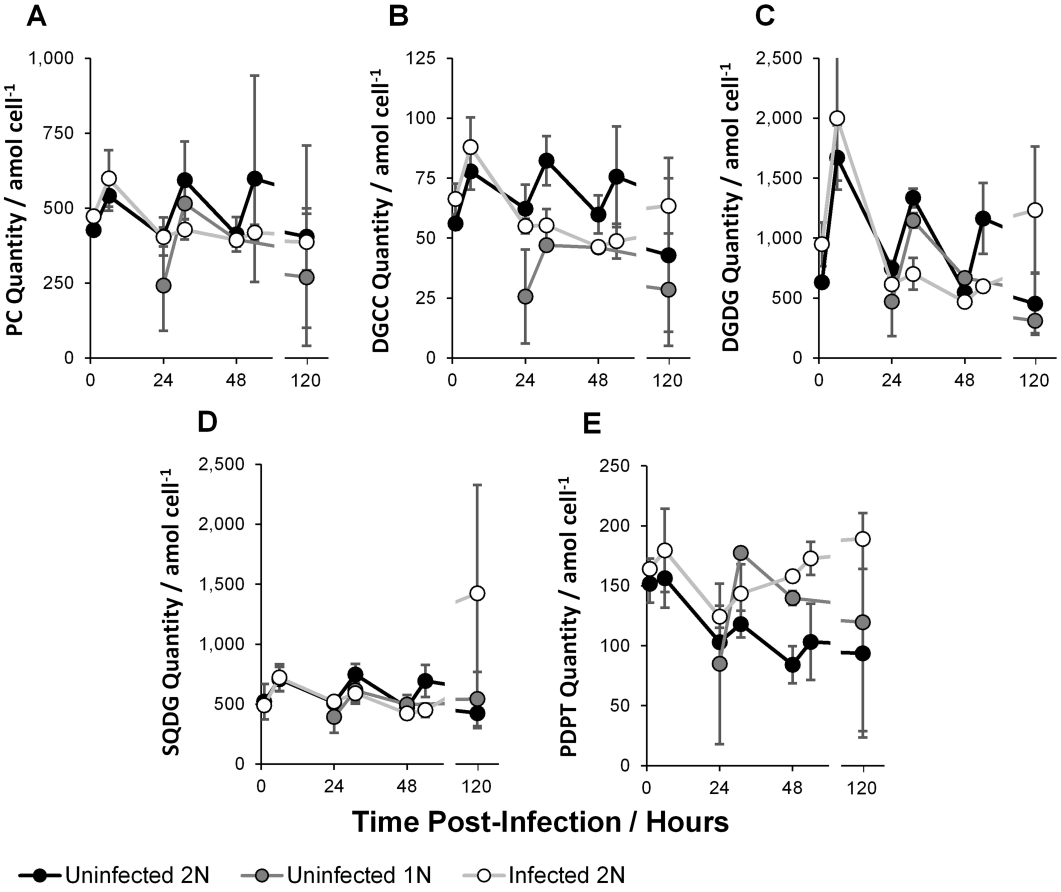


Figure 7.12: Polar glycerolipid quantity per *E. huxleyi* cell: Glycerophosphatidylcholine (PC, A); diacylglyceryl-3-O-carboxy-(hydroxymethyl)-choline (DGCC, B); Digalactosyldiacylglycerol (DGDG, C); sulfoquinovosyldiacylglycerol (SQDG, D) and phosphatidyl-S,S-dimethylpropanethiol (PDPT, E). Data represent the mean average of two biological replicates with error bars of one standard deviation. Statistically significant variations between cultures are described in the following section. P values were indicated as *$p<0.025$; **$p<0.0025$ where appropriate, as outlined in the Materials and Methods.

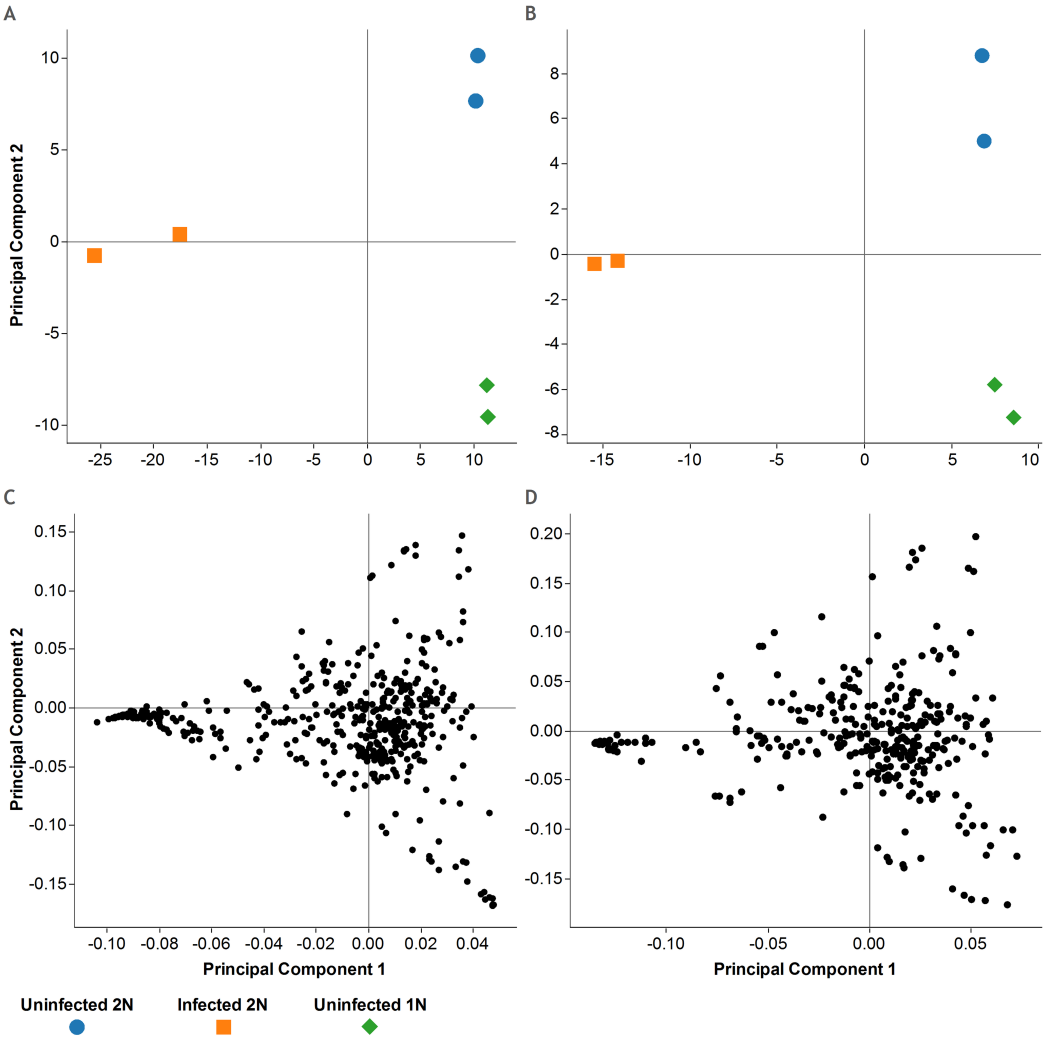


Figure 7.13: Partial Least Squares Discriminant Analysis (PLS-DA) of Untargeted Lipidomic Data (48 h): Positive Ion Mode, Score Plot (A); Loading Plot (C); Negative Ion Mode, Score Plot (B); Loading Plot (D); First Principal Component (PC1); Second Principal Component (PC2). The score plots (A+B) represent the model score for a given sample, colour coded by culture type. In the loading plots (C+D), each point represents the loading (or the contribution of a given molecular ion toward the respective scores). For example, ions with the most negative loadings on PC1 contribute greatest toward the infected diploid scores. We consider these ions indicative of the infected diploid cultures and hence biomarkers thereof.

A

Selection	Rank	MZ	R.T.	D.R.T.	ID	Adduct	ΔPPM	MS2/Other Evidence for Assignment.	MS2 ΔPPM
Uninfected 2N	1	872.6824	15.8	*	sGSL + H_2	(M+H)+	-0.3370	Neutral loss of 268.0794 (Sialic glycosyl headgroup)	0.1119
	2	868.6510	14.4	*	sGSL - H_2	(M+H)+	-0.1853	Neutral loss of 268.0794 (Sialic glycosyl headgroup)	-1.3800
	3	870.6666	15.2	*	sGSL	(M+H)+	-0.0870	Neutral loss of 268.0794 (Sialic glycosyl headgroup)	-1.7532
	4	892.6484	15.3	*	sGSL	(M+Na)+	0.0643	Coelution with other sGSL species	N/A
	5	752.5226	12.3	*	PC(34:5)	(M+H)+	-0.2163	Daughter of 184.0733. (Phosphatidyl choline headgroup)	0.1087
Uninfected 1N	1	549.5356	15.3	*	WE(31:0)	(M+2ACN+H)+	-0.5252	No co-eluting C31:0 fatty acid [M-H]- detected in negative	N/A
	2	508.5091	15.3	*	WE(31:0)	(M+ACN+H)+	-0.5653	Coelution with the 549.5356 species	N/A
	3	802.6768	17.3	*	GSL(t40:1)	(M+H)+	0.5615	Neutral loss of 180.0634 (glycosyl headgroup)	0.3373
	4	800.6614	17.0	*	GSL(t40:2)	(M+H)+	0.1505	Neutral loss of 180.0634 (glycosyl headgroup)	-0.6606
	5	738.5155	11.9	*	MGDG(32:5)	(M+NH_4)+	-0.5809	Neutral loss of 197.0899 (glycosyl headgroup + NH_3)	-0.1108
Infected 2N	1	840.7079	18.2	*	TAG(50:6)	(M+NH_4)+	-0.3801	Neutral loss of 245.2355 (14:0 FA+NH_3)	-0.1679
								Neutral loss of 291.2198 (18:5 FA+NH_3)	-0.3640
								Neutral loss of 299.2824 (18:1 FA+NH_3)	-0.0554
	2	858.7546	19.0	*	TAG(51:4)	(M+NH_4)+	-0.1424	Neutral loss of 259.2511 (15:0 FA+NH_3)	0.3336
								Neutral loss of 295.2511 (18:3 FA+NH_3)	-0.1770
								Neutral loss of 299.2824 (18:1 FA+NH_3)	-0.1251
	3	799.6788	19.4	*	TAG(46:1)	(M+Na)+	-0.2767	Neutral loss of 245.2355 (14:0 FA+NH_3)	-0.9463
								Neutral loss of 299.2824 (18:1 FA+NH_3)	-0.9487
	4	885.7650	18.7	*	TAG(50:4)	(M+ACN+NH_4)+	0.4524	Neutral loss of 245.2355 (14:0 FA+NH_3)	-0.5357
								Neutral loss of 295.2511 (18:3 FA+NH_3)	-1.1283
								Neutral loss of 299.2824 (18:1 FA+NH_3)	-0.4950
	5	856.7391	18.6	*	TAG(51:5)	(M+NH_4)+	-0.2433	Neutral loss of 259.2511 (15:0 FA+NH_3)	-0.6637
								Neutral loss of 293.2355 (18:4 FA+NH_3)	-0.9228
								Neutral loss of 299.2824 (18:1 FA+NH_3)	-0.7866

Figure 7.14: All ion MS2 fragmentation evidence (positive ions) in support of the untargeted lipidomic assignments presented in Figure 5.5. MZ represents the mass to charge ratio of a given molecular ion. Null represents species for which no identity hits were found in our lipid database. Time represents chromatographic retention time. ΔPPM is the difference between the measured and database MZ values for the parent ion in parts per million. R.T is chromatographic retention time and an asterix in the adjacent D.R.T. column designates a diagnostic retention time for the assigned species. MS2 ΔPPM is the difference between the measured and predicted MZ values for the fragment ion described in parts per million.

B

Selection	Rank	MZ	R.T.	D.R.T.	ID	Adduct	ΔPPM	MS2/Other Evidence for Assignment.	MS2 ΔPPM
Uninfected 2N	1	870.6694	15.8	*	sGSL + H_2	(M-H)-	-2.0473	Co-elution with sGSL species identified in positive ion mode	N/A
	2	868.6532	15.2	*	sGSL	(M-H)-	-1.4703	Co-elution with sGSL species identified in positive ion mode	N/A
	3	854.5800	12.6	*	BLL(18:1/22:6)	(M-H)-	-1.4577	Daughter of 190.0710 (Headgroup-H_2O) in positive ion mode	-2.6306
	4	827.4963	10.9	*	MGDG(36:9)	(M+Hac-H)-	-1.4944	Neutral loss of 197.0894 (Glycosyl headgroup + NH_3) from 786.5151 $[M+NH_4]+$ in positive ion mode	-1.8832
	5	949.5362	14.9	*	SQDG(40:8)	(M+Hac-H)-	-0.9924	Daughter of 225.00745 (Headgroup-H_2O) in negative ion mode	-0.6666
Uninfected 1N	1	623.3450	6.2	*	LMGDG(22:6)	(M+Hac-H)-	-2.0609	Daughter of 253.0918 (Glycosyl headgroup + Glycerol) in negative ion mode	0.3951
								Daughter of 327.2330 (22:6 FA) in negative ion mode	-1.0696
	2	696.6158	17.3	*	Cer(d18:1/22:0(OH))	(M+Hac-H)-	-1.5529	Daughter of 282.2791 (d18:1 long chain base - H_2O) in positive ion mode	0.2126
								Neutral loss of 355.3450 (22:0-OH fatty amide) from 638.6086 $[M+H]+$ in positive ion mode	-1.6592
	3	694.6004	17.1	*	Cer(d18:1/22:1(OH))	(M+Hac-H)-	-1.8571	Daughter of 282.2791 (d18:1 long chain base - H_2O) in positive ion mode	-0.4960
								Neutral loss of 353.3294 (22:1-OH fatty amide) from 636.5925 $[M+H]+$ in positive ion mode	-0.3530
	4	941.5497	11.2	*	DGDG(32:5)	(M+Hac-H)-	-1.8670	Daughter of 415.1457 (Diglycosyl headgroup + Glycerol) in negative ion mode	0.4577
	5	833.4686	12.0	*	DGDG(30:6)	(M-H_2O-H)-	0.7829	Daughter of 415.1457 (Diglycosyl headgroup + Glycerol) in negative ion mode	1.4212
Infected 2N	1	838.6196	16.3	*	vGSL	(M+Cl)-	-1.8107	Co-elution with vGSL species identified in negative ion mode	N/A
	2	802.6404	16.3	*	vGSL	(M-H)-	1.1583	Daughter of 268.2635 (t17:0 long chain base -$2H_2O$) in positive ion mode	-1.0810
								Neutral loss of 180.0634 (glycosyl headgroup) from 804.6559 in positive ion mode	-0.8165
	3	900.5646	10.9	*	BLL(22:6/22:6)	(M-H)-	-1.6144	Daughter of 190.0710 (Headgroup-H_2O) in positive ion mode	0.0000
	4	816.6564	16.7	*	vGSL + CH_2	(M-H)-	0.7207	Neutral loss of 180.6034 (glycosyl headgroup) from 818.6710 in positive ion mode	0.1409
	5	531.5149	18.0	*	WE(36:2)	(M-H)-	-0.4470	Daughter of 281.2486 (18:1 fatty acid) in negative ion mode	-0.3558

Figure 7.15: All ion MS2 fragmentation evidence (negative ions) in support of the untargeted lipidomic assignments presented in Figure 5.5. MZ represents the mass to charge ratio of a given molecular ion. Null represents species for which no identity hits were found in our lipid database. Time represents chromatographic retention time. ΔPPM is the difference between the measured and database MZ values for the parent ion in parts per million. R.T is chromatographic retention time and an asterix in the adjacent D.R.T. column designates a diagnostic retention time for the assigned species. MS2 ΔPPM is the difference between the measured and predicted MZ values for the fragment ion described in parts per million.

A

Selection	Rank	MZ	R.T.	ID	Adduct	P Value	PC1	PC2	Normalised Abundance		
									Uninfected 2N	Uninfected 1N	Infected 2N
Uninfected 2N	1	872.6824	15.8	sGSL + H2	(M+H)+	0.0055	0.02	0.14	0.81	0.00	0.19
	2	868.6510	14.4	sGSL - H2	(M+H)+	0.0050	0.01	0.13	0.79	0.00	0.21
	3	870.6666	15.2	sGSL	(M+H)+	0.0114	0.01	0.13	0.78	0.00	0.22
	4	914.6305	15.4	sGSL	(M+2Na-H)+	0.0289	0.01	0.13	0.77	0.00	0.23
	5	783.6397	17.4	Null	Null	0.0213	0.02	0.13	0.80	0.01	0.19
Uninfected 1N	1	549.5356	15.3	WE(31:0)	(M+2ACN+H)+	0.0136	0.05	-0.17	0.00	1.00	0.00
	2	809.7131	19.6	Null	Null	0.0420	0.05	-0.17	0.01	0.99	0.00
	3	508.5091	15.3	WE(31:0)	(M+ACN+H)+	0.0039	0.05	-0.17	0.00	1.00	0.00
	4	754.6710	19.3	Null	Null	0.0042	0.04	-0.16	0.00	0.97	0.02
	5	795.6976	19.3	Null	Null	0.0140	0.05	-0.16	0.02	0.98	0.00
Infected 2N	1	840.7079	18.2	TAG(50:6)	(M+NH4)+	0.0404	-0.10	-0.01	0.00	0.00	1.00
	2	858.7546	19.0	TAG(51:4)	(M+NH4)+	0.0373	-0.10	-0.01	0.00	0.00	1.00
	3	799.6788	19.4	TAG(46:1)	(M+Na)+	0.0348	-0.10	-0.01	0.00	0.00	1.00
	4	885.7650	18.7	TAG(48:2)	(M+2ACN+H)+	0.0331	-0.10	-0.01	0.00	0.00	1.00
	5	814.6978	19.5	Null	Null	0.0317	-0.10	-0.01	0.00	0.00	1.00

B

Uninfected 2N	1	870.6694	15.8	sGSL + H2	(M-H)-	0.0129	0.03	0.19	0.82	0.00	0.18
	2	868.6532	15.2	sGSL	(M-H)-	0.0177	0.02	0.18	0.79	0.00	0.21
	3	1,032.6599	15.3	Null	Null	0.0457	0.02	0.17	0.77	0.03	0.20
	4	950.6565	15.3	Null	Null	0.0309	0.02	0.17	0.75	0.03	0.22
	5	830.5065	11.0	Null	Null	0.0133	0.03	0.11	0.65	0.19	0.16
Uninfected 1N	1	623.3450	6.2	LMGDG(22:6)	(M+Hac-H)-	0.0113	0.06	-0.17	0.06	0.87	0.08
	2	944.6105	14.4	Null	Null	0.0002	0.05	-0.17	0.04	0.86	0.11
	3	862.6068	14.3	Null	Null	0.0000	0.05	-0.17	0.04	0.84	0.13
	4	916.5790	14.4	Null	Null	0.0000	0.04	-0.16	0.04	0.81	0.15
	5	696.6158	17.3	Null	Null	0.0013	0.02	-0.14	0.03	0.70	0.27
Infected 2N	1	1,108.6733	16.3	Null	Null	0.0018	-0.14	-0.01	0.00	0.00	1.00
	2	838.6196	16.3	vGSL	(M+Cl)-	0.0023	-0.13	-0.01	0.00	0.00	1.00
	3	1,026.6705	16.3	Null	Null	0.0004	-0.13	-0.01	0.00	0.00	1.00
	4	1,054.7020	17.1	Null	Null	0.0000	-0.13	-0.01	0.00	0.00	1.00
	5	931.6549	15.8	Null	Null	0.0004	-0.13	-0.01	0.00	0.00	1.00

Figure 7.16: Top 5 loadings including unassigned species for each culture type in the PLS-DA model: Positive ions (A) and Negative ions (B). Normalised abundance represents the average abundance of an ion in the specified culture type divided by the total abundance of that ion in all three cultures. MZ represents the mass to charge ratio of a given molecular ion. Null represents species for which no identity hits were found in our lipid database. R.T. represents chromatographic retention time. PC1 and PC2 represent the loading values for each ion in the PLS-DA model.

Chemical structures of minor lipid species observed in Chapter 5 are included below. Stereochemistry and regiochemistry of hydroxyl and double bond moeities is unknown and these are included for illustrative purposes.

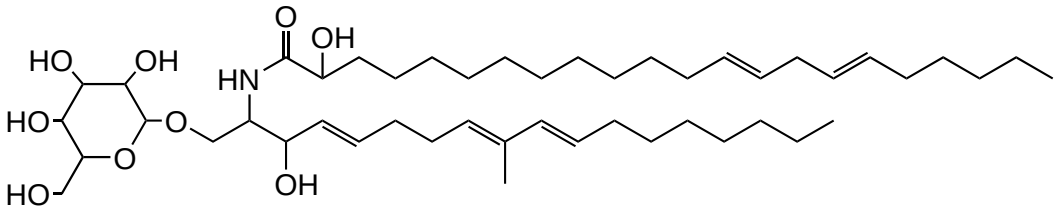


Figure 7.17: Chemical structure of host glycosphingolipid (hGSL) - adapted from reference (**Vardi et al.**, 2012).

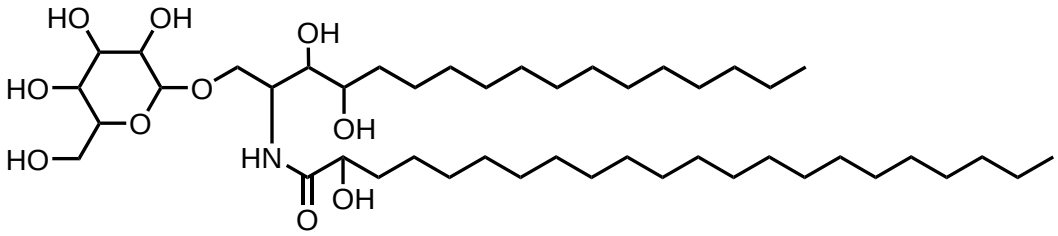


Figure 7.18: Chemical structure of viral glycosphingolipid (vGSL) - adapted from references (**Vardi et al.**, 2009, 2012)

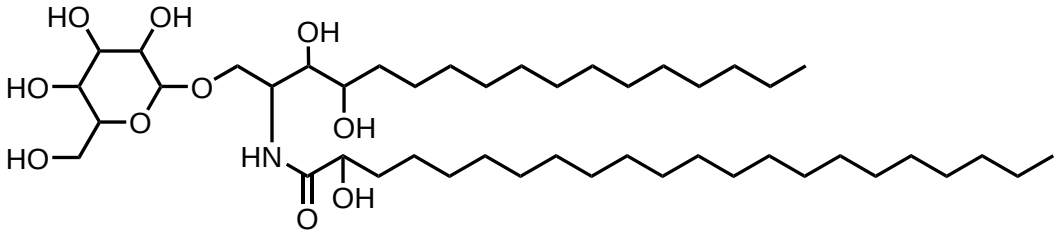


Figure 7.19: Chemical structure of sialic glycosphingolipid (sGSL) - adapted from reference (**Fulton et al.**, 2014)

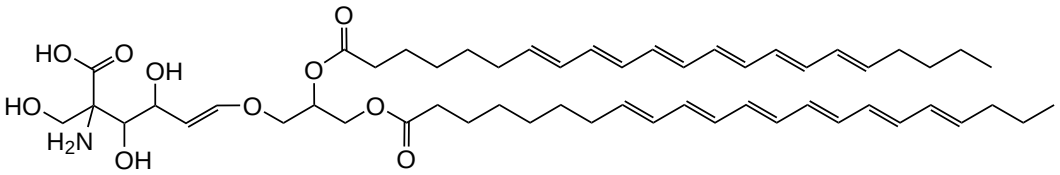


Figure 7.20: Chemical structure of betaine like lipid (BLL) - adapted from reference (**Fulton et al.**, 2014)

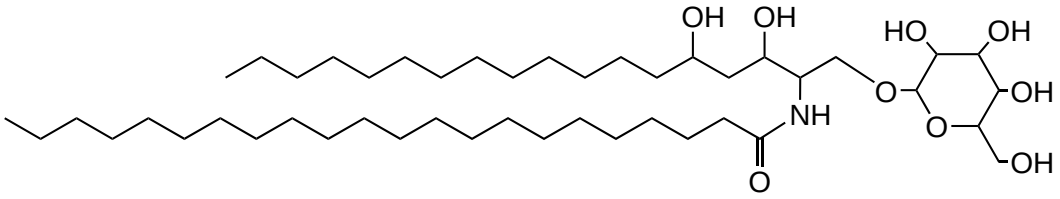


Figure 7.21: Tentative chemical structure of haploid glycosphingolipid (GSL(t40:0)) - see Chapter 5.

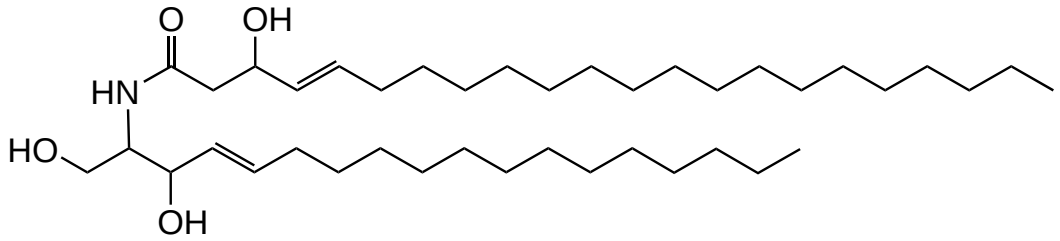


Figure 7.22: Tentative chemical structure of haploid ceramide (Cer(18:1/20:1(OH))) - see Chapter 5.

7.4 Appendix 4 - Supplementary Information for Chapter 5

Targeted and Untargeted Lipidomics of *Emiliania huxleyi* Viral Infection and Life Cycle Phases Highlights Molecular Biomarkers of Infection, Susceptibility, and Ploidy - Published Manuscript

frontiers in Marine Science

ORIGINAL RESEARCH
published: 13 October 2015
doi: 10.3389/fmars.2015.00081

Targeted and untargeted lipidomics of *Emiliania huxleyi* viral infection and life cycle phases highlights molecular biomarkers of infection, susceptibility, and ploidy

Jonathan E. Hunter[1,2]*, *Miguel J. Frada*[3], *Helen F. Fredricks*[4], *Assaf Vardi*[3] and *Benjamin A. S. Van Mooy*[4]

[1] Ocean and Earth Science, National Oceanography Centre, University of Southampton, Southampton, UK, [2] Institute for Life Sciences, University of Southampton, Southampton, UK, [3] Department of Plant and Environmental Sciences, Weizmann Institute of Science, Rehovot, Israel, [4] Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA, USA

Marine viruses that infect phytoplankton strongly influence the ecology and evolution of their hosts. *Emiliania huxleyi* is characterized by a biphasic life cycle composed of a diploid (2N) and haploid (1N) phase; diploid cells are susceptible to infection by specific coccolithoviruses, yet haploid cells are resistant. Glycosphingolipids (GSLs) play a role during infection, but their molecular distribution in haploid cells is unknown. We present mass spectrometric analyses of lipids from cultures of uninfected diploid, infected diploid, and uninfected haploid *E. huxleyi*. Known viral GSLs were present in the infected diploid cultures as expected, but surprisingly, trace amounts of viral GSLs were also detected in the uninfected haploid cells. Sialic-acid GSLs have been linked to viral susceptibility in diploid cells, but were found to be absent in the haploid cultures, suggesting a mechanism of haploid resistance to infection. Additional untargeted high-resolution mass spectrometry data processed via multivariate analysis unveiled a number of novel biomarkers of infected, non-infected, and haploid cells. These data expand our understanding on the dynamics of lipid metabolism during *E. huxleyi* host/virus interactions and highlight potential novel biomarkers for infection, susceptibility, and ploidy.

Keywords: ***Emiliania huxleyi*, lipidomics, coccolithovirus, glycosphingolipid, glycerolipids, haploid**

OPEN ACCESS

Edited by:
Anton F. Post,
Coastal Resources Center, University of Rhode Island, USA

Reviewed by:
Kathleen Scott,
University of South Florida, USA
Joaquin Martinez Martinez,
Blgelow Laboratory for Ocean Sciences, USA

***Correspondence:**
Jonathan E. Hunter
j.hunter@soton.ac.uk

Specialty section:
This article was submitted to Aquatic Microbiology, a section of the journal Frontiers in Marine Science

Received: *20 May 2015*
Accepted: *25 September 2015*
Published: *13 October 2015*

Citation:
Hunter JE, Frada MJ, Fredricks HF, Vardi A and Van Mooy BAS (2015) Targeted and untargeted lipidomics of Emiliania huxleyi viral infection and life cycle phases highlights molecular biomarkers of infection, susceptibility, and ploidy. Front. Mar. Sci. 2:81. doi: 10.3389/fmars.2015.00081

INTRODUCTION

Emiliania huxleyi (Lohmann) is the numerically dominant coccolithophore in the modern oceans and an important component of phytoplankton assemblages, inhabiting all but extreme polar oceans. Moreover, it forms large, dense blooms in high-latitude coastal and shelf ecosystems that exert a critical impact upon the global carbon cycle and the earth's climate (Westbroek et al., 1993; Paasche, 2001; Tyrrell and Merico, 2004). In later stages these blooms become visible to satellites as large scale cell death leads to the mass shedding of the highly scattering calcium carbonate coccoliths that normally coat the surface of *E. huxleyi* cells (Holligan et al., 1983; Vardi et al., 2012; Lehahn et al., 2014).

Marine viruses are the most abundant biological agents in the oceans (Fuhrman, 1999; Suttle, 2007). Specific, giant, lytic double-stranded DNA *E. huxleyi* viruses (EhV), belonging to the *phycodnaviridae* family that infect microalgae (Van Etten et al., 2002) are heavily implicated in the decay of *E. huxleyi* blooms (Bratbak et al., 1993; Brussaard et al., 1996; Vardi et al., 2012; Lehahn et al., 2014). Viruses that induce host cell lysis are thought to release particulate carbon and other nutrients into the water column (the "viral shunt"), thus circumventing the export of particulate organic matter to the deep ocean by way of the biological pump (Fuhrman, 1999; Suttle, 2007; Jover et al., 2014). Conversely, viral infection is also known to induce increased production of transparent exopolymer particles (TEP) in *E. huxleyi,* that accelerate the formation of sinking particulates, enhancing the biological pump, and removing virus particles from the upper water column (Vardi et al., 2012). The efficacy of the biological pump, and thus the effect of viruses on phytoplankton, has direct implications upon atmospheric carbon dioxide (Suttle, 2007).

At the cellular level, as a large dsDNA virus with high metabolic demand for the building blocks of DNA, lipids and protein synthesis, EhV triggers a rapid remodeling of the host metabolism (Rosenwasser et al., 2014; Schatz et al., 2014). In particular, recent studies have highlighted the crucial role of membrane lipids in the progression and regulation of EhV infection (Vardi et al., 2009, 2012; Rosenwasser et al., 2014). Evidence from genome and microscopic investigation suggests that EhV86 utilizes an animal-like, membrane-dependent infection strategy. EhV entry, by membrane fusion or endocytosis and the acquisition of host membrane lipids via budding (Mackinder et al., 2009; Schatz et al., 2014), seem to be localized to membrane lipid raft regions (Rose et al., 2014). Furthermore, the EhV genome contains a cluster of genes composing a nearly complete sphingolipid biosynthetic pathway analogous to the host pathway (Wilson et al., 2005; Monier et al., 2009).

Glycosphingolipids (GSLs) bearing a sphingoid base derived from palmitoyl-CoA—host glycosphingolipids (hGSL)—are abundant in uninfected *E. huxleyi*. Under lytic infection however, viral glycosphingolipids (vGSLs) derived from myristoyl-CoA are synthesized de novo, and vGSLs are known to play a role in the regulation of cell death in infected cells and are enriched in the membrane of newly formed virions (Vardi et al., 2009; Fulton et al., 2014; Rosenwasser et al., 2014). These vGSLs were detected in coccolithophore populations in the North Atlantic, which highlights their potential as biomarkers for viral infection in the oceans (Vardi et al., 2009, 2012).

A glycosphingolipid with a sialic acid modified glycosyl headgroup (sGSL) was recently described to have a direct relationship with susceptibility to infection. Across 11 strains of *E. huxleyi* (Fulton et al., 2014), sGSL was only detected at greater than trace levels in susceptible host strains, but not in viral-resistant host strains. Given the presence of a sialidase gene in the EhV genome (Wilson et al., 2005), it has been speculated that sGSL is a target for hydrolysis by EhV sialidases and/or a ligand for attachment by EhV lectin proteins during infection (Fulton et al., 2014). This mechanism is of an analogous fashion to a range of human viral pathogens including influenza (Stray et al., 2000). Additionally, betaine-like glycerolipids (BLL) were also recently described whose fatty acid composition appears highly indicative of the progression of infection. It was reported that uninfected *E. huxleyi* BLL composition was almost exclusively C16:0/C22:6 and C18:1/C22:6, while under viral infection the total BLL composition shifted to contain 50% C22:6/22:6 (Fulton et al., 2014).

A basic feature of *E. huxleyi* is the possession of a biphasic, haplodiplontic, and heteromorphic life cycle comprising a diploid (2N) coccolith-bearing phase that is non-motile and involved in the formation of blooms and a contrasting haploid (1N) form bearing flagella and non-mineralized organic body-scales (Green et al., 1996; Houdan et al., 2004). The vast majority of research on *E. huxleyi* has been conducted with 2N cells, and yet recent transcriptomic analyses have revealed a dramatic differentiation between 1N and 2N cells, with less than 50% of transcripts estimated to be shared between the two phases, unraveling a deep degree of physiological segregation (Von Dassow et al., 2009; Rokitta et al., 2011). Interestingly, whilst the 2N form is generally susceptible to EhV infection, the 1N form appears completely resistant to EhV (Frada et al., 2008). Furthermore, when 2N *E. huxleyi* are subject to EhV infection, a transition toward the resistant 1N flagellated form is induced, likely allowing for continuity of *E. huxleyi* following viral bloom termination (Frada et al., 2008, 2012).

Given the central role that membrane lipids play in the progression and regulation of EhV infection and the resistance to EhV infection exhibited by 1N *E. huxleyi*, a number of questions arise that are as yet unanswered. Early targeted analyses have shown similar compositions with respect to the major structural lipids, storage lipids, and pigments between 1N and 2N cells in the case of a single strain of *E. huxleyi* (Bell and Pond, 1996). Minor or novel lipid classes however, are as yet uncharacterized in the 1N cell lipidome and it is unknown whether the susceptibility marker sGSL is absent from EhV resistant 1N cells. By characterizing these lipids and the lipidome as a whole, we can gain insight into the mechanism of 1N *E. huxleyi* resistance to EhV infections and potentially highlight biomarkers of each of the life cycle phases.

We present herein, a detailed characterization of the lipidomes of cultured *E. huxleyi*, an uninfected 2N strain (RCC 1216), the 2N strain under infection with coccolithovirus (RCC 1216 + EhV201), and an uninfected 1N strain (RCC 1217). Total lipid extracts derived from these cultures collected over 120 h post-infection were characterized by mass spectrometry. We used two approaches, targeted analysis for quantification of known GSL/glycerolipid species and untargeted analyses for screening for unknown lipids.

MATERIALS AND METHODS

Culturing Procedures

The calcifying, 2N *Emiliania huxleyi* strain RCC1216 and the non-calcified, flagellated 1N *E. huxleyi* RCC1217 (isolated from RCC1216 following a partial phase change (2N to 1N), were used for this study (Houdan et al., 2005). Cells were cultured

in K/2 medium (Keller et al., 1987) and incubated at 18°C with a 16 : 8 h, light: dark illumination cycle. Light intensity was provided at 100 μmol photons m^{-2} s^{-1} with cool white LED lights. All experiments were performed in duplicate. The virus used for this study is the lytic *E. huxleyi* virus EhV201 (Schroeder et al., 2002) used at an initial multiplicity of infection (MOI) of 0.2 viral particles $cell^{-1}$. Samples of 25 mL were collected daily over 120 h post-infection by gentle vacuum-filtration onto pre-combusted GFF filters (Whatmann) and stored at −80°C until further analysis.

Targeted Lipid Analysis

Lipid abbreviations as indicated in the results section. Total lipid extracts were prepared from the cell isolates by a modified Bligh and Dyer extraction (Bligh and Dyer, 1959; Popendorf et al., 2013), with addition of the internal standard DNP-PE(16:0/16:0) (2,4-dinitrophenyl modified PE). The prepared total lipid extracts were then subjected to targeted lipid analysis by normal phase high performance liquid chromatography tandem mass spectrometry (HPLC-MS/MS) on an Agilent 1200 HPLC coupled to a Thermo Scientific TSQ Vantage triple quadrupole MS. Chromatography and mass spectrometry conditions were as described by Popendorf et al. (2013).

Lipid classes were identified by retention time and characteristic MS^2 fragmentation (Popendorf et al., 2013; Fulton et al., 2014) and quantified based upon peak area within a given MS^2 mass chromatogram. Quantification was achieved relative to external standard response factor calibrations. These calibrations were generated immediately before each analytical run, from a mixture of standards in a dilution series. PC, PG, PE, and DNP-PE (Avanti Polar Lipids, Alabaster, AL, USA); MGDG, DGDG (Matreya LLC, Pleasant Gap, PA, USA), and SQDG (Lipid Products, South Nutfield, UK) standards were used to generate calibrations for the glycerolipids. DGCC was quantified based on a purified extract from cultured *Thalassiosira Pseudonana* (Popendorf et al., 2013). DGTS was quantified based upon the DGCC calibration subject to a scaling factor (Popendorf et al., 2011b). The internal standard mixture was run after every seven samples as a control upon instrument variability. Quantities were corrected based upon the quality control run prior to a given sample. BLL and PDPT were quantified directly from the DNP-PE internal standard, as a semi-quantitative solution in the absence of an available internal standard. Quantification of hGSL, vGSL, and sGSL was based upon the response factor calibration of soy glucocerebroside extract (Avanti Polar Lipids, Alabaster, AL, USA) (Fulton et al., 2014).

While uninfected 1N cells have been reported to have a smaller diameter than calcified uninfected 2N cells, this difference is thought to be due primarily due to the presence of the calcified skeletons as opposed to large difference in cytoplasmic volume (Mausz and Pohnert, 2014). Furthermore, uncalcified 2N cells were found to be of comparable diameter to the 1N cells (Klaveness, 1972; Mausz and Pohnert, 2014). Therefore, we assume that the biovolume of 1N and 2N cells was similar, and the measured lipid quantities were normalized to cell concentration for ease of interpretation.

All data represent the mean average of two biological replicates with error bars of one standard deviation. Statistical significance for the targeted analyses was determined by two-tailed, paired equal-variance *T*-test. A Bonferroni correction was applied to the significance threshold to account for dual comparisons (uninfected 2N/uninfected 1N, uninfected 2N/infected 2N), thus a *p*-value of <0.025 was considered statistically significant. Any such significant variation is described in the results section. *P*-values were indicated as $^{*}p < 0.025$; $^{**}p < 0.0025$ where appropriate.

Untargeted Lipid Analysis

Total lipid extracts from the 48 h samples were also analyzed by untargeted reverse phase HPLC-MS methodology on an Agilent 1200 HPLC coupled to a Thermo Scientific Exactive orbitrap high resolution mass spectrometer. Chromatography and mass spectrometry conditions were as described by Hummel et al. (2011) with the exception of a Waters XBridge C8 column (5 uM packing, 150 × 2.1 mm). During each sample run, the mass spectrometer continuously cycled between full positive, full negative, positive all ion fragmentation, and negative all ion fragmentation modes, generating spectra with high mass accuracy.

The data was processed with the Thermo Scientific Sieve software package using the component extraction algorithm for chromatographic alignment, peak detection, and integration. Identified peak areas were then normalized to the DNP-PE internal standard and number of cells isolated. These unbiased analyses yield relative quantification; hence the abundance of a given molecular species is only comparable with the abundance of the same molecular species in other samples. The extracted data was filtered to remove molecular species where the deviation in abundance between replicates exceeded a factor of 10.

Peak areas were mean centered, level scaled to their means (Van den Berg et al., 2006) and used to build a Partial Least Square Discriminant Analysis Model (PLS-DA), using the Classification Toolbox for MATLAB (Ballabio and Consonni, 2013). PLS-DA models (Positive ions/Negative ions) were built using the following parameters: two components, Bayes assignment, six cross validation (CV) groups in contiguous blocks (i.e., Leave-One-Out validation as six samples were used). The positive model described 94% of the variance and had an error and CV error rate of 0. The negative model described 89% of the variance and had an error and CV error rate of 0. Ions from the mass spectrometry data were ranked upon their contribution (loading) toward the model score of samples of a given culture type (Supplementary Figure 2).

The top five ions ranked as indicative of each culture type in positive and negative ion mode were identified by matching to an extensive, accurate mass, structure query language (SQL) lipid database. The database was populated by permutations of fatty acids (chain length/degree of unsaturation) and common glycerolipids/sphingolipids. The complete LIPID MAPS (version 20130306) structural database (Sud et al., 2007) and MaConDa mass spectrometry contaminants database (Weber et al., 2012) were also included. The chemical formulae of database entries were then used to calculate accurate mass m/z values based upon

a list of common molecular ion adducts in ESI-MS (Huang et al., 1999). Database hits were within 2.5 ppm of the measured m/z. Supporting MS^2 fragmentation and diagnostic retention time information was recorded for the assignments presented in the Results and is included in Supplementary Figures 3, 4.

RESULTS

Host Cell and Viral Dynamics and Relative Abundances

Uninfected 2N and uninfected 1N cultures grew at a comparable average rate of 0.023 h^{-1}, reaching a population density of $4.07 \times 10^6 \pm 0.39 \times 10^6$ cells mL^{-1} and $4.79 \times 10^6 \pm 0.28 \times 10^6$cells mL^{-1} respectively within 120 h (**Figure 1A**). In contrast, the infected 2N culture population peaked at $5.07 \times 10^5 \pm 1.15 \times 10^4$ cells mL^{-1} at 31 h and rapidly declined thereafter concomitant with the emergence of EhV particles in the medium. EhV concentration peaked at 54 h at $9.17 \times 10^7 \pm 0.71 \times 10^7$ virions mL^{-1} in the infected 2N cultures (**Figure 1C**). The decline phase was accompanied by an increase in the percentage of non-calcified (low scatter) cells from 17.70 ± 0.73% at 0 h to 80.05 ± 2.63% at 48 h. Bacteria concentration (**Figure 1D**), whilst remaining low in the uninfected 2N control, increased at a rate of 0.037 h^{-1} in the infected 2N cultures after the onset of infection reaching a maximum of $1.15 \times 10^7 \pm 8.39 \times 10^5$ cells mL^{-1} at 120 h.

Glycerolipid Targeted Lipidomics

The polar glycerolipids phosphatidylcholine (PC), diacylglyceryl-3-O-carboxy-(hydroxymethyl)-choline (DGCC), digalactosyldiacylglycerol (DGDG), sulfoquinovosyldiacyl glycerol (SQDG), and the sulfur containing phospholipid phosphatidyl-S,S-dimethylpropanethiol (PDPT) (Fulton et al., 2014) did not vary between cultures (Supplementary Figure 1). However, several other classes of polar glycerolipid showed interesting differences.

Phosphatidylglycerol (PG) quantity per cell (**Figure 2A**) in the uninfected 2N and uninfected 1N cultures was statistically similar through time. PG per cell in the infected 2N cultures initially declined by 0.27 ± 0.023* fold at 31 h relative to the uninfected 2N control. A large increase in infected 2N PG per cell of 30.29 ± 21.80 fold was observed at 120 h, although relative to the uninfected 2N control cultures the increase was not statistically significant ($p = 0.12$), due to a large variation between the biological replicates at this time point.

Phosphatidylethanolamine (PE) quantity per cell (**Figure 2B**) showed a decreasing trend through time in the uninfected 1N and uninfected 2N cultures. PE was slightly elevated in the uninfected 1N cultures compared to the uninfected 2N control, this was statistically significant only at 24 h, where PE per cell in the uninfected 1N was 1.57 ± 0.16* fold higher than in uninfected 2N. PE quantity per cell in the infected 2N cultures diverged from the uninfected 2N control after 24 h, progressively increasing to 4.58 ± 0.18** fold greater at 54 h. In common with PG, a large increase of infected 2N PE per cell at 120 h, of 180.10 ± 105.19 fold, was not statistically significant ($p = 0.10$), due to variability in the replicate samples at this time point.

Two non-phosphorous polar glycerolipids showed interesting and significant dynamics. Monogalactosyldiacylglycerol (MGDG) quantity per cell (**Figure 2C**) was statistically similar in all of the cultures before 52 h, but showed a decreasing trend with time. However, at 120 h MGDG in the infected 2N cultures increased by 9.31 ± 5.68* fold, relative to the uninfected 2N cultures. Diacylglyceryltrimethylhomoserine (DGTS) quantity per cell (**Figure 2D**) appeared to be approximately half as abundant in the uninfected 1N compared with the uninfected 2N control cultures at 0.55 ± 0.12* fold less at 24 h and 0.66 ± 0.01** fold less at 48 h. DGTS per cell in the infected 2N cultures did not vary statistically significantly from the uninfected 2N control, with the exception of a 1.29 ± 0.02** fold increase at 48 h. Similar to PG, DGTS levels in the uninfected 2N and uninfected 1N cultures remained relatively consistent through time.

Total Betaine-Like Lipid (BLL; Fulton et al., 2014) quantity per cell (**Figure 3A**) varied considerably in all cultures, ranging from a minimum of 35.55 ± 21.35 amol $cell^{-1}$ at 24 h in the uninfected 1N cultures to a maximum of 86.07 ± 27.21 amol

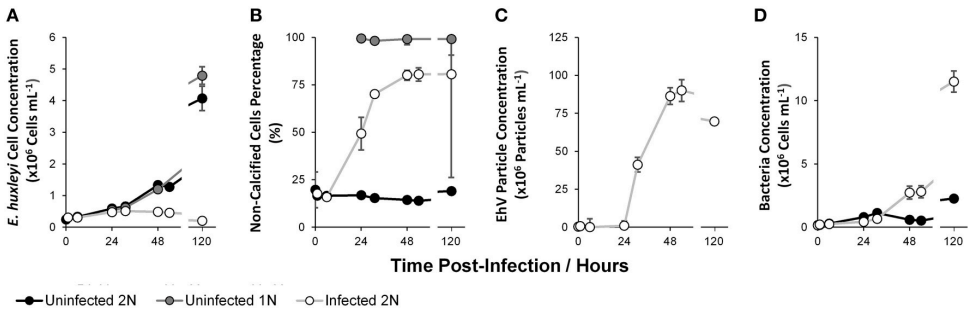


FIGURE 1 | Host cell and viral dynamics during infection: Total Host Cell Concentration (A); Non-calcified Cell Percentage of Total (Host Cell, B); EhV Particle Concentration (C); Bacterial Cell Concentration (D). Data represent the mean average of two biological replicates with error bars of one standard deviation.

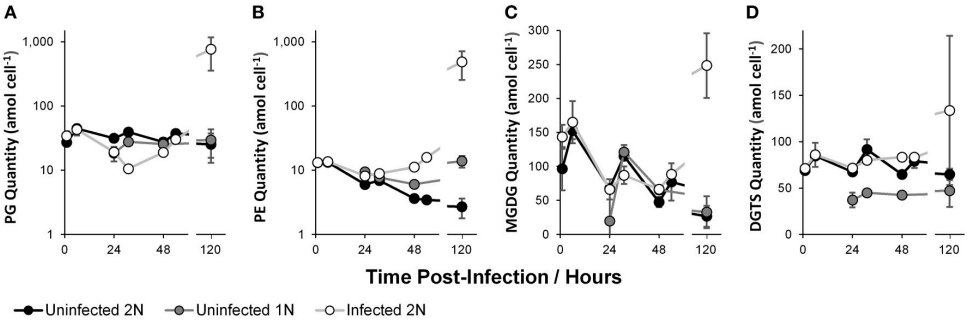


FIGURE 2 | Polar glycerolipid quantity per E. huxleyi cell: Phosphatidylglycerol (PG, A); Phosphatidylethanolamine (PE, B); Monogalactosyldiacylglycerol (MGDG, C); and Diacylglyceryltrimethylhomoserine (DGTS, D). Data represent the mean average of two biological replicates with error bars of one standard deviation. See Supplementary Table 1 for tabulated data.

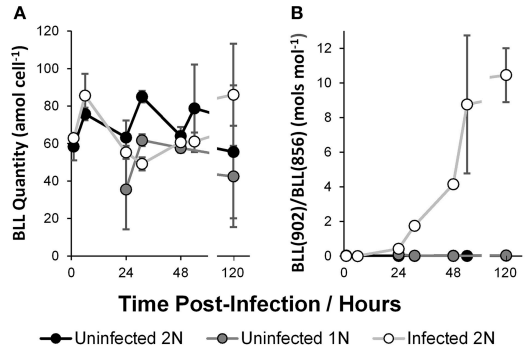


FIGURE 3 | Betaine-Like Lipid (BLL), total quantity per cell (A), and ratio of BLL(902; 22:6/22:6) quantity to BLL(856; 18:1/22:6) quantity (B). Data represent the mean average of *two* biological replicates with error bars of *one* standard deviation. See Supplementary Table 1 for tabulated data.

cell^{-1} at 120 h in the infected 2N cultures. Only at 31 h were significant differences observed, where the infected 2N cultures were 0.58 ± 0.05 fold less and the uninfected 1N cultures were 0.73 ± 0.05* fold less than in the uninfected 2N control. No clear trend was evident in the total BLL quantity per cell, in response to time or between the different cultures. However, the ratio between two BLL molecular species, BLL(22:6/22:6) and BLL(18:1/22:6) showed marked differences (**Figure 3B**). In the uninfected 2N and uninfected 1N cultures, the ratio was <0.1 throughout, representing the relative absence of BLL(22:6/22:6) compared to BLL(18:1/22:6) when uninfected. The ratio of these species rose rapidly under infection reaching 10.46 ± 1.56 at 120 h after inoculation of the infected 2N cultures.

Glycosphingolipid Targeted Lipidomics

Concentrations of two classes of glycosphingolipids were highly dynamic through the course of infection. The infection marker viral glycosphingolipid (vGSL, **Figure 4A**) was abundant only in the infected 2N cultures. Infected 2N vGSL concentration rose consistently through time from absent at 0 h to 169.32 ± 73.77 amol cell^{-1} at 120 h, concomitant with the increase in EhV concentrations. Interestingly, lower levels of vGSL (4–8 amol cell^{-1}) were also detected in the uninfected 1N cultures, reaching a maximum of 7.63 ± 0.13 amol cell^{-1} at 120 h.

The susceptibility marker sialic glycosphingolipid (sGSL, **Figure 4C**) was abundant in the uninfected 2N control, with no clear time dependence, between a range of 132.32 ± 4.09 amol cell^{-1} at 24 h and 193.76 ± 3.20 amol cell^{-1} at 6 h. The concentrations of sGSL in the infected 2N cultures displayed an approximately decreasing trend from 191.07 ± 35.35 amol cell^{-1} at 6 h to 51.59 ± 0.57 amol cell^{-1} at 54 h. Importantly, sGSL was absent from the uninfected 1N cultures. In contrast to vGSL and sGSL, a third class of glycosphingolipid, the intrinsic host glycosphingolipid (hGSL, **Figure 4B**) presented no statistically significant variation between cultures, and no trend was evident with the progression of time.

Untargeted Lipidomics and Biomarker Selection

Untargeted lipidomic data were collected from the incubations at 48 h and subsequent PLS-DA analysis revealed a number of potential biomarkers for *E. huxleyi* life stages and viral infection. The top 5 loadings that were indicative for each culture type, bearing significant differences by univariate statistics ($p =< 0.05$ by single factor ANOVA), were assigned by database matching (**Table 1**). All ion fragmentation MS^2 data was used to confirm database hits (Supplementary Figures 3, 4).

The potential biomarkers for the 2N control cultures were sGSLs, bearing various adducts and plus or minus one double bond resulting in the highest positive loadings on PC2. In positive ion mode (**Table 1A**), tentatively assigned sGSL+H_2, sGSL–H_2, and sGSL (with two different adducts) gave normalized abundances of 0.81, 0.79, 0.78, and 0.72 respectively, where the sGSL of calculated m/z 870.6665 $[M+H]^+$ is regarded as the archetypal sGSL (Fulton et al., 2014). A 752.5226 m/z species, with a database hit to PC(34:5) showed a normalized abundance of 0.60. In negative ion mode (**Table 1B**), sGSL+H_2, sGSL and

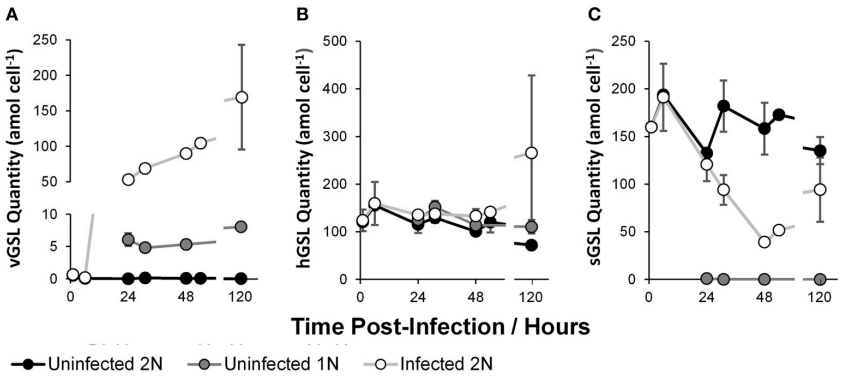


FIGURE 4 | Glycosphingolipid quantity per E. huxleyi cell: Viral glycosphingolipid (vGSL, A); Host glycosphingolipid (hGSL, B); Sialic-acid glycosphingolipid (sGSL, C). Data represent the mean average of *two* biological replicates with error bars of *one* standard deviation. See Supplementary Table 1 for tabulated data.

BLL(18:1/22:6) showed a normalized abundance of 0.82, 0.79, and 0.61 respectively. Species of m/z 827.4963 and 949.5362 with hits to MGDG(36:9) and SQDG(40:8) respectively showed normalized abundances of 0.60 and 0.58.

The top potential biomarker for uninfected 1N cells, in positive ion mode, was a wax ester/alkanoate WE(31:0) with two and one acetonitrile adducts, which was detected solely in the 1N cultures yielding normalized abundances of 1.00 and 1.00 respectively. Two GSLs, GSL(t40:0) and GSL(t40:1), distinct to previously documented GSLs in the *E. huxleyi* host/virus system, had normalized abundances of 0.67 and 0.65 respectively. Finally, MGDG(32:5) was ranked fifth with a normalized abundance of 0.70. In negative ion mode, a lyso-MGDG (bearing a single fatty acid rather than a pair), LMGDG(22:6) was most strongly indicative of the uninfected 1N cultures with a normalized abundance of 0.87. Ceramide species, chemically similar to sGSL but lacking any headgroup and bearing a hydroxy fatty amide, Cer(d18:1/22:0(OH)), and Cer(d18:1/22:1(OH)) in the uninfected 1N cultures had normalized abundances of 0.70 and 0.61 respectively. Lastly, digalactosyldiacylglycerols and DGDG(32:5) and DGDG(30:6) were enriched in the uninfected 1N cultures with normalized abundances of 0.71 and 0.69 respectively.

The potential biomarkers for the infected 2N cells were highly diagnostic, with 7 out of the top 10 yielding normalized abundances of 1.00. In positive ion mode, a series of triacylglycerols (TAGs) gave the highest negative loadings on PC1 and were ranked highest. TAG(50:6); TAG(51:4); TAG(46:1); TAG(48:2); and TAG(51:5) each had normalized abundances of 1.00. In negative ion mode, previously described biomarkers vGSL (with a chloride adduct); vGSL; BLL(22:6/22:6); and vGSL + CH_2 had normalized abundances of 1.00, 1.00, 0.98, 0.97, and 0.86 respectively, where vGSL of calculated m/z 802.6414 $[M\text{-}H]^-$ is regarded as the archetypal vGSL (Vardi et al., 2009). A fatty acid/wax ester like FA/WE(36:2) species was ranked fifth in negative ion mode with a normalized abundance of 0.86.

Further to the assigned species (**Table 1**) there were many molecular species of interest that could not be assigned by accurate mass database matching (Supplementary Table 2).

DISCUSSION

Host Cell and Viral Dynamics and Relative Abundances

Uninfected 2N and uninfected 1N *E. huxleyi* populations increased exponentially through time with a similar growth rate and maximum population, but the infected 2N cultures peaked at 31 h and declined thereafter, consistent with typical lytic viral infection by EhV (Bratbak et al., 1993; Wilson et al., 2002) (**Figure 1A**). This decline was concomitant with a rapid increase in EhV particles and non-calcified, low side scatter *E. huxleyi* cell concentration in the growth media, indicating both viral burst from infected cells and the demise of 2N calcified cells (**Figures 1B,C**). These growth dynamics reflect identical trends to those previously described (Frada et al., 2008, 2012). High scatter cells were predominate in the uninfected 2N control, representing the coccolith bearing 2N phase. Conversely, cells in the uninfected 1N cultures were mostly low scatter, indicative of the non-calcifying 1N phase.

The cultures used in this study were non-axenic, and a rapid increase in bacterial numbers in the infected 2N cultures was likely fuelled by the release of organic carbon during cell lysis. However, in terms of overall biomass, the level of contamination from bacteria is comparatively minor: considering the cell concentrations of *E. huxlyei* and bacteria, and cellular carbon quotas of 2 pmol C $cell^{-1}$ and 1 fmol C $cell^{-1}$, respectively (Gundersen et al., 2002; Borchard and Engle, 2012), *E. huxleyi* dominate biomass by approximately a factor of 20. Based on the cell counts of the uninfected 2N and uninfected 1N cultures, we would expect this factor to be orders of magnitude higher. Thus, except in the case of a few specific lipid molecules (discussed

TABLE 1 | Top 5 loadings (assigned species only) for each culture type in the PLS-DA model: Positive ions (A) and Negative ions (B).

A - Positive Ions

Selection	Rank	M/Z	Retention Time (Min)	Identity	Adduct	P	Normalised Abundance Uninfected 2N	Uninfected 1N	Infected 2N
Uninfected 2N	1	872.6824	15.8	sGSL + H_2	(M+H)+	*	0.81	0.00	0.19
	2	868.6510	14.4	sGSL - H_2	(M+H)+	*	0.79	0.00	0.21
	3	870.6666	15.2	sGSL	(M+H)+	*	0.78	0.00	0.22
	4	892.6484	15.3	sGSL	(M+Na)+	*	0.72	0.02	0.26
	5	752.5226	12.3	PC(34:5)	(M+H)+	*	0.60	0.15	0.25
Uninfected 1N	1	549.5356	15.3	WE(31:0)	(M+2AcN+H)+	*	0.00	1.00	0.00
	2	508.5091	15.3	WE(31:0)	(M+AcN+H)+	**	0.00	1.00	0.00
	3	802.6768	17.3	GSL(t40:0)	(M+H)+	**	0.04	0.67	0.29
	4	800.6614	17.0	GSL(t40:1)	(M+H)+	**	0.05	0.65	0.30
	5	738.5155	11.9	MGDG(32:5)	(M+NH4)+	**	0.22	0.70	0.09
Infected 2N	1	840.7079	18.2	TAG(50:6)	(M+NH4)+	*	0.00	0.00	1.00
	2	858.7546	19.0	TAG(51:4)	(M+NH4)+	*	0.00	0.00	1.00
	3	799.6788	19.4	TAG(46:1)	(M+Na)+	*	0.00	0.00	1.00
	4	885.7650	18.7	TAG(50:4)	(M+AcN+NH4)+	*	0.00	0.00	1.00
	5	856.7391	18.6	TAG(51:5)	(M+NH4)+	*	0.00	0.00	1.00

B - Negative Ions

Uninfected 2N	1	870.6694	15.8	sGSL + H_2	(M-H)-	*	0.82	0.00	0.18
	2	868.6532	15.2	sGSL	(M-H)-	*	0.79	0.00	0.21
	3	854.5800	12.6	BLL(18:1/22:6)	(M-H)-	*	0.61	0.27	0.13
	4	827.4963	10.9	MGDG(36:9)	(M+HAc-H)-	**	0.60	0.23	0.17
	5	949.5362	14.9	SQDG(40:8)	(M+HAc-H)-	**	0.58	0.26	0.16
Uninfected 1N	1	623.3450	6.2	LMGDG(22:6)	(M+HAc-H)-	*	0.06	0.87	0.08
	2	696.6158	17.3	Cer(d18:1/22:0(OH))	(M+HAc-H)-	**	0.03	0.70	0.27
	3	694.6004	17.1	Cer(d18:1/22:1(OH))	(M+HAc-H)-	**	0.06	0.61	0.33
	4	941.5497	11.2	DGDG(32:5)	(M+HAc-H)-	*	0.19	0.71	0.11
	5	833.4686	12.0	DGDG(30:6)	(M-H2O-H)-	*	0.22	0.69	0.09
Infected 2N	1	838.6196	16.3	vGSL	(M+Cl)-	**	0.00	0.00	1.00
	2	802.6404	16.3	vGSL	(M-H)-	**	0.00	0.00	1.00
	3	900.5646	10.9	BLL(22:6/22:6)	(M-H)-	**	0.01	0.01	0.98
	4	816.6564	16.7	vGSL + CH_2	(M-H)-	**	0.00	0.03	0.97
	5	531.5149	18.0	WE(36:2)	(M-H)-	**	0.06	0.07	0.86

*Normalized abundance represents the average abundance of an ion in the specified culture type divided by the total abundance of that ion in all three cultures. M/Z represents the mass to charge ratio of a given molecular ion. AcN and HAc represent acetonitrile and acetate adducts respectively. *p = <0.05, **p = <0.005 by single factor ANOVA.*

below) we would expect the lipidomes of our samples to be overwhelmingly dominated by contributions from *E. huxleyi*.

Glycerolipid Targeted Lipidomics

For the most part, polar glycerolipids per cell were largely invariant between cultures and with time, with only a few notable exceptions. Despite the relatively small contributions of bacterial biomass, the observed increase in PG concentration in the infected 2N cultures (**Figure 2A**) was most likely due to an increase in the bacterial population (**Figure 1D**). Thus, this result is not entirely unexpected since PG generally composes about half of the glycerolipids in marine bacteria, while in *E. huxleyi* PG only

represents 0.2% of the lipidome (Fulton et al., 2014; Carini et al., 2015).

In the infected 2N cultures, PE quantity also increased compared to the uninfected 2N control after 24 h (**Figure 2B**). Like PG, PE is also a major component of bacterial membranes but is scarce in *E. huxleyi* (Van Mooy et al., 2009; Popendorf et al., 2011a; Carini et al., 2015), and thus a contribution from bacteria is possible (Fulton et al., 2014). Alternatively, this increase in PE may be derived from the host cell's autophagy machinery (Schatz et al., 2014). Autophagy is a highly conserved eukaryotic mechanism for the degradation of damaged organelles and unwanted macromolecules (Shemi et al., 2015). It has been demonstrated that autophagy is induced during the lytic phase of EhV infection and plays a role in the propagation of the virus (Schatz et al., 2014). A well-known hallmark of autophagy is PE lipidation of the Atg8 protein, as also reported in *E. huxleyi* under EhV infection (Schatz et al., 2014). Thus, the observed increase in PE under EhV infection may be attributable to the induction of autophagy via lipidation of the Atg8 protein. Interestingly, total PE appeared slightly more abundant per cell in the uninfected 1N cultures, relative to the uninfected 2N control.

MGDG quantity per cell was similar between uninfected 2N and uninfected 1N cultures, while the infected 2N cultures showed a large increase in MGDG quantity per cell at 120 h post-inoculation (**Figure 2C**). MGDG is often associated with the thylakoid membranes of the chloroplast (Sakurai et al., 2006) and hence variation may indicate a consequence of infection upon the host photosynthetic machinery. In contrast, a decrease in MGDG under EhV infection of *E. huxleyi* has been shown elsewhere (Fulton et al., 2014). Furthermore, a recent study on another haptophyte *Haptolina ericina* infected with a dsDNA virus observed a decrease in glycolipid abundance under infection (Ray et al., 2014). This decrease was suggested to derive from cell disruption during infection, resulting in a loss of chloroplasts during sample preparation. The haptophyte *Phaeocystis pouchetii* did not display a major loss of cellular MGDG when infected with a dsDNA virus (Ray et al., 2014). MGDG is one of the most abundant lipids in *E. huxleyi* (Fulton et al., 2014), but this molecule is scarce in bacteria because they have a clear preference for synthesizing phospholipids over glycolipids when phosphate is abundant, as in the K/2 medium used here (Popendorf et al., 2011a; Carini et al., 2015); thus, it is unlikely that the MGDG observed in the infected 2N cultures is of bacterial origin. Clearly at this stage we have an incomplete understanding of the behavior of MGDG under viral infection in haptophytes such as *E. huxleyi*.

Total DGTS abundance per cell in the uninfected 1N cultures was consistently around half of its abundance in the uninfected 2N control (**Figure 2D**). Betaine lipids such as DGTS are known to substitute for membrane phospholipids (primarily PC) under phosphorus stress in many eukaryotic phytoplankton. As such the BL/PC ratio has been considered as a measure for phosphorus limitation in the environment (Van Mooy et al., 2009; Martin et al., 2011). Reduced DGTS abundance from 1N cells within an environmental population would impact such a measure. Furthermore, when coupled with a slight increase in PE in the 1N case, outlined earlier, this may shift the glycerolipid stoichiometry to an N: P ratio lower than the 2N counterpart (Van Mooy et al., 2009; Carini et al., 2015).

Whilst total BLL quantity per cell was generally similar (**Figure 3A**), BLL(22:6/22:6) was present only at trace level in the uninfected 2N cultures and increased dramatically in abundance in the infected 2N cultures. BLL(18:1/22:6) was present in the uninfected 2N and uninfected 1N cultures and we found that the ratio between BLL(22:6/22:6)/BLL(18:1/22:6) was a strong indicator of infection in these experiments (**Figure 3B**). We suggest that in light of these and previous observations (Fulton et al., 2014), that BLL(22:6/22:6) and ratios thereof may be a useful biomarker of EhV infection to compliment vGSL. Application as a biomarker would require further investigation to verify BLL(22:6/22:6) presence/absence under the infected/host states is conserved across different *E. huxleyi* and viral strains and specific to *E. huxleyi*. The presence/absence relationship has been to date demonstrated in *E. huxleyi* 1216 (2N), 1217 (1N), and 1216 with EhV201 in this study and *E. huxleyi* 374 (2N) with EhV86 (Fulton et al., 2014).

Notably, the polar glycerolipids PC, DGCC, DGDG, SQDG, and PDPT were effectively identical in all three cultures through time (Supplementary Figure 1). This observation, in conjunction with the many similarities in PG, PE, and MGDG indicates that the overall polar glycerolipid content is not radically affected by life cycle or even viral infection, except potentially at the very termination of the lytic phase; only DGTS and BLL (22:6/22:6) show notable earlier variations. Bell and Pond (1996) observed that the distributions of fatty acids within polar glycerolipids were also fairly similar in uninfected 2N and 1N cells. Thus, overall these molecules seem to play a largely passive role in life cycle differentiation and infection. Furthermore, since the cellular lipid content of the major glycerolipids is mostly similar in all three cultures and cellular lipid content is a first-order approximation of cellular biomass, these data validate our assumption that the different types of cells do not vary greatly in size (i.e., biomass content) and support our decision to normalize all lipid data to cell abundance.

Glycosphingolipid Targeted Lipidomics

In contrast to the polar glycerolipids, the glycosphingolipids were highly dynamic, showing marked differences between culture treatments and through time. The infection marker vGSL was absent from the uninfected 2N cultures and abundant in the infected 2N cultures, rising through time with the progression of infection (**Figure 4A**), in agreement with previous observations (Vardi et al., 2009, 2012). Interestingly, low levels of vGSL-like molecules were detected in the uninfected 1N cultures. These vGSL-like species produce the same diagnostic glycosyl headgroup fragments and co-elute with previously characterized vGSLs. Untargeted lipidomic data, discussed below, suggest however that these molecules may not be exactly the same as the vGSLs in the infected diploid cultures originally recognized by Vardi et al. (2009). Future developments in targeted lipidomic methods hold the promise of being able to distinguish true vGSL from infected diploid cells and this novel vGSL-like molecule from uninfected haploid cells.

Host intrinsic hGSL did not vary in quantity per cell between the uninfected 2N control and the infected 2N cultures, in line with previous studies (Vardi et al., 2009, 2012; Fulton et al., 2014). Furthermore, we find that hGSL cellular abundance is also statistically similar in the uninfected 1N cultures (**Figure 4B**).

The marker lipid sGSL was recently described as indicative of susceptibility to viral infection in *E. huxleyi* (Fulton et al., 2014) and we found that sGSL was abundant in the virus-susceptible 2N strain used in this study, RCC1216. Levels of sGSL were somewhat reduced with the progression of infection in the infected 2N cultures at intermediate time, before returning to a level statistically similar to that of the uninfected control (**Figure 4C**). This contrasts with a previous study with a different *E. huxleyi* strain (Fulton et al., 2014) showing an increase in sGSL cellular abundance under infection. The same report however, included mesocosm field experiments that showed decreases in sGSL abundance per cell subject to EhV infection (Fulton et al., 2014); these results were interpreted as preferential infection of sGSL rich cells and relative growth of sGSL poor virus resistant cells. In light of our finding that sGSL is absent in 1N cells, the interpretation of Fulton et al. (2014) is entirely consistent with the observations of Frada et al. (2012) from the same mesocosms showing that 1N cells became proportionally more abundant, an affirmation of the Cheshire Cat hypothesis (Frada et al., 2008). Although the dynamics of sGSL subject to EhV infection appear variable between 2N systems (Fulton et al., 2014), the discovery here that sGSL is absent in 1N *E. huxleyi* cells could make sGSL an even more important biomarker for viral infection dynamics and associated shifts in ploidy.

Since 1N cells are resistant to viral infection (Frada et al., 2008) and sGSL is linked with susceptibility in several strains of 2N *E. huxleyi* (Fulton et al., 2014), the absence of sGSL from 1N cells further implicates sGSL as playing a key role in viral infection (Fulton et al., 2014). Sialic acid moieties similar to that of sGSL are implicated as ligands for viral attachment in the cyanobacterium *Prochlorococcus* (Avrani et al., 2011; Fulton et al., 2014). It has been speculated therefore, that viral attachment and entry, facilitated by sGSL, may be the means by which this resistance is achieved (Fulton et al., 2014). Furthermore, it has been speculated that sGSL may have intrinsic function in the production or externalization of the coccoliths that characterize the exterior of the calcifying 2N phase (Fulton et al., 2014). It is noteworthy therefore, that the 1N phase that does not produce sGSL also does not produce coccoliths, further suggesting a link between sGSL and calcification.

Untargeted Lipidomics and Biomarker Selection

By using PLS-DA to analyse the untargeted lipidomic data, we have highlighted molecular species diagnostic of each 1N, 2N, and infected 2N cells (**Table 1**). These species were assigned identities based upon matches between accurate m/z and an extensive database of lipids. The included assignments are supported by secondary MS^2 data. Due to the nature of the all ion fragmentation MS^2 method applied in the interest of unbiased methodology, incorrect assignments of co-eluting species may occur in some cases. As such the assigned lipid identities in the unbiased analyses are presented as tentative. The differential abundances of each species give a quantitative representation for the discovery of potential biomarkers.

For the uninfected 2N control, the known susceptibility biomarker sGSL (Fulton et al., 2014) and related species were the top ranked assigned hits, in both positive and negative ion mode. This observation is in agreement with the targeted data, where sGSL was absent from the uninfected 1N cultures, and reduced in abundance in the infected 2N case. In negative ion mode, BLL(18:1/22:6) was ranked 3rd for the uninfected 2N cultures. BLL(18:1/22:6) is a known marker of *E. huxleyi*, whose concentration has been shown to drop under viral infection (Fulton et al., 2014). At the present time, sGSL and BLL (18:1/22:6) have not been reported in any other species than *E. huxleyi*. By contrast, PC(34:5), MGDG(36:9), and SQDG(40:8) which were ranked 5th in positive, 4th and 5th in negative ion mode respectively, are common polar membrane glycerolipids in eukaryotic microalgae (Popendorf et al., 2011a; Brandsma et al., 2012). These observations underscore the challenge of developing these biomarkers for use in the field: a molecule must be both diagnostic for *E. huxleyi* viral infection and ploidy and absent in all other marine microbes. The former criterion is addressed with the data presented here, but the latter will only be met after the lipidomes of many additional marine taxa are examined using the techniques similar to those employed here. In the meantime, the biomarkers we identified by untargeted lipidomics should be applied with due discretion.

In the uninfected 1N cultures, the top hits in positive ion mode appeared to be a form of 31:0 wax ester or alkenoate. It is possible that changes in this molecule could reflect differences in the amounts of long chain alkenones in 1N cells. Alkenones are widely used in palaeoceanography temperature reconstructions (Volkman et al., 1980; Prahl and Wakeham, 1987). Direct comparisons between this 31:0 wax ester/alkenoate and alkenones remain to be conducted.

The 3rd and 4th ranked species in 1 N cultures in positive ion mode were matched to a GSL(t40:0) and GSL(t40:1) species. These GSLs appear to share some structural similarities with archetypal forms of hGSL (GSL(d41:5(OH)) and vGSL (GSL(t39:0(OH)) identified by Vardi et al. (2009, 2012). These haploid specific GSLs may be the source of the observed vGSL-like species in the targeted uninfected haploid analyses, discussed previously. In light of the aforementioned targeted GSL data, these haploid GSLs certainly warrant further characterization. In negative ion mode, Cer(d18:1/22:0(OH)) and Cer(d18:1/22:1(OH)) hydroxyceramide species were ranked 2nd and 3rd for the uninfected 1N cultures. These ceramides are effectively the same chemical composition as GSL minus the glycosyl headgroup. Furthermore, it is likely that the GSL(t40:1) highlighted is in fact the Cer(d18:1/22:0(OH)) ceramide moiety with a glycosyl headgroup, although this could not be confirmed from the MS2 data. Interestingly, upregulation of enzymes responsible for the production of ceramides from the hydrolysis of the beta-glycosidic linkages of GSLs is a characteristic response of 2N cells to viral infection (Rosenwasser et al., 2014). To our knowledge, the GSL(t40:0) and GSL(t40:1) species have not been previously reported in any marine microbial organism,

although they are closely related chemically to the other GSLs. Thus, these GSL species represent, pending further validation and characterization, candidate biomarkers for 1N *E. huxleyi*. 1N *E. huxleyi* are not easily identifiable in field samples by classic microscopy techniques (Von Dassow et al., 2009) and lipid biomarkers have particular utility in bulk environmental samples, such as filtered particulate organic material collected from the water column or sinking particulate trap matter, where other methods may not be applicable.

Finally, a number of glyceroglycolipids appeared indicative of the uninfected 1N cultures. MGDG(32:5) was ranked 5th in positive ion mode and LMGDG(22:6), DGDG(32:5) and DGDG(30:6) were ranked 1st, 4th, and 5th in negative ion mode. Glyceroglycolipids are typically associated with the thylakoid membranes of the chloroplast (Petroutsos et al., 2014). Our analysis highlights differences in these photosynthetic membrane lipids and alludes to some implication of ploidy upon the photosynthetic apparatus. However, these are fairly common molecules that are unlikely to be of much utility as biomarkers for *E. huxleyi* viral infection or ploidy (Popendorf et al., 2011a; Brandsma et al., 2012).

For the infected 2N cultures, the PLS-DA model highlights a series of TAGs found under positive ionization as indicative. TAG biosynthesis is commonly upregulated in phytoplankton in response to a variety of stressors, and TAGs are found primarily in lipid bodies. TAGs are unlikely to be attributable to heterotrophic bacterial contaminants as they are not generally produced in significant quantities by bacteria (Alvarez and Steinbüchel, 2002). It has been proposed that EhV capsid assembly occurs via a mechanism similar to that of the Hepatitis C virus (Herker and Ott, 2011; Fulton et al., 2014), where lipids are incorporated from the interface of lipid bodies and the endoplasmic reticulum. In uninfected 2N *E. huxleyi*, these lipid bodies are composed mainly of alkenones, alkenoates, and small quantities of TAGs and other lipids (Eltgroth et al., 2005). Research on human hepatocyte cells shows that TAGs are critical to Hepatitis C capsid assembly (Liefhebber et al., 2014). Thus, we speculate that the observed increase in a number of TAG species in *E. huxleyi* under EhV infection may result from the virus upregulating TAG biosynthesis in order to assist capsid assembly.

In negative ion mode, vGSL and BLL(22:6/22:6) represent the top four assigned species indicative of the infected cultures, in line with previous observations (Vardi et al., 2012; Fulton et al., 2014) and the targeted data discussed previously, while a 36:2 wax ester/alkenoate species is ranked fifth. None of these molecules are likely to be derived from contaminating bacteria in the cultures, but instead are of decidedly coccolithophore origin (Eltgroth et al., 2005; Fulton et al., 2014; Ray et al., 2014). A molecule with the exact same elemental formula, 36:2 methyl alkenoate, was identified in *E. huxleyi* lipid bodies (Eltgroth et al., 2005). Closely related alkenones are abundant in EhV virions and alkenone content in infected *E. huxleyi* has previously been demonstrated to be increased (Fulton et al., 2014). As mentioned previously, alkenoates and alkenones are thought to occur primarily in lipid bodies (Eltgroth et al., 2005). We suggest therefore, that this 36:2 wax ester/alkenoate species we observed was likely localized to lipid bodies. If true, this result, in conjunction with the TAG data mentioned above, further implicates the role of lipid bodies in viral assembly (Fulton et al., 2014).

In addition to the assigned species discussed previously, many novel biomarkers could not be assigned from our database and/or corroborated by all ion fragmentation MS^2 (Supplementary Table 2). Many of these species demonstrate absence/presence behavior between cultures and have potential as biomarkers of EhV infection or ploidy in *E. huxleyi*. The assigned and unassigned candidate biomarkers discovered by these unbiased analyses are tantalizing. A further biomarker validation study, utilizing targeted MS^2/MS^3 fragmentation in conjunction with high resolution mass spectrometry has the potential to assign and confirm the identities of the additional molecular species.

CONCLUSIONS

We have presented critical new data on the glycerolipids and glycosphingolipids of 1N *E. huxleyi* by the application of a targeted lipidomics approach. Comparison of these data to that of infected and uninfected 2N *E. huxleyi* has provided new insights on ploidy and viral infection. Firstly, we note the detection of trace levels of the EhV type viral glycosphingolipid (vGSL) in the uninfected 1N samples. The implications of this observation are as yet unclear but should be considered in the application of vGSL as a biomarker of EhV infection. Secondly, we find that the sialic glycosphingolipid (sGSL), a proposed marker of susceptibility to EhV infection in *E. huxleyi,* was absent from the uninfected 1N cell. This absence provides further evidence for the role of sGSL in EhV infection and may confer 1N *E. huxleyi* its documented resistance to EhV.

In addition, we have highlighted promising lipid biomarker candidates for each of the uninfected 2N, 1N, and infected 2N cases by way of untargeted lipidomics. Differentially enriched biomarker candidates have been tentatively identified, such as novel glycosphingolipids, hydroxyceramides and a wax ester/alkanoate that are highly indicative of the 1N phase. Following further biomarker validation and structural studies these lipids may yield powerful lipid biomarkers for the determination of 1N *E. huxleyi*, identification of which is not feasible by classic microscopy techniques.

These findings contribute to our understanding of the critical role of lipids in *E. huxleyi*/EhV interactions. Moreover, our findings further the potential of lipid based biomarkers as indicators of the progression of infection and life cycle in *E. huxleyi*. Extrapolation from simplified cell culture models to complex environmental systems must be approached with due caution and the discussed biomarkers require further validation in culture and *in situ*. However, such biomarkers have the potential to yield great insight into the processes that dictate the characteristics of *E. huxleyi* blooms, translating to substantial implications for global carbon cycling and climate.

ACKNOWLEDGMENTS

The authors would like to thank J. Tagliafere and J. Ossolinski for assistance in the laboratory and J. Collins, J. Fulton, and B. Edwards for helpful discussions about data analysis. This work was funded by the University of Southampton—Vice Chancellor's

Scholarship Award (JH), University of Southampton—Diamond Jubilee Fellowship (BV), Graduate School of the National Oceanography Centre WHOI Exchange Award (JH), and the Gordon and Betty Moore Foundation through Grant GBMF3301 (BV and AV).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmars.2015.00081

REFERENCES

Alvarez, H., and Steinbüchel, A. (2002). Triacylglycerols in prokaryotic microorganisms. *Appl. Microbiol. Biotechnol.* 60, 367–376. doi: 10.1007/s00253-002-1135-0

Avrani, S., Wurtzel, O., Sharon, I., Sorek, R., and Lindell, D. (2011). Genomic island variability facilitates Prochlorococcus-virus coexistence. *Nature* 474, 604–608. doi: 10.1038/nature10172

Ballabio, D., and Consonni, V. (2013). Classification tools in chemistry. Part 1: linear models. PLS-DA. *Anal. Methods* 5, 3790–3798. doi: 10.1039/C3AY40582F

Bell, M. V., and Pond, D. (1996). Lipid composition during growth of motile and coccolith forms of *Emiliania huxleyi*. *Phytochemistry* 41, 465–471. doi: 10.1016/0031-9422(95)00663-X

Bligh, E. G., and Dyer, W. J. (1959). A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol.* 37, 911–917. doi: 10.1139/o59-099

Borchard, C., and Engle, A. (2012). Organic matter exudation by *Emiliania huxleyi* under simulated future ocean conditions. *Biogeosciences* 9, 3405–3423. doi: 10.5194/bg-9-3405-2012

Brandsma, J., Hopmans, E. C., Philippart, C. J. M., Veldhuis, M. J. W., Schouten, S., and Sinninghe Damsté, J. S. (2012). Low temporal variation in the intact polar lipid composition of North Sea coastal marine water reveals limited chemotaxonomic value. *Biogeosciences* 9, 1073–1084. doi: 10.5194/bg-9-1073-2012

Bratbak, G., Egge, J. K., and Heldal, M. (1993). Viral mortality of the marine alga *Emiliania huxleyi* (Haptophyceae) and termination of algal blooms. *Mar. Ecol. Prog. Ser.* 93, 39–48. doi: 10.3354/meps093039

Brussaard, C., Gast, G., van Duyl, F., and Riegman, R. (1996). Impact of phytoplankton bloom magnitude on a pelagic microbial food web. *Mar. Ecol. Prog. Ser.* 144, 211–221. doi: 10.3354/meps144211

Carini, P., Van Mooy, B. A. S., Thrash, J. C., White, A., Zhao, Y., Campbell, E. O., et al. (2015). SAR11 lipid renovation in response to phosphate starvation. *Proc. Natl. Acad. Sci. U.S.A.* 112, 7767–7772. doi: 10.1073/pnas.1505034112

Eltgroth, M. L., Watwood, R. L., and Wolfe, G. V. (2005). Production and cellular localization of neutral long-chain lipids in the haptophyte algae isochrysis galbana and *Emiliania huxleyi*. *J. Phycol.* 41, 1000–1009. doi: 10.1111/j.1529-8817.2005.00128.x

Frada, M. J., Bidle, K. D., Probert, I., and de Vargas, C. (2012). *In situ* survey of life cycle phases of the coccolithophore *Emiliania huxleyi* (Haptophyta). *Environ. Microbiol.* 14, 1558–1569. doi: 10.1111/j.1462-2920.2012.02745.x

Frada, M., Probert, I., Allen, M. J., Wilson, W. H., and de Vargas, C. (2008). The "Cheshire Cat" escape strategy of the coccolithophore *Emiliania huxleyi* in response to viral infection. *Proc. Natl. Acad. Sci. U.S.A.* 105, 15944–15949. doi: 10.1073/pnas.0807707105

Fuhrman, J. A. (1999). Marine viruses and their biogeochemical and ecological effects. *Nature* 399, 541–548. doi: 10.1038/21119

Fulton, J. M., Fredricks, H. F., Bidle, K. D., Vardi, A., Kendrick, B. J., DiTullio, G. R., et al. (2014). Novel molecular determinants of viral susceptibility and resistance in the lipidome of *Emiliania huxleyi*. *Environ. Microbiol.* 16, 1137–1149. doi: 10.1111/1462-2920.12358

Green, J. C., Course, P. A., and Tarran, G. A. (1996). The life-cycle of *Emiliania huxleyi*: a brief review and a study of relative ploidy levels analysed by flow cytometry. *J. Mar. Syst.* 9, 33–44. doi: 10.1016/0924-7963(96)00014-0

Gundersen, K., Heldal, M., Norland, S., Purdie, D. A., and Knap, A. H. (2002). Elemental C, N, and P content of individual bacteria collected at the Bermuda Atlantic Time-Series (BATS) site. *Limnol. Oceanogr.* 47, 1525–1530. doi: 10.4319/lo.2002.47.5.1525

Herker, E., and Ott, M. (2011). Unique ties between hepatitis C virus replication and intracellular lipids. *Trends Endocrinol. Metab.* 22, 241–248. doi: 10.1016/j.tem.2011.03.004

Holligan, P. M., Viollier, M., Harbour, D. S., Camus, P., and Champagne-Philippe, M. (1983). Satellite and ship studies of coccolithophore production along a continental shelf edge. *Nature* 304, 339–342. doi: 10.1038/304339a0

Houdan, A., Billard, C., Marie, D., Not, F., Sáez, A. G., Young, J. R., et al. (2004). Holococcolithophore—heterococcolithophore (Haptophyta) life cycles: flow cytometric analysis of relative ploidy levels. *Syst. Biodivers.* 1, 453–465. doi: 10.1017/S1477200003001270

Houdan, A., Probert, I., Van Lenning, K., and Lefebvre, S. (2005). Comparison of photosynthetic responses in diploid and haploid life-cycle phases of *Emiliania huxleyi* (Prymnesiophyceae). *Mar. Ecol. Prog. Ser.* 292, 139–146. doi: 10.3354/meps292139

Huang, N., Siegel, M. M., Kruppa, G. H., and Laukien, F. H. (1999). Automation of a Fourier transform ion cyclotron resonance mass spectrometer for acquisition, analysis, and e-mailing of high-resolution exact-mass electrospray ionization mass spectral data. *J. Am. Soc. Mass Spectrom.* 10, 1166–1173. doi: 10.1016/S1044-0305(99)00089-6

Hummel, J., Segu, S., Li, Y., Irgang, S., Jueppner, J., and Giavalisco, P. (2011). Ultra performance liquid chromatography and high resolution mass spectrometry for the analysis of plant lipids. *Front. Plant Sci.* 2:54. doi: 10.3389/fpls.2011.00054

Jover, L. F., Effler, T. C., Buchan, A., Wilhelm, S. W., and Weitz, J. S. (2014). The elemental composition of virus particles: implications for marine biogeochemical cycles. *Nat. Rev. Micro.* 12, 519–528. doi: 10.1038/nrmicro3289

Keller, M. D., Selvin, R. C., Claus, W., and Guillard, R. R. L. (1987). Media for the culture of oceanic ultraphytoplankton. *J. Phycol.* 23, 633–638. doi: 10.1111/j.1529-8817.1987.tb04217.x

Klaveness, D. (1972). Coccolithus huxleyi (Lohm.) Kamptn II. The flagellate cell, aberrant cell types, vegetative propagation and life cycles. *Br. Phycol. J.* 7, 309–318. doi: 10.1080/00071617200650321

Lehahn, Y., Koren, I., Schatz, D., Frada, M., Sheyn, U., Boss, E., et al. (2014). Decoupling physical from biological processes to assess the impact of viruses on a mesoscale algal bloom. *Curr. Biol.* 24, 2041–2046. doi: 10.1016/j.cub.2014.07.046

Liefhebber, J. M. P., Hague, C. V., Zhang, Q., Wakelam, M. J. O., and McLauchlan, J. (2014). Modulation of triglyceride and cholesterol ester synthesis impairs assembly of infectious Hepatitis C virus. *J. Biol. Chem.* 289, 21276–21288. doi: 10.1074/jbc.M114.582999

Mackinder, L. C. M., Worthy, C. A., Biggi, G., Hall, M., Ryan, K. P., Varsani, A., et al. (2009). A unicellular algal virus, *Emiliania huxleyi* virus 86, exploits an animal-like infection strategy. *J. Gen. Virol.* 90, 2306–2316. doi: 10.1099/vir.0.011635-0

Martin, P., Van Mooy, B. A. S., Heithoff, A., and Dyhrman, S. T. (2011). Phosphorus supply drives rapid turnover of membrane phospholipids in the diatom *Thalassiosira pseudonana*. *ISME J.* 5, 1057–1060. doi: 10.1038/ismej.2010.192

Mausz, M. A., and Pohnert, G. (2014). Phenotypic diversity of diploid and haploid *Emiliania huxleyi* cells and of cells in different growth phases revealed by comparative metabolomics. *J. Plant Physiol.* 172, 137–148. doi: 10.1016/j.jplph.2014.05.014

Monier, A., Pagarete, A., de Vargas, C., Allen, M. J., Read, B., Claverie, J. -M., et al. (2009). Horizontal gene transfer of an entire metabolic pathway between a eukaryotic alga and its DNA virus. *Genome Res.* 19, 1441–1449. doi: 10.1101/gr.091686.109

Paasche, E. (2001). A review of the coccolithophorid *Emiliania huxleyi* (Prymnesiophyceae), with particular reference to growth, coccolith formation,

and calcification-photosynthesis interactions. *Phycologia* 40, 503–529. doi: 10.2216/i0031-8884-40-6-503.1

Petroutsos, D., Amiar, S., Abida, H., Dolch, L.-J., Bastien, O., Rébeillé, F., et al. (2014). Evolution of galactoglycerolipid biosynthetic pathways—From cyanobacteria to primary plastids and from primary to secondary plastids. *Prog. Lipid Res.* 54, 68–85. doi: 10.1016/j.plipres.2014.02.001

Popendorf, K. J., Fredricks, H. F., and Van Mooy, B. A. S. (2013). Molecular ion-independent quantification of polar glycerolipid classes in marine plankton using triple quadrupole MS. *Lipids* 48, 185–195. doi: 10.1007/s11745-012-3748-0

Popendorf, K. J., Lomas, M. W., and Van Mooy, B. A. S. (2011a). Microbial sources of intact polar diacylglycerolipids in the Western North Atlantic Ocean. *Org. Geochem.* 42, 803–811. doi: 10.1016/j.orggeochem.2011.05.003

Popendorf, K. J., Tanaka, T., Pujo-Pay, M., Lagaria, A., Courties, C., Conan, P., et al. (2011b). Gradients in intact polar diacylglycerolipids across the Mediterranean Sea are related to phosphate availability. *Biogeosciences* 8, 3733–3745. doi: 10.5194/bg-8-3733-2011

Prahl, F. G., and Wakeham, S. G. (1987). Calibration of unsaturation patterns in long-chain ketone compositions for palaeotemperature assessment. *Nature* 330, 367–369. doi: 10.1038/330367a0

Ray, J. L., Haramaty, L., Thyrhaug, R., Fredricks, H. F., Van Mooy, B. A. S., Larsen, A., et al. (2014). Virus infection of *Haptolina ericina* and Phaeocystis pouchetii implicates evolutionary conservation of programmed cell death induction in marine haptophyte–virus interactions. *J. Plankton Res.* 36, 943–955. doi: 10.1093/plankt/fbu029

Rokitta, S. D., de Nooijer, L. J., Trimborn, S., de Vargas, C., Rost, B., and John, U. (2011). Transcriptome analyses reveal differential gene expression patterns between the life-cycle stages of *Emiliania huxleyi* (Haptophyta) and reflect specialization to different ecological niches. *J. Phycol.* 47, 829–838. doi: 10.1111/j.1529-8817.2011.01014.x

Rose, S. L., Fulton, J. M., Brown, C. M., Natale, F., Van Mooy, B. A. S., and Bidle, K. D. (2014). Isolation and characterization of lipid rafts in *Emiliania huxleyi*: a role for membrane microdomains in host–virus interactions. *Environ. Microbiol.* 16, 1150–1166. doi: 10.1111/1462-2920.12357

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. doi: 10.1105/tpc.114.125641

Sakurai, I., Shen, J.-R., Leng, J., Ohashi, S., Kobayashi, M., and Wada, H. (2006). Lipids in oxygen-evolving photosystem II complexes of cyanobacteria and higher plants. *J. Biochem.* 140, 201–209. doi: 10.1093/jb/mvj141

Schatz, D., Shemi, A., Rosenwasser, S., Sabanay, H., Wolf, S. G., Ben-Dor, S., et al. (2014). Hijacking of an autophagy-like process is critical for the life cycle of a DNA virus infecting oceanic algal blooms. *New Phytol.* 204, 854–863. doi: 10.1111/nph.13008

Schroeder, D. C., Oke, J., Malin, G., and Wilson, W. H. (2002). Coccolithovirus (Phycodnaviridae): characterisation of a new large dsDNA algal virus that infects Emiliana huxleyi. *Arch. Virol.* 147, 1685–1698. doi: 10.1007/s00705-002-0841-3

Shemi, A., Ben-Dor, S., and Vardi, A. (2015). Elucidating the composition and conservation of the autophagy pathway in photosynthetic eukaryotes. *Autophagy* 11, 701–715. doi: 10.1080/15548627.2015.1034407

Stray, S. J., Cummings, R. D., and Air, G. M. (2000). Influenza virus infection of desialylated cells. *Glycobiology* 10, 649–658. doi: 10.1093/glycob/10.7.649

Sud, M., Fahy, E., Cotter, D., Brown, A., Dennis, E. A., Glass, C. K., et al. (2007). LMSD: LIPID MAPS structure database. *Nucleic Acids Res.* 35, D527–D532. doi: 10.1093/nar/gkl838

Suttle, C. A. (2007). Marine viruses—major players in the global ecosystem. *Nat. Rev. Micro.* 5, 801–812. doi: 10.1038/nrmicro1750

Tyrrell, T., and Merico, A. (2004). "*Emiliania huxleyi*: bloom observations and the conditions that induce them," in *Coccolithophores SE—4*, eds H. R. Thierstein and J. R. Young (Berlin; Heidelberg: Springer), 75–97.

Van den Berg, R. A., Hoefsloot, H. C. J., Westerhuis, J. A., Smilde, A. K., and van der Werf, M. J. (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics* 7:142. doi: 10.1186/1471-2164-7-142

Van Etten, J. L., Graves, M. V., Müller, D. G., Boland, W., and Delaroque, N. (2002). Phycodnaviridae– large DNA algal viruses. *Arch. Virol.* 147, 1479–1516. doi: 10.1007/s00705-002-0822-6

Van Mooy, B. A. S., Fredricks, H. F., Pedler, B. E., Dyhrman, S. T., Karl, D. M., Koblížek, M., et al. (2009). Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity. *Nature* 458, 69–72. doi: 10.1038/nature07659

Vardi, A., Haramaty, L., Van Mooy, B. A. S., Fredricks, H. F., Kimmance, S. A., Larsen, A., et al. (2012). Host–virus dynamics and subcellular controls of cell fate in a natural coccolithophore population. *Proc. Natl. Acad. Sci. U.S.A.* 109, 19327–19332. doi: 10.1073/pnas.1208895109

Vardi, A., Van Mooy, B. A. S., Fredricks, H. F., Popendorf, K. J., Ossolinski, J. E., Haramaty, L., et al. (2009). Viral glycosphingolipids induce lytic infection and cell death in marine phytoplankton. *Science* 326, 861–865. doi: 10.1126/science.1177322

Volkman, J. K., Eglinton, G., Corner, E. D. S., and Sargent, J. R. (1980). Novel unsaturated straight-chain C37-C39 methyl and ethyl ketones in marine sediments and a coccolithophore *Emiliania huxleyi*. *Phys. Chem. Earth* 12, 219–227. doi: 10.1016/0079-1946(79)90106-x

Von Dassow, P., Ogata, H., Probert, I., Wincker, P., Da Silva, C., Audic, S., et al. (2009). Transcriptome analysis of functional differentiation between haploid and diploid cells of *Emiliania huxleyi*, a globally significant photosynthetic calcifying cell. *Genome Biol.* 10, 1–33. doi: 10.1186/gb-2009-10-10-r114

Weber, R. J. M., Li, E., Bruty, J., He, S., and Viant, M. R. (2012). MaConDa: a publicly accessible mass spectrometry contaminants database. *Bioinformatics* 28, 2856–2857. doi: 10.1093/bioinformatics/bts527

Westbroek, P., Brown, C. W., van Bleijswijk, J., Brownlee, C., Brummer, G. J., Conte, M., et al. (1993). A model system approach to biological climate forcing. The example of *Emiliania huxleyi*. *Glob. Planet. Change* 8, 27–46. doi: 10.1016/0921-8181(93)90061-R

Wilson, W. H., Schroeder, D. C., Allen, M. J., Holden, M. T. G., Parkhill, J., Barrell, B. G., et al. (2005). Complete genome sequence and lytic phase transcription profile of a coccolithovirus. *Science* 309, 1090–1092. doi: 10.1126/science.1113109

Wilson, W. H., Tarran, G. A., Schroeder, D., Cox, M., Oke, J., and Malin, G. (2002). Isolation of viruses responsible for the demise of an *Emiliania huxleyi* bloom in the English Channel. *J. Mar. Biol. Assoc. U.K.* 82, 369–377. doi: 10.1017/S002531540200560X

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

7.5 References

Fulton, J. M., Fredricks, H. F., Bidle, K. D., Vardi, A., Kendrick, B. J., DiTullio, G. R., et al. (2014), Novel molecular determinants of viral susceptibility and resistance in the lipidome of Emiliania huxleyi, *Environmental Microbiology*

Vardi, A., Haramaty, L., Van Mooy, B. A. S., Fredricks, H. F., Kimmance, S. A., Larsen, A., et al. (2012), Hostvirus dynamics and subcellular controls of cell fate in a natural coccolithophore population, *Proceedings of the National Academy of Sciences*

Vardi, A., Van Mooy, B. A. S., Fredricks, H. F., Popendorf, K. J., Ossolinski, J. E., Haramaty, L., et al. (2009), Viral glycosphingolipids induce lytic infection and cell death in marine phytoplankton., *Science*