

Article

Mass Spectrometry Data Repository Enhances Novel Metabolite Discoveries with Advances in Computational Metabolomics

Hiroshi Tsugawa ^{1,2,*}, Aya Satoh ^{1,†}, Haruki Uchino ^{2,3}, Tomas Cajka ^{4,5}, Makoto Arita ^{2,3,6} and Masanori Arita ^{1,7,*}

¹ Metabolome informatics research team, RIKEN Center for Sustainable Resource Science, Yokohama 230-0045, Japan; aya.hayaishi@riken.jp

² Laboratory for metabolomics, RIKEN Center for Integrative Medical Sciences, Yokohama 230-0045, Japan; haruki-uchino@keio.jp (H.U.); makoto.arita@riken.jp (M.A.)

³ Division of Physiological Chemistry and Metabolism, Graduate School of Pharmaceutical Sciences, Keio University, Minato-ku, Tokyo 105-8512, Japan

⁴ Department of Metabolomics, Institute of Physiology of the Czech Academy of Sciences, Videnska 1083, 14220 Prague, Czech Republic; tomas.cajka@fgu.cas.cz

⁵ Department of Translational Metabolism, Institute of Physiology of the Czech Academy of Sciences, Videnska 1083, 14220 Prague, Czech Republic

⁶ Cellular and Molecular Epigenetics Laboratory, Graduate School of Medical Life Science, Yokohama City University, Tsurumi, Yokohama 230-0045, Japan

⁷ National Institute of Genetics, Mishima 411-8540, Japan

* Correspondence: hiroshi.tsugawa@riken.jp (H.T.); arita@nig.ac.jp (M.A.); Tel.: +81-45-503-9491 (H.T.); +81-45-503-9491 (M.A.)

† These authors contributed equally to this work.

Received: 20 May 2019; Accepted: 19 June 2019; Published: 24 June 2019



Abstract: Mass spectrometry raw data repositories, including Metabolomics Workbench and MetaboLights, have contributed to increased transparency in metabolomics studies and the discovery of novel insights in biology by reanalysis with updated computational metabolomics tools. Herein, we reanalyzed the previously published lipidomics data from nine algal species, resulting in the annotation of 1437 lipids achieving a 40% increase in annotation compared to the previous results. Specifically, diacylglyceryl-carboxyhydroxy-methylcholine (DGCC) in *Paolova lutheri* and *Pleurochrysis carterae*, glucuronosyldiacylglycerol (GlcADG) in *Euglena gracilis*, and *P. carterae*, phosphatidylmethanol (PMeOH) in *E. gracilis*, and several oxidized phospholipids (oxidized phosphatidylcholine, OxPC; phosphatidylethanolamine, OxPE; phosphatidylglycerol, OxPG; phosphatidylinositol, OxPI) in *Chlorella variabilis* were newly characterized with the enriched lipid spectral databases. Moreover, we integrated the data from untargeted and targeted analyses from data independent tandem mass spectrometry (DIA-MS/MS) acquisition, specifically the sequential window acquisition of all theoretical fragment-ion MS/MS (SWATH-MS/MS) spectra, to increase the lipidomic annotation coverage. After the creation of a global library of precursor and diagnostic ions of lipids by the MS-DIAL untargeted analysis, the co-eluted DIA-MS/MS spectra were resolved in MRMPROBS targeted analysis by tracing the specific product ions involved in acyl chain compositions. Our results indicated that the metabolite quantifications based on DIA-MS/MS chromatograms were somewhat inferior to the MS¹-centric quantifications, while the annotation coverage outperformed those of the untargeted analysis of the data dependent and DIA-MS/MS data. Consequently, integrated analyses of untargeted and targeted approaches are necessary to extract the maximum amount of metabolome information, and our results showcase the value of data repositories for the discovery of novel insights in lipid biology.

Keywords: data repository; computational metabolomics; reanalysis; lipidomics; data processing

1. Introduction

Many studies using mass spectrometry (MS)-based untargeted metabolomics have provided novel insights in biology, and the importance of metabolomics data repositories has been recognized [1]. In addition to the international data repositories, including Metabolomics Workbench [2], MetaboLights [3], and GNPS MassIVE [4], institute-oriented repositories such as RIKEN DropMet (<http://prime.psc.riken.jp/>) are available for sharing raw MS data. MS repositories aim to (A) increase the transparency and reproducibility of MS-centric metabolomics studies, (B) provide a benchmark for testing new analytical and computational methodologies, (C) share the results of metabolome analyses for providing opportunities for data-driven hypothesis generation, and (D) reanalyze published data with continuous identification efforts to obtain novel insights and discover novel metabolites. Nevertheless, few studies have demonstrated the value of MS repositories except for the first purpose. Case studies showing remarkable results toward the other three purposes would facilitate data sharing by researchers and academic journals [5].

We published nine algal lipidomics datasets in 2015 which are available on the RIKEN DropMet. According to previous reports, 1023 lipids were annotated by integrating the results of data-dependent MS/MS acquisition (DDA-MS/MS) and data-independent MS/MS acquisition (DIA-MS/MS) with an in-silico MS/MS spectral library of lipids to assign the algal phylogenetic tree based on lipid properties [6]. Although comprehensive lipid analysis has been achieved by these previous methodologies, the coverage of algal lipids can be improved by two major methodological updates. First, the count of annotated lipids can be increased by updating the in-silico spectral library with continuous data curation efforts, where 1,051,894 spectra of 525,947 molecules from 90 lipid classes are currently registered in MS-DIAL (version 3.68), while 122,844 spectra of 61,422 molecules from 24 lipid classes were registered in the first version of MS-DIAL (version 1.82) [6–8]. Herein, the characterization of 10 newly incorporated lipid classes including lysophosphatidylserine (LPS), lysophosphatidylglycerol (LPG), phosphatidylmethanol (PMeOH), glucuronosyldiacylglycerol (GlcADG), diacylglyceryl-carboxyhydroxy-methylcholine (DGCC), and its lyso-type form (LDGCC), as well as oxidized fatty acids containing phosphatidylcholine (OxPC), phosphatidylethanolamine (OxPE) phosphatidylglycerol (OxPG), and phosphatidylinositol (OxPI) are highlighted for algal lipid profiling. Second, the peak capacity in LC-MS/MS data could be increased by the integrated analyses of MS-DIAL-based untargeted [6,9] and MRMPROBS-based targeted analyses [10,11] for DIA-MS/MS data. In DIA-MS/MS data, MS-DIAL uses the MS¹ chromatogram trace for metabolite quantification and the deconvoluted MS/MS spectrum for metabolite annotation. Therefore, the program does not completely resolve co-eluted metabolites, as only a singlet MS¹ chromatogram peak from mixed metabolite ions can be obtained [12]. On the other hand, MRMPROBS can use either the MS¹ or MS/MS chromatogram for metabolite quantification and metabolite diagnostics is performed by the integrated score of the peak groups from the user-defined precursor-product transitions library [10]. Notably, co-eluted metabolites in the MS¹ chromatogram trace can be resolved using the MS/MS chromatograms which differ based on the unique product ions from the lipid structure, resulting in increased peak capacity, i.e., increased deconvolution efficiency, in the LC-MS/MS dataset.

Herein, we showcase novel lipid discoveries using algal lipidomics data as a benchmark. We discovered novel lipid classes which have never been reported in the algal species using updated in-silico MS/MS libraries. Moreover, we demonstrate increased annotated lipid coverage by integrating the pipelines of the MS-DIAL and MRMPROBS programs. Although the integrated pipeline can be executed using other state-of-the-art program combinations, such as XCMS [13] and MZmine 2 [14], for untargeted analysis and MetDIA for targeted analysis [15], the selected programs support the direct link from untargeted to targeted analyses with a user-friendly graphical user interface (GUI) where manual

peak-picking required for targeted approaches is acceptable. This study highlights the importance of mass spectrometry data repositories to deepen our understanding of lipids in algal species.

2. Materials and Methods

2.1. Overview of Data Analysis Workflow

Since the MetaboLights database and repository was launched in 2012 by the European Bioinformatics Institute (EMBL-EBI) as the first repository for metabolomics data, data submission has continuously increased (~2.5 TB data was available in June 2019), and accessibility and awareness have been enhanced through the efforts of MetabolomeXchange (<http://www.metabolomexchange.org>) and OmicsDI [16]. RIKEN DropMet (http://prime.psc.riken.jp/menta.cgi/prime/drop_index) has also been launched in 2009 to share MS-based metabolomics data from RIKEN, in which ~300 GB of data from 29 studies are currently available; a part of this repository, i.e., the algae lipidomics data, was used in this study.

On the other hand, data processing tools like MS-DIAL [6,9], XCMS [13], and MZmine 2 [14] have continuously been updated with database curations like Metlin [17] in XCMS. Since the LipidBlast library was released in 2013 as the first public in-silico library for lipids [18], the fork libraries for quadrupole/time-of-flight mass spectrometry (QTOF-MS) with collision-induced dissociation (CID) and orbital ion trap MS with higher-energy collisional dissociation (HCD) data have been developed in MS-DIAL [6,7] owing to the continuous effort. The annotation described in this study can be executed in the MS-DIAL version 3.66 or higher. All programs, i.e., MS-DIAL, MRMPROBS, and the related spectral libraries are available on the RIKEN PRIME website (<http://prime.psc.riken.jp/>).

2.2. Mass Spectrometry Data

The DDA and DIA lipidomics data obtained in positive and negative ion modes of nine algal species were downloaded from the RIKEN DropMet website (http://prime.psc.riken.jp/menta.cgi/prime/drop_index; ID, DM0022). Briefly, the extraction of algal lipids was performed using a biphasic solvent system of cold methanol, methyl *tert*-butyl ether (MTBE), and water followed by lipid separation via reversed-phase liquid chromatography. Both DDA and DIA data were acquired using a QTOF mass spectrometer (TripleTOF 5600+, SCIEX). For DIA (SWATH-MS/MS), a 21 Da isolation window was used for selecting precursor ions shifting over an m/z 100–1250 mass range. Further details are provided in the previous study [6].

2.3. Software Programs

MS-DIAL version 3.06 and MRMPROBS version 2.44 were used herein. All programs including the latest version are freely available on the RIKEN PRIME website (<http://prime.psc.riken.jp/>).

The same parameters in MS-DIAL were used for DDA and DIA data analyses: retention time begin, 0 min; retention time end, 100 min; mass range begin, 0 Da; mass range end, 5000 Da; accurate mass tolerance (MS1), 0.01 Da; MS2 tolerance, 0.025 Da; maximum charge number, 2; smoothing method, linear weighted moving average; smoothing level, 3; minimum peak width, 5 scan; minimum peak height, 1000; mass slice width, 0.1 Da; sigma window value, 0.5; MS2Dec amplitude cut-off, 0; exclude after precursor, true; keep isotope until, 0.5 Da; keep original precursor isotopes, false; exclude after precursor, true; retention time tolerance for identification, 4 min; MS1 for identification, 0.01 Da; accurate mass tolerance (MS2) for identification, 0.05 Da; identification score cut-off, 70%; using retention time for scoring, true; relative abundance cut off, 0; top candidate report, true; retention time tolerance for alignment, 0.05 min; MS1 tolerance for alignment, 0.015 Da; peak count filter, 0; remove feature based on peak height fold-change, true; sample max/blank average, 5; keep identified and annotated metabolites, true; keep removable features and assign the tag for checking, true; replace true zero values with 1/10 of the minimum peak height over all samples, false. Lipid annotation was performed automatically using the in-silico MS/MS spectral library described below, and the result

was manually curated with the confirmation of the characteristic product ions and neutral losses to reduce false-positive annotations.

The parameters in MRMPROBS were set as follows: MS1 tolerance, 0.01 Da; MS2 tolerance, 0.025 Da; smoothing method, linear weighted moving average; smoothing level, 1; minimum peak width, 5 scan; minimum peak height, 200; retention time tolerance for identification, 0.1 min; amplitude tolerance for identification, 15%; minimum posterior, 70%; the abundance ratios in reference library were automatically generated by MS-DIAL. The results of metabolite annotation and peak picking were manually curated using the graphical user interface of MRMPROBS.

2.4. In-Silico MS/MS Spectral Libraries

The diagnostic ions used to characterize lipid classes were determined using authentic standards, experimental MS/MS spectra of biological samples, or MS/MS spectral information reported in the literature. The MS/MS spectra of PMeOH, LPS, and LPG were confirmed using the standard compounds PMeOH 16:0–16:0, LPG 18:1, and LPS 18:1 (Avanti Polar Lipids, Inc., Alabaster, AL, USA). The DGCC and LDGCC spectra were examined in the DDA-MS/MS data of *Pavlova lutheri* because these lipids were previously discovered in *P. lutheri* [19] and the corresponding literature's MS/MS spectrum was utilized to create an in-silico MS/MS library [20]; the MS/MS spectra that have electronically been described in a peer-review journal but not recorded in publicly and commercially available spectral databases such as MassBank and NIST are referred to as the literature's MS/MS. The in-silico MS/MS spectral libraries for oxidized phospholipids were developed considering our previously published data [21]. The library creation for GlcADG was based on the literature's MS/MS spectrum [22]. Information regarding ion abundances in the MS/MS spectral libraries was based on our LC-MS/MS experimental conditions and the detailed analytical conditions were described in a previous study [7]. Briefly, the MS data were acquired in information-dependent mode (IDA), i.e., DDA, using SCIEX TripleTOF 5600+ or 6600 systems. The mass range, collision energy, and collision energy spread were set to m/z 70–1250, 45 V, and 15 V, respectively.

3. Results and Discussion

3.1. Novel Lipid Characterizations in Algae with Enriched In-Silico Spectral Libraries

The global lipid profiling of nine algal lipids was achieved in 2015 and 15 lipid classes were characterized [6]. These classes include free fatty acid (FFA), di- and triacylglycerols (DAG and TAG), seven phospholipid classes (phosphatidylcholine, PC; phosphatidylethanolamine, PE; phosphatidylglycerol, PG; phosphatidylinositol, PI; phosphatidylserine, PS; lysophosphatidylcholine, LPC; and lysophosphatidylethanolamine, LPE), mono- and digalactosyldiacylglycerol (MGDG and DGDG), sulfoquinovosyldiacylglycerol (SQDG), diacylglyceryltrimethylhomoserine (DGTS), and its lyso-type form (LDGTS). Of these, the most common lipid classes in the photosynthetic membranes of plants, cyanobacteria, and algae, which include PG, SQDG, MGDG, and DGDG, have been characterized in all algal species. In contrast, *N,N,N*-trimethylammonium cation-containing lipids, i.e., PC and DGTS, were characterized as species-specific lipid classes. For example, *Chlamydomonas reinhardtii* only contains DGTS, while *Chlorella* species only contain PC as their characteristic positively charged membrane lipids. Since the specificity of lipid metabolism is highly influenced by genetics, evolution, and the environment of living organisms, increasing lipidomics coverage is an emerging requirement in biology.

Herein, 17 lipid classes were newly characterized in algal species using enriched MS/MS spectral libraries and include LPS, LPG, oxidized phospholipids (OxPC, OxPE, OxPG, OxPI), seven ceramide classes (Cer-NS, Cer-NDS, Cer-AP, Cer-NP, Cer-AS, Cer-ADS, and HexCer-AP), PMeOH, GlcADG, DGCC, and LDGCC (Table 1). The diagnostic ions are summarized in Table 2, and the annotation strategy and nomenclature of the ceramides are reported elsewhere [7]. The phytoceramide species, Cer-AP and Cer-NP, were characterized in all algal species, while several lipid classes were determined to be algal species-specific (Figure 1).

Table 1. Summary of annotated lipids.

Super Class	Class	<i>Auxenochlorella protothecoides</i>	<i>Chlorella sorokiniana</i>	<i>Chlorella variabilis</i>	<i>Chlamydomonas reinhardtii</i>	<i>Dunaliella salina</i>	<i>Euglena gracilis</i>	<i>Nannochloropsis oculata</i>	<i>Pavlova lutheri</i>	<i>Pleurochrysis carterae</i>	Total
Fatty acids	FA	5	5	5	5	4	9	6	6	9	11
Glycerolipids	DAG	13	18	21	16	2	61	27	5	13	100
Glycerolipids	TAG	91	80	144	97	152	481	231	126	121	622
Phospholipids	PC	25	21	26	0	10	33	42	1	0	75
Phospholipids	PE	12	18	19	5	5	14	16	4	1	46
Phospholipids	PG	12	16	13	9	9	9	13	8	1	28
Phospholipids	PS	3	2	3	0	0	1	0	0	0	6
Phospholipids	PI	8	8	9	8	6	7	14	8	1	19
Phospholipids	PMeOH	0	0	0	0	0	29	0	0	0	29
Phospholipids	LPC	4	8	6	0	0	1	6	0	0	10
Phospholipids	LPE	2	6	2	1	0	0	3	0	0	8
Phospholipids	LPG	0	2	1	2	0	1	0	0	0	2
Phospholipids	LPS	0	1	1	0	0	0	0	0	0	1
Oxidized lipids	OxPC	0	0	3	0	0	0	0	0	0	3
Oxidized lipids	OxPE	0	0	3	0	0	0	0	0	0	3
Oxidized lipids	OxPG	0	2	6	0	1	0	0	0	0	7
Oxidized lipids	OxPI	0	0	3	0	0	0	0	0	0	3
Algal lipids	LDGTS/LDGTA	0	0	0	14	4	9	21	1	3	34
Algal lipids	DGTS/DGTA	0	0	0	68	12	12	68	14	4	134
Algal lipids	LDGCC	0	0	0	0	0	0	0	7	14	16
Algal lipids	DGCC	0	0	0	0	0	0	0	35	38	47
Plant lipids	MGDG	25	40	35	26	16	42	18	24	27	82
Plant lipids	DGDG	19	22	28	23	15	39	26	12	28	64
Plant lipids	SQDG	16	15	10	20	8	25	22	16	24	41
Plant lipids	GlcADG	0	0	0	0	0	3	0	1	4	5
Ceramides	Cer-AP	8	8	8	7	8	8	8	8	8	10
Ceramides	Cer-NP	8	7	8	7	6	6	10	7	6	10
Ceramides	Cer-NS	0	0	0	0	0	1	4	0	2	5
Ceramides	Cer-NDS	1	2	4	1	2	3	4	2	2	6
Ceramides	Cer-AS	1	1	1	1	1	3	1	2	1	4
Ceramides	Cer-ADS	0	0	3	0	0	1	0	0	0	4
Ceramides	HexCer-AP	0	0	0	0	0	0	0	2	0	2
Total		253	282	362	310	261	798	540	289	307	1437

The count of annotated lipids is provided for each lipid class and algal species. DGTS/DGTA indicates that these lipid classes are annotated using the same characteristic ions.

Table 2. Diagnostic ions for lipid characterizations.

Lipid Class	Ion Mode	Adduct Type	Example	Diagnostic Ions (Lipid Class)	Diagnostic Ions (Acyl Chains)
LDGCC	Positive	[M+H] ⁺	LDGCC 18:0	<i>m/z</i> 104.107 C ₅ H ₁₄ NO ⁺ , <i>m/z</i> 132.102 C ₆ H ₁₄ NO ₂ ⁺	-
DGCC	Positive	[M+H] ⁺	DGCC 18:0-20:4	<i>m/z</i> 104.107 C ₅ H ₁₄ NO ⁺ , <i>m/z</i> 132.102 C ₆ H ₁₄ NO ₂ ⁺	NL of SN1 (<i>m/z</i> 538.374 C ₃₀ H ₅₂ NO ₇ ⁺), NL of SN1+H ₂ O (<i>m/z</i> 520.363 C ₃₀ H ₅₀ NO ₆ ⁺), NL of SN2 (<i>m/z</i> 518.405 C ₂₈ H ₅₆ NO ₇ ⁺), NL of SN2+H ₂ O (<i>m/z</i> 500.395 C ₂₈ H ₅₄ NO ₆ ⁺)
OxPC	Negative	[M+HCOO] ⁻	OxPC 18:0-20:4+2O	NL of HCOO+CH ₃ (<i>m/z</i> 826.56 C ₄₅ H ₈₁ NO ₁₀ P ⁻)	SN1 (<i>m/z</i> 283.264 C ₁₈ H ₃₅ O ₂ ⁻), SN2 (<i>m/z</i> 335.223, C ₂₀ H ₃₁ O ₄ ⁻), SN2-H ₂ O (<i>m/z</i> 317.212 C ₂₀ H ₂₉ O ₃ ⁻) *SN2-2H ₂ O (<i>m/z</i> 299.202 C ₂₀ H ₂₇ O ₂ ⁻)
OxPE	Negative	[M-H] ⁻	OxPE 18:0-20:4+2O	<i>m/z</i> 196.038 C ₅ H ₁₁ NO ₅ P ⁻	SN1 (<i>m/z</i> 283.264 C ₁₈ H ₃₅ O ₂ ⁻), SN2 (<i>m/z</i> 335.223, C ₂₀ H ₃₁ O ₄ ⁻), SN2-H ₂ O (<i>m/z</i> 317.212 C ₂₀ H ₂₉ O ₃ ⁻) *SN2-2H ₂ O (<i>m/z</i> 299.202 C ₂₀ H ₂₇ O ₂ ⁻)
OxPG	Negative	[M-H] ⁻	OxPG 18:0-20:4+2O	<i>m/z</i> 152.995 C ₃ H ₆ O ₅ P ⁻	SN1 (<i>m/z</i> 283.264 C ₁₈ H ₃₅ O ₂ ⁻), SN2 (<i>m/z</i> 335.223, C ₂₀ H ₃₁ O ₄ ⁻), SN2-H ₂ O (<i>m/z</i> 317.212 C ₂₀ H ₂₉ O ₃ ⁻) *SN2-2H ₂ O (<i>m/z</i> 299.202 C ₂₀ H ₂₇ O ₂ ⁻)
OxPI	Negative	[M-H] ⁻	OxPI 18:0-20:4+2O	<i>m/z</i> 297.038 C ₉ H ₁₄ O ₉ P ⁻ , <i>m/z</i> 241.012 C ₆ H ₁₀ O ₈ P ⁻	SN1 (<i>m/z</i> 283.264 C ₁₈ H ₃₅ O ₂ ⁻), SN2 (<i>m/z</i> 335.223, C ₂₀ H ₃₁ O ₄ ⁻), SN2-H ₂ O (<i>m/z</i> 317.212 C ₂₀ H ₂₉ O ₃ ⁻) *SN2-2H ₂ O (<i>m/z</i> 299.202 C ₂₀ H ₂₇ O ₂ ⁻)
PMeOH	Negative	[M-H] ⁻	PMeOH 18:0-20:4	<i>m/z</i> 167.012 C ₄ H ₈ O ₅ P ⁻ , <i>m/z</i> 110.985 CH ₄ O ₄ P ⁻	SN1 (<i>m/z</i> 283.264 C ₁₈ H ₃₅ O ₂ ⁻), SN2 (<i>m/z</i> 303.233 C ₂₀ H ₃₁ O ₂ ⁻)
GlcADG	Negative	[M-H] ⁻	GlcADG 18:0-20:4	<i>m/z</i> 249.062 C ₉ H ₁₃ O ₇	SN1 (<i>m/z</i> 283.264 C ₁₈ H ₃₅ O ₂ ⁻), SN2 (<i>m/z</i> 303.233 C ₂₀ H ₃₁ O ₂ ⁻)
LPG	Negative	[M-H] ⁻	LPG 18:0	<i>m/z</i> 152.995 C ₃ H ₆ O ₅ P ⁻	SN1 SN2 (<i>m/z</i> 283.264 C ₁₈ H ₃₅ O ₂ ⁻)
LPS	Negative	[M-H] ⁻	LPS 18:0	NL of C ₃ H ₆ NO ₂ (<i>m/z</i> 437.267 C ₂₁ H ₄₂ O ₇ P ⁻)	SN1 SN2 (<i>m/z</i> 283.264 C ₁₈ H ₃₅ O ₂ ⁻)

The product ion *m/z* or neutral loss information is summarized. In this study, LDGCC and DGCC are characterized in positive ion mode and the other lipid classes are characterized in negative ion mode. The example lipid molecule is used for the showcase of fragment ions. The term "20:4" describes the fatty acid moiety of lipids including 20 carbons and 4 double bonds. The asterisk denotes a minor fragment ion that is rarely detected under our experimental conditions.

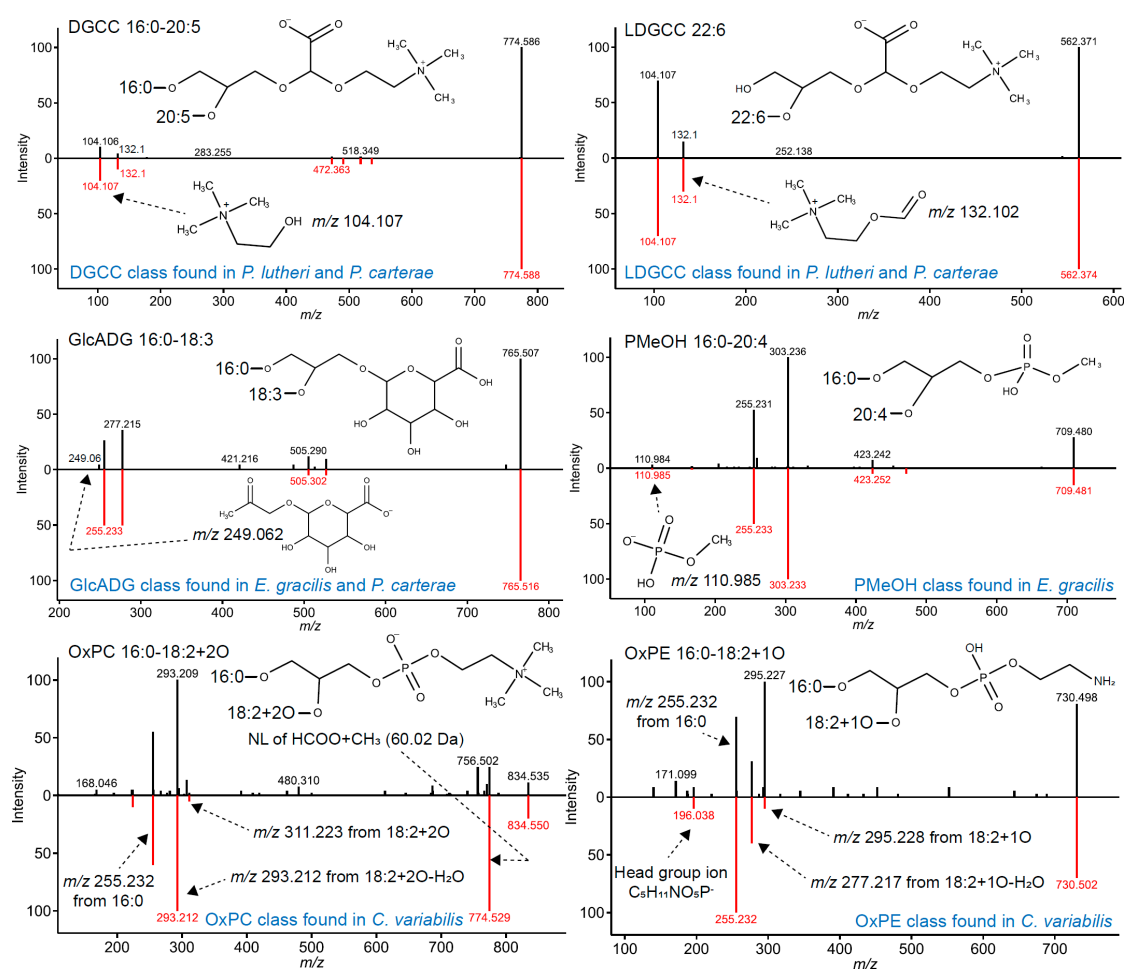


Figure 1. Lipid characterization using enriched in-silico MS/MS spectral libraries. The in-silico MS/MS characterizations for phosphatidylmethanol (PMeOH), glucuronosyldiacylglycerol (GlcADG), oxidized fatty acids containing phosphatidylcholine (OxPC) and phosphatidylethanolamine (OxPE), diacylglyceryl-carboxyhydroxy-methylcholine (DGCC), and its lyso-type form (LDGCC) are showcased. The upper- and bottom spectra show the experimental and in-silico MS/MS spectra, respectively. The string character indicates the abbreviation of fatty acids, and the ester link of the fatty acids is also described by string characters. NL refers to neutral loss. The algal species where the classified lipids are observed is also highlighted.

For example, DGCC and LDGCC lipids were only characterized in *Pavlova lutheri* and *Pleurochrysis carterae*. DGCC is well-known as a major betaine lipid of non-plastid membranes in *P. lutheri* [19], while this lipid class has never been reported in *P. carterae*. Thus, further investigation in *P. carterae* is required to define its exact stereochemistry. The MS-based lipidomics platform does not resolve the stereochemistry of acyl chains and sometimes lipid classes cannot uniquely be assigned, although a large variety of lipid molecules can be covered by tracing the lipid class-specific product ions and neutral losses. For example, DGTS and diacylglyceryl hydroxymethyl-*N,N,N*-trimethyl- β -alanine (DGTA), the major lipid class in *P. lutheri*, are characterized by the same diagnostic ions (m/z 144.102 and m/z 236.149) [23] under our experimental conditions, so the annotation must be determined by considering the genetic background in mass spectrometry-based metabolite annotations [9]. GlcADG, also known as diacylglyceryl glucuronide (DGGA), was observed in *Euglena gracilis*, *P. lutheri*, and *P. carterae*. GlcADG is known to be accumulated in response to phosphorus starvation in *Arabidopsis thaliana* and *Oryza sativa* [24], and this lipid class is commonly observed in several algal species. A previous study has also reported its existence in *P. lutheri*. Furthermore, the rare phospholipid

PMeOH class was characterized in *E. gracilis*, although it could also be detected as an artifact of extraction [25]. Finally, several oxidized fatty acid-containing phospholipids, including OxPC, OxPE, OxPG, and OxPI, were characterized in *Chlorella variabilis*. Although further investigation of these discovered lipids is required to determine whether the lipid class is endogenously biosynthesized in a specific algal species [26], our results indicated that the reanalysis of the published data with the updated annotation workflow could provide new insights and hypotheses not previously reported.

3.2. Strategy to Link Untargeted- and Targeted Analyses for Increasing Lipid Coverage

We further demonstrated the increased lipid profiling coverage by integrating untargeted and targeted analysis approaches (Figure 2). Although MS-DIAL involves a deconvolution algorithm, MS2Dec, to process the DIA-MS/MS data, MS2Dec requires at least two data point peak-top differences of co-eluted peaks for chromatogram deconvolution [6,12]. Moreover, MS-DIAL uses MS¹ chromatogram traces for metabolite quantification, while the MS/MS chromatogram can effectively be used to annotate and quantify the target metabolite in DIA-MS/MS data [15]. Therefore, the MRMPROBS program was used, in which the user-friendly GUI was available for data curation in addition to its favorable algorithm aspects [10,11], to compensate for the shortcomings of the MS-DIAL program. First, both DDA- and DIA-MS/MS data were analyzed using MS-DIAL. Second, all spectrometrically ‘matched’ candidates to an MS/MS spectrum were obtained and the function was newly developed. In this study, we utilized the DDA-MS/MS spectral data to examine the co-eluted metabolite profile because of the better spectrum quality than that of the DIA-MS/MS spectra. Third, the reference format file, which contains (1) metabolite name (2) retention time, (3) precursor ion *m/z*, and product ion *m/z* list, and (4) ion abundance ratios, was generated to cover all matched candidates. Finally, the DIA-MS/MS data were analyzed by MRMPROBS using the reference library where the peak left and right edges of the metabolite peak in each MS/MS chromatogram trace were manually refined.

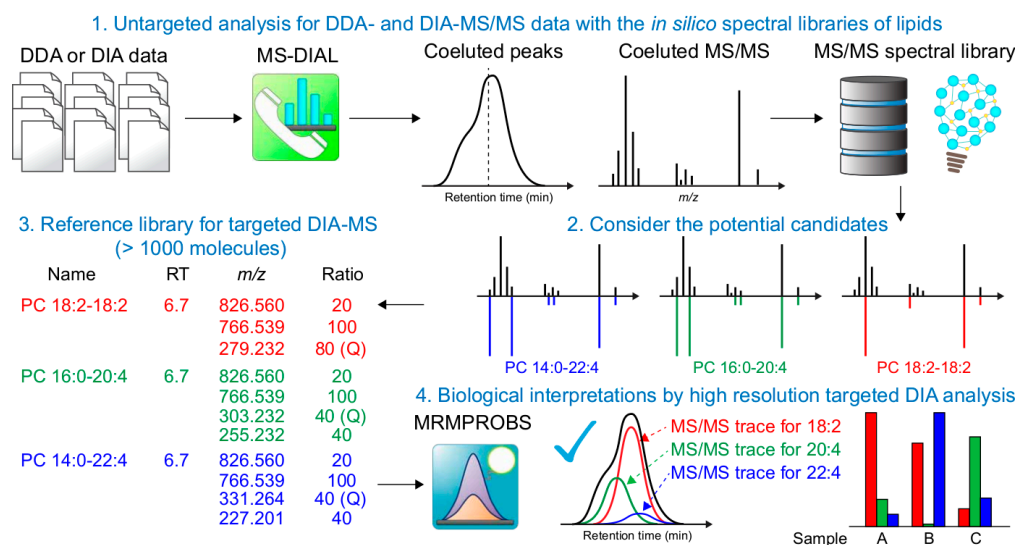


Figure 2. Integrated strategy of untargeted and targeted analyses to increase the coverage of annotated lipids. First, the data dependent (DDA) and independent (DIA) acquisition data are analyzed using the untargeted analysis pipeline where peak-picking, MS/MS assignment, and peak alignment are performed. Second, all potential lipid candidates that exceeded the cut-off of mass spectral similarity are obtained. Third, a reference library containing the target metabolite name, retention time, precursor ion *m/z*, product ion *m/z* list, and ion abundance ratios is automatically generated. In our study, the ratio “100” indicates that the trace is used to quantify the metabolite, and the diagnostic ion for characterizing the metabolite is described by “Q”. Finally, the MS/MS chromatograms are analyzed by the targeted analysis pipeline where the chromatogram traces are evaluated with the reference libraries combined with manual curations.

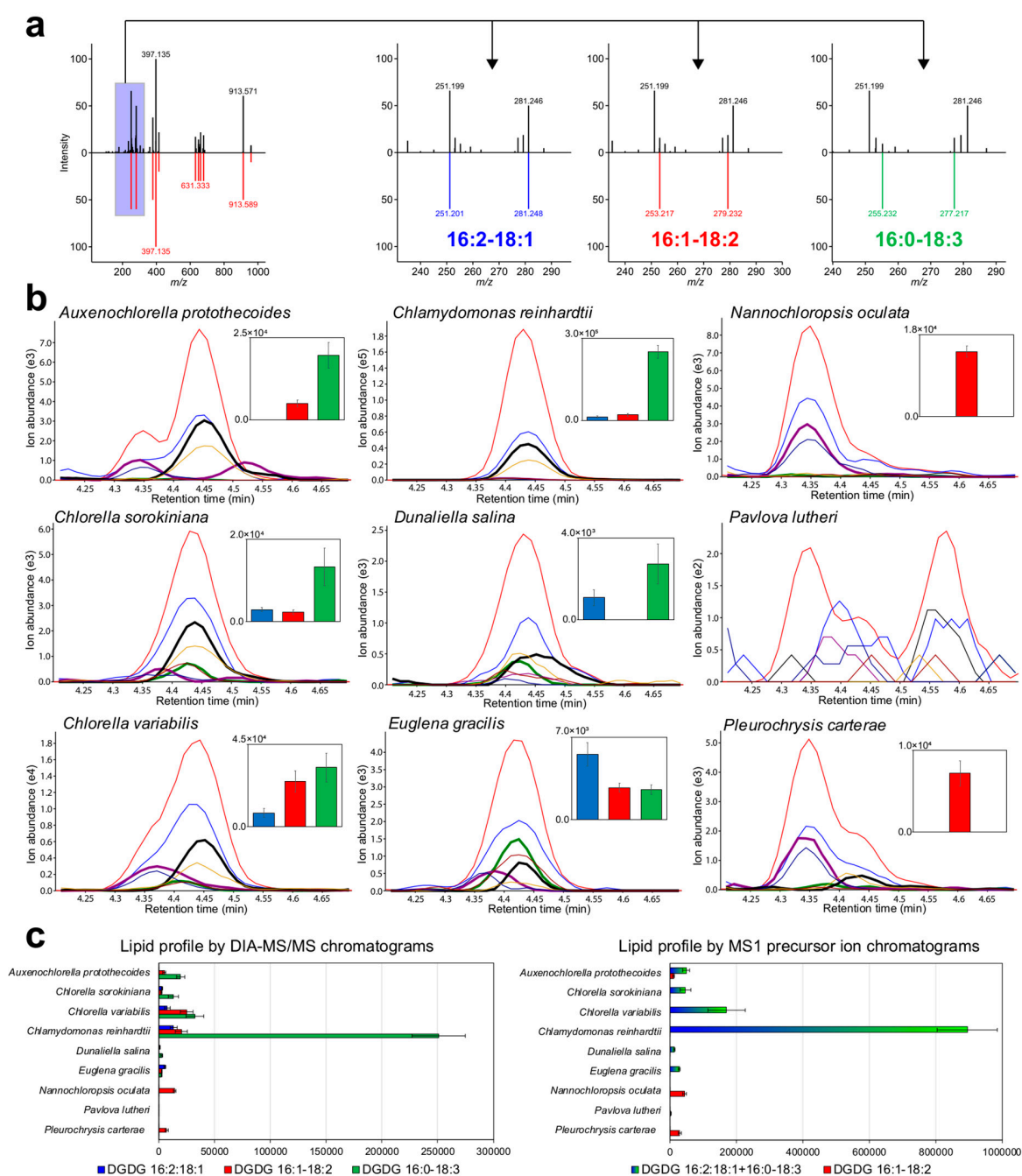
Importantly, smoothing level 1 was used in the MRMPROBS program and it was set to 3 in the MS-DIAL program. The higher smoothing level allows for the determination of peak left and right edges in the automated data analysis pipeline owing to the reduced noise level. Therefore, the higher smoothing value was used in the untargeted analysis. Conversely, the co-eluted peaks are often merged as a single peak. Because all chromatogram peaks could manually be checked and modified in the GUI, the lower smoothing value was used in the targeted analysis software.

3.3. Showcase of Newly Resolved Lipid Profiles by MS/MS-Centric Data Analysis

We showcased the methodology by profiling three co-eluted lipid molecules, DGDG 16:2–18:1, DGDG 16:1–18:2, and DGDG 16:0–18:3, with the same precursor m/z and similar retention times (Figure 3a). In the analysis of *E. gracilis*, only DGDG 16:2–18:1 was annotated correctly in MS-DIAL based on the spectral match score, while the MS/MS spectrum was partially interpreted as DGDG 16:1–18:2 and 16:0–18:3. Using the principles of MS-DIAL, the spectra of co-eluted metabolites are annotated using the representative metabolite with the highest spectral matching score, although it could be manually modified. Therefore, MS-DIAL generated the MRMPROBS reference format for the three lipid candidates, and the lipids were quantified using the MRMPORBS program. As a result, DGDG 16:0–18:3 was determined to be the major component of the co-eluted lipids in *Chlamydomonas reinhardtii*, *Auxenochlorella protothecoides*, *C. sorokiniana*, *C. variabilis*, and *Dunaliella salina*, DGDG 16:2–18:1 the major component in *E. gracilis*, and DGDG 16:1–18:2 in *Nannochloropsis oculata* and *P. carterae* (Figure 3b), where the lipid differences clearly reflected the differences in the phylum. Importantly, these differences could not be resolved using MS-DIAL untargeted analysis (Figure 3c) because the program uses MS¹-centric peak quantification, i.e., the red traces in the chromatograms shown in Figure 3b. This indicated that the DIA-MS/MS data enabled the increased coverage of metabolic profiling, as described elsewhere [27], and the two program suites MS-DIAL and MRMPROBS provided a solution to fully utilize the information-rich MS/MS spectral data for comprehensive metabolome analyses.

3.4. Comparison of Untargeted- and Targeted Analysis Results

We characterized 1437 molecules in nine algal species, and the total count of annotated lipids was 40% higher than that (1023) of the lipids annotated using the previously developed methodology (Figure 4a, Table 1, Supplementary Data 1, 2, and 3). Moreover, we examined correlations among the quantification methods, including MS¹-centric peak height in DDA and DIA-MS/MS data, and MS/MS-centric peak height and area in the DIA-MS/MS data. As expected, the correlation between peak height and area in DIA-MS/MS data were high (Figure 4b, right-bottom), but the correlation between peak heights in DDA and DIA-MS/MS data could be affected by the mass spectrometry settings [6] where the total scan cycle times were different in DDA (650 ms) and DIA-MS/MS (730 ms) data acquisition (Figure 4b, top-left). These differences were also caused by different LC-MS analysis days, where the MS sensitivities differed for each lipid class. Surprisingly, our results indicated that the correlations between MS¹ and MS/MS chromatograms were highly dependent on the lipid classes. This indicates that the sensitivity of the product ions is different for each lipid class, and the correlation value is high when abundances are compared within each lipid class. The dynamic range using MS/MS chromatograms was narrower than that of the MS¹ chromatograms, and the saturation behavior was observed in the correlation plots, especially for TAG lipids, between MS¹- and MS/MS-centric quantifications. In fact, the SWATH-MS/MS channels of the TripleTOF 5600+ instrument have a lower linear dynamic range compared to MS¹. These results suggest that MS¹-centric metabolite quantification is slightly superior to that of the MS/MS-centric quantification, while the annotation coverage in MS/MS-centric analyses outperformed the untargeted analysis pipeline.



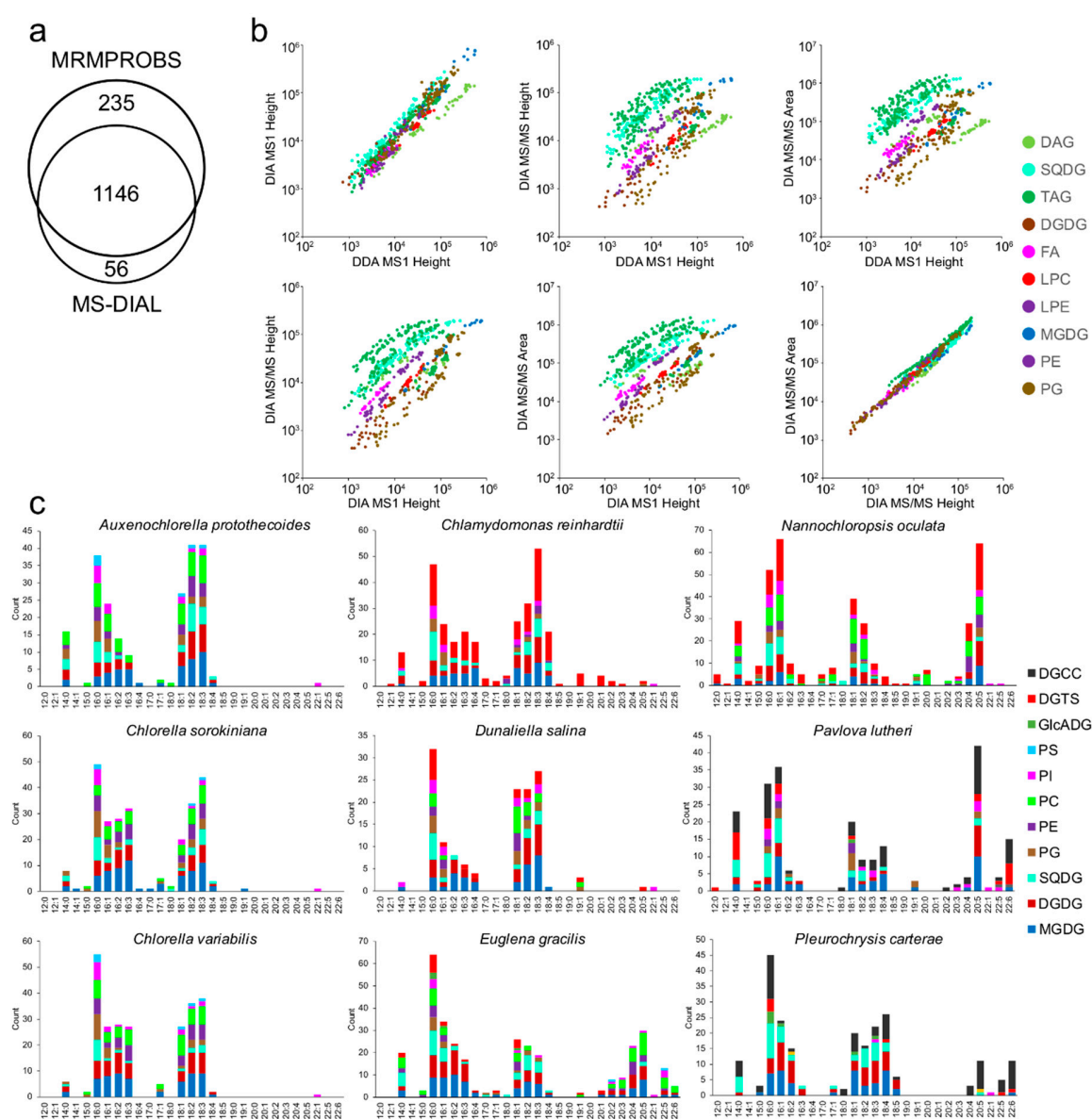


Figure 4. Summary of lipid annotations and comparison of MS¹- and MS/MS-centric peak quantifications. **(a)** Counts of the annotated lipids in untargeted (MS-DIAL) and targeted (MRMPROBS) analyses. **(b)** The correlations of ion abundances where the MS¹-centric peak heights in the DDA and DIA data and MS/MS-centric peak height and area are examined. **(c)** The statistics of acyl chains in all lipid classes except for triacylglycerol (TAG) for each algal species. The x- and y-axis shows the fatty acid information and frequency, respectively. The TAG lipids are not used for the counts because many combinations are observed in a single MS/MS spectrum, which may include false-positive identifications.

Finally, a detailed investigation of fatty acid properties revealed the uniqueness of the acyl chains in each algal species (Figure 4c). Importantly, we used the fatty acid counts included in the lipid classes instead of the ion abundances because lipid ionization efficiency is highly dependent on the lipid class and retention time. In Plantae, 16:0, 16:1, 16:2, 16:3, 18:0, 18:1, 18:2, and 18:3 are known to be major acyl chains [28], while 16:4 and 18:4 are highly distributed in Chlorophyceae including *C. reinhardtii* and *D. salina* compared to Trebouxiophyceae including *Chlorella* species. Moreover, the acyl chain of 18:0 is not often observed in glycerolipids and glycerophospholipids in nine algal species. In Chromista and Protozoa, polyunsaturated fatty acids (PUFAs) are enriched in addition to

the common 16 and 18 carbon length series. These results show that various PUFAs, such as 20:4, 20:5, 22:5, and 22:6, were observed in *E. gracilis* while 20:5 and 22:6 were enriched in *P. lutheri*, and 20:4 and 20:5 in the lipids of *N. oculata*. These observations were achieved using the MS/MS chromatogram traces for lipid quantification, and the approach is effective for investigating lipid profiles in living organisms to deepen our understanding of lipid metabolism and its connection with gene expression and enzyme activities.

In general, untargeted analysis searches are conducted for as many metabolites as possible to generate hypotheses in biology, though the rate of false-positive annotations becomes higher than that in targeted analysis; therefore, data analysts should devote much time and effort in curating annotation results. On the other hand, the targeted analysis focuses on a limited number of metabolites with less false-positive rate, though curation is still needed to modify the peak-picking results. Although the automated pipelines with the estimation of false discovery rate (FDR) in annotations have also been developed in metabolomics [29–31] for large-scale datasets like cohort studies, data analysts should recognize the pitfall in annotations that may be mentioned as false-positive metabolites in the metabolome and lipidome data sheet for statistical analyses. Therefore, the integrated analysis by untargeted and targeted techniques is important to reduce misleading results in biological studies. In fact, the metabolites of interest obtained from the integrated results must be validated using authentic standard compounds and further biological experiments to compensate for the lack of current MS instruments and informatic techniques providing limited stereochemical and isomer information.

4. Conclusions

Consequently, the reanalysis of published data was demonstrated where 17 lipid classes were newly characterized in nine algal species in addition to the 15 lipid classes annotated previously. In effect, the coverage of lipid classes was doubled by updating the computational mass spectrometry techniques and mass spectral libraries, and our reanalysis indicates the value of MS data repositories where the raw data could be utilized as a benchmark for new software programs and data-driven hypothesis generation. The lipidomics workflow is also executable with hydrophilic interaction chromatography (HILIC) [32] or supercritical fluid chromatography (SFC) [33], in which the molecules can be separated based on the specific chemical properties of each lipid class, enabling efficient exclusion of false-positive annotations from incorrect lipid classes. Ion mobility MS provides another diagnostic criterion, viz. collision cross-section (CCS), to increase the confidence in lipid annotation [34]. Although we only showcased the increase in lipid profiling coverage, this strategy could also be applied to more diverse metabolites with experimentally acquired spectral libraries. There are three types of spectral databases: (1) completely open-access, i.e., all records are browsable and downloadable (e.g., MassBank [35], PlaSMA [9], Fiehnlib [36], and GNPS [4]), (2) limited access, i.e., browsable but not downloadable (e.g., Metlin [17] and mzCloud (<https://www.mzcloud.org/>)), and (3) licensed databases such as NIST and Wiley, and the integrated databases cover the MS/MS spectra of approximately 12,000 unique metabolites [9,26]. Importantly, all these databases have increasingly been updated by the continuous effort of mass spectrometrists; therefore, success similar to that obtained herein is achievable by reanalyzing public data using the upgraded databases.

In conclusion, the science of metabolomics and lipidomics now enters a new era owing to state-of-the-art analytical techniques and informatics platforms where metabolic profiling is semi-automatically executable [37]. Therefore, MS data repositories will become increasingly important to reach a 'standard' in genomics and transcriptomics data sciences. Our computational workflow could be used as a pipeline for metabolomics and lipidomics data processing and the understanding of metabolism is deepened by advances in computational metabolomics.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2218-1989/9/6/119/s1>. Supplementary Data 1. MS-DIAL original output for DIA-MS/MS data. The peak height is exported. The representative SMILES and InChIKey codes are also exported although the stereochemistry is not determined spectrometrically; Supplementary Data 2. MS-DIAL original output for DDA-MS/MS data. The peak height is

exported. The representative SMILES and InChIKey codes are also exported although the stereochemistry is not determined spectrometrically; Supplementary Data 3. MRMPROBS output for DIA-MS/MS data. The peak height and area values are exported.

Author Contributions: H.T., T.C., M.A. (Makoto Arita) and M.A. (Masanori Arita) designed the research. H.T. wrote all source codes for MS-DIAL and MRMPROBS. H.T. and H.U. analyzed standard compounds and created in-silico MS/MS spectral library. A.S. performed the reanalysis of the algal lipidomics data. H.T. wrote the manuscript, and all authors helped to improve the manuscript.

Funding: This work was partially supported by JSPS KAKENHI (15H05897, 15H05898, 17H03621, 18H02432, 18K19155) and JST National Bioscience Database Center (NBDC).

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

References

1. Haug, K.; Salek, R.M.; Steinbeck, C. Global Open Data Management in Metabolomics. *Curr. Opin. Chem. Biol.* **2017**, *36*, 58–63. [[CrossRef](#)] [[PubMed](#)]
2. Sud, M.; Fahy, E.; Cotter, D.; Azam, K.; Vadivelu, I.; Burant, C.; Edison, A.; Fiehn, O.; Higashi, R.; Nair, K.S.; et al. Metabolomics Workbench: An International Repository for Metabolomics Data and Metadata, Metabolite Standards, Protocols, Tutorials and Training, and Analysis Tools. *Nucleic Acids Res.* **2016**, *44*, D463–D470. [[CrossRef](#)] [[PubMed](#)]
3. Kale, N.S.; Haug, K.; Conesa, P.; Jayseelan, K.; Moreno, P.; Rocca-Serra, P.; Nainala, V.C.; Spicer, R.A.; Williams, M.; Li, X.; et al. MetaboLights: An Open-Access Database Repository for Metabolomics Data. *Curr. Protoc. Bioinform.* **2016**, *53*, 14.13.1–14.13.18.
4. Wang, M.; Carver, J.J.; Phelan, V.V.; Sanchez, L.M.; Garg, N.; Peng, Y.; Nguyen, D.D.; Watrous, J.; Kaponov, C.A.; Luzzatto-Knaan, T.; et al. Sharing and Community Curation of Mass Spectrometry Data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, *34*, 828–837. [[CrossRef](#)] [[PubMed](#)]
5. Spicer, R.A.; Steinbeck, C. A Lost Opportunity for Science: Journals Promote Data Sharing in Metabolomics but Do Not Enforce It. *Metabolomics* **2018**, *14*. [[CrossRef](#)] [[PubMed](#)]
6. Tsugawa, H.; Cajka, T.; Kind, T.; Ma, Y.; Higgins, B.; Ikeda, K.; Kanazawa, M.; VanderGheynst, J.; Fiehn, O.; Arita, M. MS-DIAL: Data-Independent MS/MS Deconvolution for Comprehensive Metabolome Analysis. *Nat. Methods* **2015**, *12*, 523–526. [[CrossRef](#)] [[PubMed](#)]
7. Tsugawa, H.; Ikeda, K.; Tanaka, W.; Senoo, Y.; Arita, M.; Arita, M. Comprehensive Identification of Sphingolipid Species by in Silico Retention Time and Tandem Mass Spectral Library. *J. Cheminform.* **2017**, *9*, 19. [[CrossRef](#)]
8. Tsugawa, H.; Ikeda, K.; Arita, M. The importance of bioinformatics for connecting data-driven lipidomics and biological insights. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **2017**, *1862*, 762–765. [[CrossRef](#)]
9. Tsugawa, H.; Nakabayashi, R.; Mori, T.; Yamada, Y.; Takahashi, M.; Rai, A.; Sugiyama, R.; Yamamoto, H.; Nakaya, T.; Yamazaki, M.; et al. A Cheminformatics Approach to Characterize Metabolomes in Stable-Isotope-Labeled Organisms. *Nat. Methods* **2019**, *16*, 295–298. [[CrossRef](#)]
10. Tsugawa, H.; Arita, M.; Kanazawa, M.; Ogiwara, A.; Bamba, T.; Fukusaki, E. MRMPROBS: A Data Assessment and Metabolite Identification Tool for Large-Scale Multiple Reaction Monitoring Based Widely Targeted Metabolomics. *Anal. Chem.* **2013**, *85*, 5191–5199. [[CrossRef](#)]
11. Tsugawa, H.; Kanazawa, M.; Ogiwara, A.; Arita, M. MRMPROBS Suite for Metabolomics Using Large-Scale MRM Assays. *Bioinformatics* **2014**, *30*, 2379–2380. [[CrossRef](#)] [[PubMed](#)]
12. Lai, Z.; Tsugawa, H.; Wohlgemuth, G.; Mehta, S.; Mueller, M.; Zheng, Y.; Ogiwara, A.; Meissen, J.; Showalter, M.; Takeuchi, K.; et al. Identifying Metabolites by Integrating Metabolome Databases with Mass Spectrometry Cheminformatics. *Nat. Methods* **2017**, *15*, 53–56. [[CrossRef](#)] [[PubMed](#)]
13. Mahieu, N.G.; Genenbacher, J.L.; Patti, G.J. A Roadmap for the XCMS Family of Software Solutions in Metabolomics. *Curr. Opin. Chem. Biol.* **2016**, *30*, 87–93. [[CrossRef](#)] [[PubMed](#)]
14. Pluskal, T.; Castillo, S.; Villar-Briones, A.; Oresic, M. MZmine 2: Modular Framework for Processing, Visualizing, and Analyzing Mass Spectrometry-Based Molecular Profile Data. *BMC Bioinform.* **2010**, *11*, 395. [[CrossRef](#)] [[PubMed](#)]
15. Li, H.; Cai, Y.; Guo, Y.; Chen, F.; Zhu, Z.J. MetDIA: Targeted Metabolite Extraction of Multiplexed MS/MS Spectra Generated by Data-Independent Acquisition. *Anal. Chem.* **2016**, *88*, 8757–8764. [[CrossRef](#)] [[PubMed](#)]

16. Perez-Riverol, Y.; Bai, M.; da Veiga Leprevost, F.; Squizzato, S.; Park, Y.M.; Haug, K.; Carroll, A.J.; Spalding, D.; Paschall, J.; Wang, M.; et al. Discovering and Linking Public Omics Data Sets Using the Omics Discovery Index. *Nat. Biotechnol.* **2017**, *35*, 406–409. [[CrossRef](#)] [[PubMed](#)]
17. Koellensperger, G.; Guijas, C.; Benton, H.P.; Huan, T.; Wolan, D.W.; Warth, B.; Aisporna, A.E.; Hermann, G.; Domingo-Almenara, X.; Spilker, M.E.; et al. METLIN: A Technology Platform for Identifying Knowns and Unknowns. *Anal. Chem.* **2018**, *90*, 3156–3164.
18. Kind, T.; Liu, K.-H.; Lee, D.Y.; DeFelice, B.; Meissen, J.K.; Fiehn, O. LipidBlast in Silico Tandem Mass Spectrometry Database for Lipid Identification. *Nat. Methods* **2013**, *10*, 755–758. [[CrossRef](#)]
19. Eichenberger, W.; Gribi, C. Lipids of *Pavlova Lutheri*: Cellular Site and Metabolic Role of DGCC. *Phytochemistry* **1997**, *45*, 1561–1567. [[CrossRef](#)]
20. Schleyer, G.; Shahaf, N.; Ziv, C.; Dong, Y.; Meoded, R.A.; Helfrich, E.J.N.; Schatz, D.; Rosenwasser, S.; Rogachev, I.; Aharoni, A.; et al. In Plaque-Mass Spectrometry Imaging of a Bloom-Forming Alga during Viral Infection Reveals a Metabolic Shift towards Odd-Chain Fatty Acid Lipids. *Nat. Microbiol.* **2019**, *4*, 527–538. [[CrossRef](#)]
21. Aoyagi, R.; Ikeda, K.; Isobe, Y.; Arita, M. Comprehensive Analyses of Oxidized Phospholipids Using a Measured MS/MS Spectra Library. *J. Lipid Res.* **2017**, *58*, 2229–2237. [[CrossRef](#)] [[PubMed](#)]
22. Okazaki, Y.; Otsuki, H.; Narisawa, T.; Kobayashi, M.; Sawai, S.; Kamide, Y.; Kusano, M.; Aoki, T.; Hirai, M.Y.; Saito, K. A New Class of Plant Lipid Is Essential for Protection against Phosphorus Depletion. *Nat. Commun.* **2013**, *4*, 1510. [[CrossRef](#)] [[PubMed](#)]
23. Li, Y.; Lou, Y.; Mu, T.; Xu, J.; Zhou, C.; Yan, X. Simultaneous Structural Identification of Diacylglycerol-N-Trimethylhomoserine (DGTS) and Diacylglycerolhydroxymethyl-N, N, N-Trimethyl- β -Alanine (DGTA) in Microalgae Using Dual Li^+/H^+ Adduct Ion Mode by Ultra-Performance Liquid Chromatography/Quadrupole Time-Of-Flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2017**, *31*, 824. [[PubMed](#)]
24. Okazaki, Y.; Nishizawa, T.; Takano, K.; Ohnishi, M.; Mimura, T.; Saito, K. Induced accumulation of glucuronosyldiacylglycerol in tomato and soybean under phosphorus deprivation. *Physiol. Plant.* **2015**, *155*, 33–42. [[CrossRef](#)] [[PubMed](#)]
25. Roughan, P.G.; Slack, C.R.; Holland, R. Generation of phospholipid artefacts during extraction of developing soybean seeds with methanolic solvents. *Lipids* **1978**, *13*, 497–503. [[CrossRef](#)]
26. Reis, A.; Spickett, C.M. Chemistry of Phospholipid Oxidation. *Biochim. Biophys. Acta-Biomembr.* **2012**, *1818*, 2374–2387. [[CrossRef](#)]
27. Bonnera, R.; Hopfgartner, G. SWATH data independent acquisition mass spectrometry for metabolomics. *TrAc Trends Anal. Chem.* **2018**. [[CrossRef](#)]
28. Ohlrogge, J.; Browse, J. Lipid Biosynthesis. *Plant Cell* **2007**, *7*, 957–970. [[CrossRef](#)]
29. Matsuda, F.; Tsugawa, H.; Fukusaki, E. Method for Assessing the Statistical Significance of Mass Spectral Similarities Using Basic Local Alignment Search Tool Statistics. *Anal. Chem.* **2013**, *85*, 8291–8297. [[CrossRef](#)]
30. Palmer, A.; Phapale, P.; Chernyavsky, I.; Lavigne, R.; Fay, D.; Tarasov, A.; Kovalev, V.; Fuchser, J.; Nikolenko, S.; Pineau, C.; et al. FDR-Controlled Metabolite Annotation for High-Resolution Imaging Mass Spectrometry. *Nat. Methods* **2016**, *14*, 57–60. [[CrossRef](#)] [[PubMed](#)]
31. Scheubert, K.; Hufsky, F.; Petras, D.; Wang, M.; Nothias, L.-F.; Dührkop, K.; Bandeira, N.; Dorrestein, P.C.; Böcker, S. Significance Estimation for Large Scale Metabolomics Annotations by Spectral Matching. *Nat. Commun.* **2017**, *8*, 1494. [[CrossRef](#)] [[PubMed](#)]
32. Pham, T.H.; Zaeem, M.; Fillier, T.A.; Nadeem, M.; Vidal, N.P.; Manful, C.; Cheema, S.; Cheema, M.; Thomas, R.H. Targeting Modified Lipids during Routine Lipidomics Analysis Using HILIC and C30 Reverse Phase Liquid Chromatography Coupled to Mass Spectrometry. *Sci. Rep.* **2019**, *9*. [[CrossRef](#)] [[PubMed](#)]
33. Lísá, M.; Holčápek, M. High-Throughput and Comprehensive Lipidomic Analysis Using Ultrahigh-Performance Supercritical Fluid Chromatography-Mass Spectrometry. *Anal. Chem.* **2015**, *87*, 7187–7195. [[CrossRef](#)] [[PubMed](#)]
34. Blaženović, I.; Shen, T.; Mehta, S.S.; Kind, T.; Ji, J.; Piparo, M.; Cacciola, F.; Mondello, L.; Fiehn, O. Increasing Compound Identification Rates in Untargeted Lipidomics Research with Liquid Chromatography Drift Time-Ion Mobility Mass Spectrometry. *Anal. Chem.* **2018**, *90*, 10758–10764. [[CrossRef](#)] [[PubMed](#)]
35. Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.; Tanaka, S.; Aoshima, K.; et al. MassBank: A Public Repository for Sharing Mass Spectral Data for Life Sciences. *J. Mass Spectrom.* **2010**, *45*, 703–714. [[CrossRef](#)] [[PubMed](#)]

36. Kind, T.; Wohlgemuth, G.; Lee, D.Y.; Lu, Y.; Palazoglu, M.; Shahbaz, S.; Fiehn, O. FiehnLib: Mass Spectral and Retention Index Libraries for Metabolomics Based on Quadrupole and Time-of-Flight Gas Chromatography/Mass Spectrometry. *Anal. Chem.* **2009**, *81*, 10038–10048. [[CrossRef](#)]
37. Tsugawa, H. Advances in Computational Metabolomics and Databases Deepen the Understanding of Metabolisms. *Curr. Opin. Biotechnol.* **2018**, *54*, 10–17. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).