Marine Management Organisation

Evaluating and advancing methodologies to validate the origin of capture of fish / seafood (MMO1193)

INVESTORS IN PEOPLE | Bronze

MMO1193: Evaluating and advancing methodologies to validate the origin of capture of fish / seafood

October 2019

Report prepared by: Fera Science Ltd

Project funded by: European Maritime and Fisheries Fund, grant number ENG3487

Version	Author	Note
0.1	Michael Dickinson	First draft
1.0	Chris Graham	MMO Comments
2.0	Michael Dickinson	Second draft
3.0	Michael Dickinson	Final

© Marine Management Organisation 2019

You may use and re-use the information featured on this publication (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence. Visit www.nationalarchives.gov.uk/doc/open-government-licence/ to view the licence or write to:

Information Policy Team
The National Archives
Kew
London
TW9 4DU
Email: psi@nationalarchives.gsi.gov.uk

Information about this publication and further copies are available from:

Marine Management Organisation
Lancaster House
Hampshire Court
Newcastle upon Tyne
NE4 7YH

Tel: 0300 123 1032
Email: info@marinemanagement.org.uk
Website: www.gov.uk/mmo

Disclaimer

This report contributes to the Marine Management Organisation (MMO) evidence base which is a resource developed through a large range of research activity and methods carried out by both MMO and external experts.

The opinions expressed in this report do not necessarily reflect the views of MMO nor are they intended to indicate how MMO will act on a given set of facts or signify any preference for one research activity or method over another. MMO is not liable for the accuracy or completeness of the information contained nor is it responsible for any use of the content.

Report Citation

MMO (2019). Evaluating and advancing methodologies to validate the origin of capture of fish / seafood. A report produced for the Marine Management Organisation, MMO Project No: 1193 October 2019, 20pp.

Contents

Executive Summary ... 6

1. Introduction ... 7

2. Sample description ... 8

3. Methodologies ... 9
3.1. Sample preparation ... 9
3.2. Sample extraction ... 9
3.2.1. Lipidomics analysis – non-targeted ... 9
3.2.2. GC-FID analysis – targeted ... 9
3.3. Data acquisition and analysis ... 9
3.3.1. Lipidomics analysis – non-targeted ... 9
3.3.2. GC-FID analysis – targeted ... 10

4. Results ... 11
4.1. LC-HRMS analysis – non-targeted ... 11
4.2. GC-FID analysis, targeted fatty acid determination ... 14

5. Discussion ... 17
5.1. Non-targeted lipid analysis ... 17
5.2. Targeted fatty acid analysis ... 18

6. Conclusions and Recommendations ... 18

7. References ... 19

Figures

Figure 1 - PCA of all samples (purple) and QC samples (blue), in negative mode from LC-HRMS analysis. ... 11
Figure 2 - PCA of all samples (purple) and QC samples (blue), in positive mode from LC-HRMS analysis. ... 11
Figure 3 - Box and whisker plot for compound *m/z* 452.3131 (from the positive mode LC-HRMS analysis) tentatively identified as the lipid class compound: 1-(1Z-tetradecenyl)-sn-glycero-3-phosphocholine (PC(P-14:0/0:0). ... 14
Figure 4 - Box and whisker plot for compound *m/z* 301.2159 (from the positive mode LC-HRMS analysis) tentatively identified as the lipid class compound: 8,11,14-Eicosatetraynoic acid. ... 14
Figure 5 - PCA of Dorset (green) and Ancona (red) samples from the targeted fatty acid profile data ... 17

Tables

Table 1 - Sample details; samples also used for targeted GC-FID analysis, highlighted in green 8
Table 2 - List of compounds showing significantly different abundances between sample sets, tentatively identified in negative mode by LC-HRMS.. 12
Table 3 - List of compounds showing significantly different abundances between sample sets, tentatively identified in positive mode by LC-HRMS. 13
Table 4 - Summary of fatty acids detected and their % of the total, from 6 samples selected from the Ancona sample set........ 15
Table 5 - Summary of fatty acids detected and their % of the total, from 6 samples selected from the Dorset sample set 16

Executive Summary

A proof of principle study was undertaken to assess if monitoring the lipid profiles of manilla clams (*Tapes philippinarum*) can help authenticate the original area of capture.

The study was undertaken using both a non-targeted (looking for unknowns) and targeted study (fatty acid content) of the clam's biochemical make up.

Results indicate lipid biochemical profile of clams can distinguish between a region in the English Channel versus a region in the Adriatic. In this study, differentiation was established using both non-targeted and targeted approaches.

The level of implementation of lipid analysis when assessed against the Technology Readiness Levels (TRL) as defined by the European Research Council (2014) (where the higher the number (1-9) means the more advanced and operationally applicable) suggests this approach is likely to be currently TRL 4 - Technology validated in laboratory.

Further work is required to validate this approach, however in principle the MMO could begin to contract laboratories to use the lipid profile to answer the origin question if known authentic samples of the given species could be provided alongside suspect ones.

1. Introduction

As part of the Marine Management Organisation's (MMO) ongoing work to enhance provenance and traceability, the MMO sought an assessment of methods for confirming where fish and shellfish sold ashore were originally caught to potentially support and verify other systems for managing compliance and enforcement.

The International Council for the Exploration of the Sea (ICES) is a global organisation that advises competent authorities on marine policy and management issues related to the impacts of human activities on marine ecosystems and the sustainable use of living marine resources. The organisation consists of 20 members (countries) including the United Kingdom. The oceans around these member countries are divided into ICES Areas and fishing rights are awarded for specific ICES Areas. The logistics of travel to a specific ICES Area and differences in abundances of fish within different ICES Areas create a potential for unscrupulous fishermen to capture fish from one area and declare it to an area they have a greater quota allowance in. Therefore, tools that verify the origin of capture to specific ICES areas are required to ensure correct management of the oceans.

This project investigated the use of lipid profiling to verify origin of capture. The lipid composition of biological species is determined by multiple factors, such as genetics, season, sex, and also geographical origin. Lipid analyses have been applied successfully to discriminate the geographical provenance of marine species (Garrido et al., (2017) Murzina et al., (2013)). As a proof of principle investigation this project employed non-targeted analyses of lipids (lipidomics) using liquid chromatography - high resolution mass spectrometry (LC-HRMS) to Manilla clam samples, collected in two different regions. Sample collection was carefully planned to minimise the effect of other factors that may influence the lipid status of the chosen species, such as storage conditions after capture. This enabled preliminary statistical analysis to investigate the potential of the methodology for the intended purpose.

In addition, targeted analyses of a selection of the same Manilla clam samples were performed by gas chromatography flame ionisation detection (GC-FID) for Fatty Acid Methyl Esters (FAMEs) profiles, to assess the fatty acid profile of each sample set. A total of thirty-nine Manila clam samples, from 2 different locations (Poole, Dorset in the English Channel and Ancona in the Adriatic) provided by Southern Inshore Fisheries Conservation Authority (IFCA) and The University of Portsmouth have been analysed by LC-HRMS and GC-FID.

This report describes the analytical methodologies employed, the data acquired for each sample, suggested conclusions and outlook for further work.

2. Sample description

Assignment from location A (Ancona, Adriatic) was received on 30th April 2019 from The University of Portsmouth. 135 frozen individual Manilla clams still in shell were received, weighed (average weight: 5.23g ± 1.11g), and kept on dry ice.

Composite samples were created of 5 clams each, resulting in 27 samples in total. 19 of these (randomly chosen) were taken for the main analyses and 8 were used for method development.

Assignment from location B (Poole, Dorset) was received on 19th June 2019 from Southern IFCA. 101 frozen individual Manilla clams still in shell were weighed (average weight: 16.41g ± 3.62g) frozen, and kept on dry ice. 20 composites were created consisting of 5 clams each and taken for analysis.

All composite samples were logged into the Fera Laboratory Information Management System and assigned a unique sample number; details of analysed composites are given in Table 1. Samples were stored at less than -40°C, prior to preparation and extraction.

Table 1 - Sample details; samples also used for targeted GC-FID analysis, highlighted in green

Composite Clam (n=5) sample number	Location A - Ancona, Fera LIMS	Location B – Dorset, Fera LIMS
1	Not analysed	S19-032388
2	S19-020924	S19-032389
3	S19-020925	S19-032390
4	S19-020926	S19-032391
5	S19-020927	S19-032392
6	S19-020928	S19-032393
7	S19-020929	S19-032394
8	S19-020930	S19-032395
9	S19-020931	S19-032396
10	S19-020932	S19-032397
11	S19-020933	S19-032398
12	S19-020934	S19-032399
13	S19-020935	S19-032400
14	S19-020936	S19-032401
15	S19-020937	S19-032402
16	S19-020938	S19-032403
17	S19-020939	S19-032404
18	S19-020940	S19-032405
19	S19-020941	S19-032406
20	S19-020924	S19-032407

3. Methodologies

3.1. Sample preparation

Composite samples were defrosted briefly in the fridge, before the clam muscles were separated from the shells, transferred to crystallisation dishes and frozen with liquid nitrogen, ready for freeze-drying for a minimum of 96 hours.

Freeze-dried composite samples were directly homogenised (using an IKA grinder) to a fine homogenous powder, transferred to storage containers and kept in desiccators at room temperature, prior to extraction.

3.2. Sample extraction

3.2.1. Lipidomics analysis – non-targeted

From each clam composite, a sub-sample was taken (500mg ± 20mg). To this sub-sample 5ml of hexane/ethanol mixture (1:1, volume per volume) was added. The sample containers were briefly vortexed (for approximately 10 seconds), before being placed in a heated water bath 50°C for 90 min, whilst being shaken at 300rpm. After brief vortexing (for approximately 10 seconds), samples were centrifuged for 10 minutes (3220*g*, 20°C) before a second higher speed centrifugation step (21,000*g*, 20°C). Nine-hundred µL (micro-litre) of the supernatants were transferred to vials for analysis by LC-HRMS. A quality control (QC) sample was prepared by combining 100µL aliquots of each sample extract and briefly shaken.

All extracts were stored in the fridge at 4°C, prior to analysis.

3.2.2. GC-FID analysis – targeted

Fatty acid profiles were derived from six randomly selected composite samples per location (table 1). Profiles were determined as described in Christie (1990). This involved the esterification of all fatty acids by heating with methanol in the presence of a sulphuric acid catalyst. Esterification created fatty acid methyl esters (FAMEs) that were extracted in the organic solvent hexane before determination.

A reference olive oil sample with known fatty acid composition (quality control sample, S12-006382) was also prepared and analysed by GC-FID using identical procedures described.

Identities of the fatty acids were established using a reference FAME standard (Sigma Aldrich) analysed alongside the samples.

3.3. Data acquisition and analysis

3.3.1. Lipidomics analysis – non-targeted

LC-HRMS analysis was performed on an AccelaHigh Speed Liquid Chromatography system from Thermo Fisher Scientific. The column used was an ACE 3Q 150x3mm, 3µm (Advanced Chromatography Technologies). Mobile phases were 0.1% formic

acid in water (Mobile Phase A, (MPA)) and 0.1% formic acid in acetonitrile (Mobile Phase B, (MPB)). Gradient applied was 100% MPA for 5 minutes before increasing to 100% MPB over 15 minutes. 100% MPB was held for 10 minutes before reverting to 100% MPA and held for 2 minutes. Injection volume was 10μl, flow rate was 0.4ml/min and column temperature was 25°C. The mass spectrometer used was a Thermo Exactive (Thermo Fisher Scientific). Data were acquired in two separate batches to cover both positive and negative ionisation modes.

Data were evaluated using Progenesis QI (Waters Corporation). Progenesis selected all potential peaks from each mass spectrometer data file, aligned the files using retention time information and looked for significant differences in the peaks found and their relative abundances between files. A principal component analysis (PCA) was produced to observe the data in a non-supervised fashion in order to associate potential trends and / or outliers within the whole data set.

The list of peaks found to be significantly different between files depends on filters chosen by the analyst. In these experiments all peaks detected with a retention time <1 minute and with a width <0.06 seconds were firstly removed (>95% of these are background "noise"). Two data analysis experiments were undertaken in negative and positive ion mode: Manilla clams from Ancona (Adriatic, Italy) versus Manila clams from Poole (Dorset, UK) was assessed. Peaks were only included in a final significantly different list if $p < 0.01$ (using Student's t-test analysis, with a correction for false detection rate using the q-value), mean fold change abundance between groups >5 and % Coefficient Variation (% CV) of the response of the peaks in each sample set <30%.

The top 18 substances of highest abundance (peak area), detected in the samples for both ionisation modes, were taken for tentative identification using the publicly available online library Metlin (Scripps Research) as described in Levin, et al (2016).

3.3.2. GC-FID analysis – targeted

Analysis was undertaken using Gas Chromatography – Flame Ionisation Detection [GC-FID] (Agilent 6890) by split injection (split ratio 20:1) of 1μL (microlitre) of the extract onto a SP-2560 GC column (75m x 0.18mm i.d., 0.14 μm (micrometre) film thickness, Supelco). Temperature gradient was as follows: 140°C (held for 5 minutes) and ramped at 4°C/minute to 240°C (held for 15 minutes). The injector was kept at 260°C and helium was used as a carrier gas at a constant pressure of 415kPa. The quality control material: extra virgin olive oil was derivatised (resulting in known FAMEs profile) alongside the samples. All samples, including QC, were run randomised against analytical standards of FAMEs and solvent blanks for identification purposes.

Data compilation and t-test statistics were undertaken in Excel 365 (Microsoft). Multivariate analysis (PCA) was undertaken in Matlab version R2011b (MathWorks).

4. Results

4.1. LC-HRMS analysis – non-targeted

The differences in the biochemical profiles between both clam populations are shown in Figures 1 and 2. Ancona samples were separated from the Dorset samples on Principal Component 1 (PC1, All Ancona samples group to the left and Dorset to the right on PC1). This shows there are clear biochemical differences between the two sample sets as received.

Figure 1 - PCA of all samples, Ancona (blue) and Dorset (purple), in negative mode from LC-HRMS analysis.

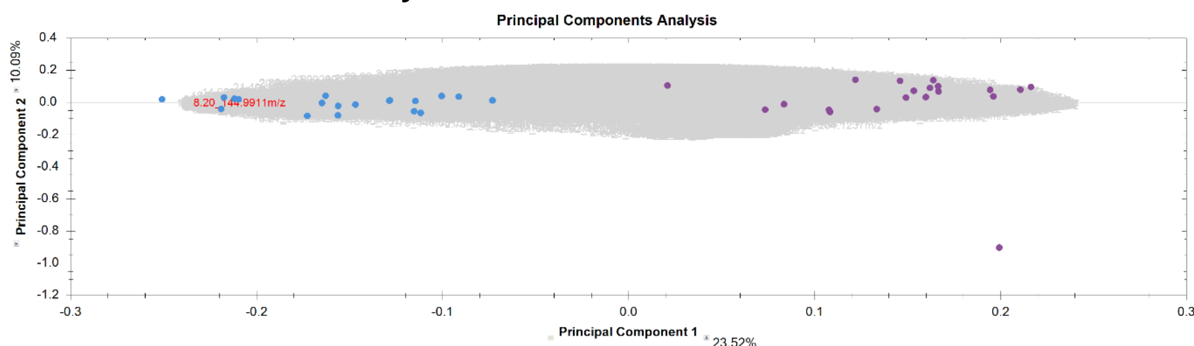


Figure 2 - PCA of all samples, Ancona (blue) and Dorset (purple), in positive mode from LC-HRMS analysis.

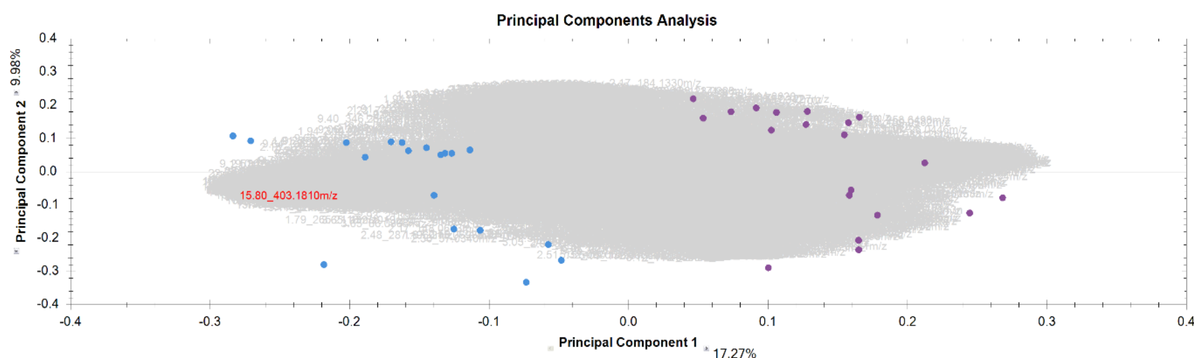


Table 2 provides combined information for the most abundant 18 peaks (per ionisation mode) and their potential identification. The compound identifications are tentative only and have not been confirmed with a reference analytical standard.

Figure 3 and Figure 4 provide box and whisker plots for 2 example compounds. These plots show the significantly different mean abundance for the compound between sample sets along with the abundance spread (variation) within each set.

Table 2 - List of compounds showing significantly different abundances between sample sets, tentatively identified in negative mode by LC-HRMS. A grey box indicates not identified. A = Ancona , D = Dorset sample set. Compounds in italics are lipids or lipid based molecules.

No	Compound mass/z	Retention time	Highest group	Tentative identification	Suggested formula
1	385.1789	16.7	A		
2	365.1969	14.2	A	20-hydroxy-Resolvin E1	C20H30O6
3	552.3314	34.8	A	PS(20:0/0:0)	C26H52NO9P
4	411.2309	21.0	A	Benzoic acid, 4-(2-butoxybenzamido)-, 2-(diethylamino) ethyl ester	C24H32N2O4
5	381.1919	13.3	A	Cinncassiol A	C20H30O7
6	363.1814	14.3	A	3b,8a-Dihydroxy-6b-angeloyloxy-7(11)-eremophilen-12,8-olide;	C20H28O6
7	443.1841	15.4	D		
8	299.0982	2.9	A		
9	379.1765	15.1	A	(1R,5S,6S)-3-(hydroxymethyl)-6-isopropyl-5-(((Z)-2-methylbut-2-enoyl)oxy)-2-oxocyclohex-3-en-1-yl (E)-4-hydroxy-2-methylbut-2-enoate	C20H28O7
10	411.2025	14.3	D	5-(2,3-Dihydroxy-3-methylbutyl)-4-(3,4-epoxy-4-methylpentanoyl)-3,4-dihydroxy-2-isopentanoyl-2-cyclopenten-1-one	C21H32O8
11	339.1814	13.9	A	11-dehydro-2,3-dinor Thromboxane B2	C18H28O6/ C19H24N4O2
12	323.1864	14.8	A	*Cibaric acid*, 12-oxo-14,18-dihydroxy-9Z,13E,15Z-octadecatrienoic acid	C18H28O5
13	643.4802	16.0	A	PE-Cer(d15:2(4E,6E)/19:0	C35H69N2O6P
14	347.1988	17.8	A	2,4-Pyrimidinediamine, 5-(3-amino-4-(methyl(phenylmethyl)amino)phenyl)-6-ethyl-	C19H28N2O4
15	265.1809	18.0	A	10,11-epoxy-3,7,11-trimethyl-2E,6E-tridecadienoic acid	C16H26O3
16	267.1966	19.5	A	(5Z,8Z)-16-Hydroxy-5,8-hexadecadienoic acid	C16H28O3
17	417.1686	14.5	D		
18	301.2023	14.3	A	10-hydroxy-hexadecan-1,16-dioic acid	C16H30O5/ C17H26N4O

Table 3 - List of compounds showing significantly different abundances between sample sets, tentatively identified in positive mode by LC-HRMS. A grey box indicates not identified. A = Ancona, D = Dorset sample set. Compounds in italics are lipids or lipid based molecules

No	Compound mass/z	Retention time	Highest group	Tentative identification	Suggested formula
1	452.3131	17.0	A	PC(P-14:0/0:0)	C22H46NO6P
2	403.2839	19.9	D	1α-hydroxy-26,27-dinorvitamin D3 25-carboxylic acid / 1α-hydroxy-26,27-dinorcholecalciferol 25-carboxylic acid	C25H38O4
3	301.2159	17.6	A	8,11,14-Eicosatetraynoic acid	C20H28O2
4	506.4048	17.3	A		
5	340.3935	18.8	D		
6	269.2108	17.9	A	(5Z,8Z)-16-Hydroxy-5,8-hexadecadienoic acid	C16H28O3
7	299.2002	15.5	A	All-trans-3,4-Didehydro-Retinoic acid	C20H26O2
8	312.3621	17.9	D		
9	293.2107	17.1	A	8-hydroxy-10,12-octadecadiynoic acid	C18H28O3
10	251.2003	17.9	A	9,12,15-hexadecatrienoic acid (16:3)	C16H26O2
11	299.2002	16.4	A	All-trans-3,4-Didehydro-Retinoic acid	C20H26O2
12	171.1490	13.0	D	3-Acrylamidopropyl trimethylammonium	C9H18N2O
13	427.3569	22.8	A	Stigmast-22-ene-3,6-dione; 7-Oxostigmasterol	C29H46O2
14	411.3619	23.4	A	Stigmasta-4,6-dien-3-one; (23R)-isocalysterol; 24-allenyl-cholesterol	C29H46O
15	283.2052	17.8	A	All-trans-Dehydroretinal	C20H26O
16	405.2995	20.3	D	MG(22:5(4Z,7Z,10Z,13Z,16Z)/0:0/0:0)	C25H40O4
17	317.2108	15.5	A	19-Hydroky-all-trans-retinoic acid/ 2,17beta-Dihydroxy-17-methylandrosta-1,4-dien-3-one	C20H28O3
18	151.0964	12.8	A		

Figure 3 - Box and whisker plot for compound *m/z* 452.3131 (from the positive mode LC-HRMS analysis) tentatively identified as the lipid class compound: 1-(1Z-tetradecenyl)-sn-glycero-3-phosphocholine (PC(P-14:0/0:0).

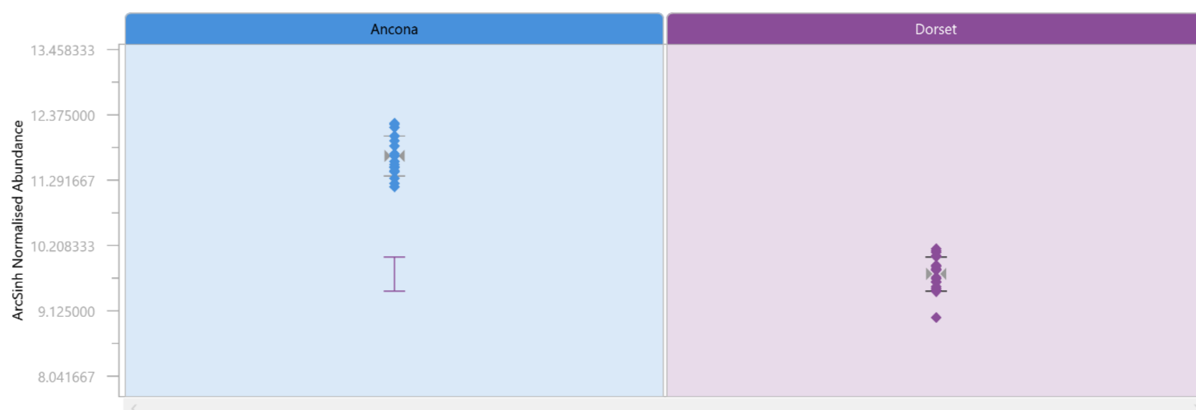
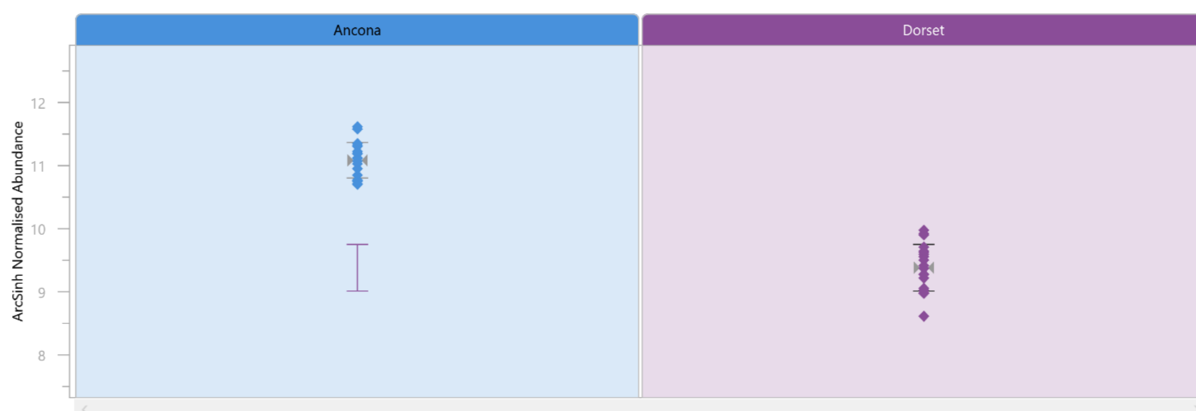


Figure 4 - Box and whisker plot for compound *m/z* 301.2159 (from the positive mode LC-HRMS analysis) tentatively identified as the lipid class compound: 8,11,14-Eicosatetraynoic acid.



4.2. GC-FID analysis, targeted fatty acid determination

Tables 4 and Figure 5 summarise the fatty acid profile of each of the six samples per location. Figure 5 shows an example PCA from this data.

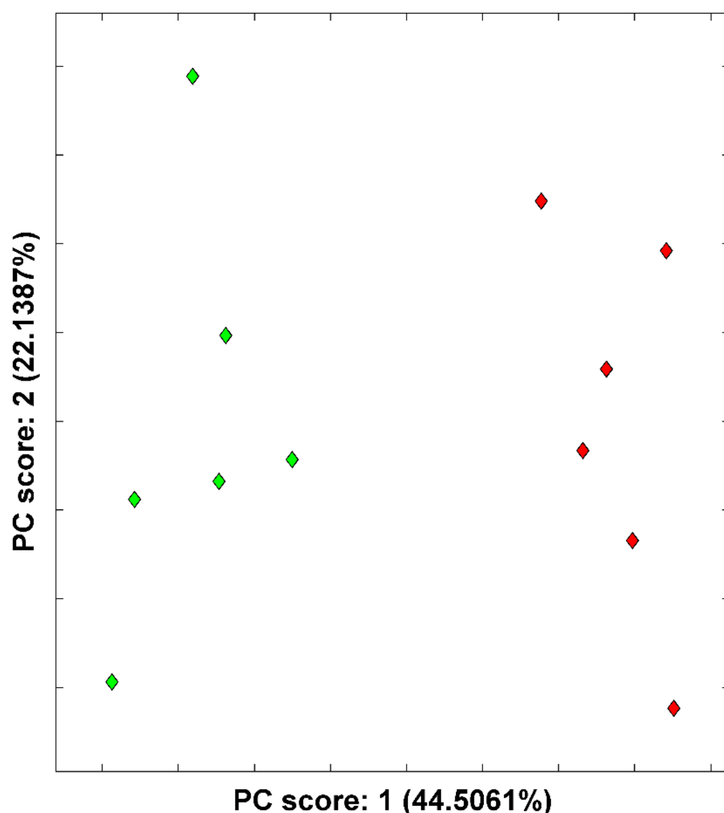
Table 4 - Summary of fatty acids detected and their % of the total, from 6 samples selected from the Ancona sample set

Fatty acid	S19-020941	S19-020932	S19-020935	S19-020938	S19-020925	S19-020929
Caproic acid	1.6	1.5	1.9	1.9	<0.5	0.9
Capriylic acid	2.5	2.6	2.8	3.1	2.9	2.9
Myristic acid	1.5	1.9	0.8	0.8	0.8	<0.5
Palmitic acid	19.0	17.7	17.2	17.8	18.6	17.8
Palmitoleic acid	4.5	4.3	3.7	4.2	4.2	4.0
Heptadecanoic acid	6.9	6.3	7.2	6.5	7.2	7.6
Stearic acid	7.3	6.9	7.1	7.1	7.2	7.4
Oleic acid	2.2	2.1	2.4	2.3	2.4	2.2
Cis-11-Eicosenoic acid	1.7	1.6	1.7	1.8	1.7	0.8
Linolenic acid	1.7	1.7	1.6	1.8	1.7	1.7
Cis-11, 14-Eicosadienoic acid	1.5	1.5	0.8	1.6	1.6	<0.5
Arachidonic acid	1.8	1.6	2.0	1.8	1.8	1.9
Tricosanoic acid	1.8	1.6	1.7	1.7	1.7	1.7
Cis-13,16-Docosadienoic acid	<0.5	<0.5	1.4	0.8	0.9	2.0
Lignoceric acid	9.3	9.3	8.5	8.8	8.7	9.0
Cis-4,7,10,13,16,19-Docosahexaenoic acid	13.2	13.3	14.7	14.1	13.9	14.5
Unknown others	23.4	26.1	24.6	23.9	24.7	25.7
% total saturated	49.9	47.8	47.1	47.7	47.1	47.3
% total monounsaturated	8.4	8.1	7.8	8.3	8.3	7.0
% total polyunsaturated	18.2	18.1	20.5	20.0	20.0	20.1

Table 5 - Summary of fatty acids detected and their % of the total, from 6 samples selected from the Dorset sample set

Fatty acid	S19-032403	S19-032396	S19-032390	S19-032407	S19-032388	S19-032399
Caproic acid	0.8	<0.5	<0.5	1.4	0.9	1.4
Capriylic acid	2.3	2.3	2.4	2.3	2.7	2.5
Myristic acid	1.7	<0.5	<0.5	0.7	<0.5	<0.5
Palmitic acid	22.6	21.4	20.9	20.2	20.8	19.7
Palmitoleic acid	3.3	2.6	2.3	2.7	2.8	2.4
Heptadecanoic acid	5.9	6.5	7.1	6.7	6.4	6.5
Stearic acid	7.5	8.4	8.4	8.0	7.8	7.7
Oleic acid	2.7	2.5	1.2	2.2	2.6	2.6
Cis-11-Eicosenoic acid	2.4	2.4	2.8	2.7	2.4	2.6
Linolenic acid	3.2	2.7	2.8	2.7	2.7	2.5
Cis-11, 14-Eicosadienoic acid	2.1	2.0	2.1	1.9	1.9	1.9
Arachidonic acid	2.0	2.4	2.6	2.5	2.4	2.4
Tricosanoic acid	1.8	1.6	1.8	1.9	1.7	1.7
Cis-13,16-Docosadienoic acid	0.7	2.1	<0.5	0.7	2.2	1.1
Lignoceric acid	8.7	9.4	9.2	8.0	8.9	8.2
Cis-4,7,10,13,16,19-Docosahexaenoic acid	11.9	12.8	15.2	12.7	13.4	13.3
Unknown others	20.4	20.7	21.2	22.9	20.4	23.6
% total saturated	51.4	49.7	49.7	49.1	49.2	47.7
% total monounsaturated	8.3	7.5	6.3	7.7	7.8	7.6
% total polyunsaturated	19.9	22.1	22.7	20.4	22.6	21.2

Figure 5 - PCA of Dorset (green) and Ancona (red) samples from the targeted fatty acid profile data



5. Discussion

5.1. Non-targeted lipid analysis

The PCA's of both data sets show that there are clear biochemical differences between the sample types as received. Although this analysis was not a purely "lipidomics" experiment, the extraction and therefore the subsequent data is steered towards the inclusion of lipid class molecules over other biochemicals. This is seen in the most significant and abundant compounds that are associated with the differences in the sample types, in that nine compounds that have a lipid classification have been tentatively identified as significantly more abundant in the Dorset samples. These are 20-hydroxy-Resolvin E1, 1-eicosanoyl-glycero-3-phosphoserine (PS 20:0/0:0), Cibaric acid, N-(nonadecanoyl)-4E,6E-tetradecasphingadienine-1-phosphoethanolamine (PE-Cer(d15:2(4E,6E)/19:0), 10,11-epoxy-3,7,11-trimethyl-2E,6E-tridecadienoic acid, (5Z,8Z)-16-Hydroxy-5,8-hexadecadienoic acid, 10-hydroxy-hexadecan-1,16-dioic acid, 1-(1Z-tetradecenyl)-sn-glycero-3-phosphocholine (PC(P-14:0/0:0)) and 8,11,14-Eicosatetraynoic acid.

For compounds detected as significantly higher in the Dorset sample set, tentative identifications did not reveal any lipids or lipid related compounds.

5.2. Targeted fatty acid analysis

Visual inspection of the data from the PCA, identified that differences are apparent between sample sets. When applying a t-test (2 tailed distribution) to each of the fatty acids % proportion of the total fat profile the following differences were found:

Palmitic acid (16:0) $p = 0.0001$, Stearic acid (18:0) $p = 0.0006$, Eicosenoic acid (20:1) $p = 0.0001$, Linolenic acid (18:3) $p < 0.0001$ and Eicosadienoic acid (20:2) $p = 0.0040$ are all significantly greater in the Dorset sample's fatty acid profile. Conversely, Caprylic acid (8:0) $p = 0.0100$ and Palmitoleic acid (16:1) $p < 0.0001$ are significantly greater in the Ancona sample's fatty acid profile.

These observations are reflected in the % total polyunsaturated fatty acids, which are significantly greater in the Dorset sample's fatty acid profile ($p = 0.01$).

6. Conclusions and Recommendations

From this targeted and non-targeted proof of principle study the following conclusions can be drawn:

- There is promising evidence to suggest that very different fishing regions such as UK vs. Adriatic based clams can be differentiated based on their fatty acid profiles and / or their biochemical profile.

- Within the differences in the biochemical profiles, lipid-based compounds play a potentially significant contribution.

- Potential biomarkers (tentatively identified) have been discovered that can differentiate clams from the Adriatic Sea compared to the English Channel.

- Both targeted and non-targeted methodologies have been applied in complementation to support this claim

- To be operationally applicable further work is required to establish and validate the approach for different regions. Different species and areas will have different lipid biomarkers however the approach of lipid biochemical profiling looks sound. To be used from today MMO would have to contract a laboratory who are capable of lipid analysis, (targeted fatty acid profiling as a minimum) and provide the lab with known authentic samples from the two (or more) regions of interest, before sending samples for analysis which are under suspicion.

- All conclusions in this report are based on samples received frozen into the laboratory. If samples were handled or stored differently this may create differences in the biochemical profile and / or fatty acid profiles which may account for the differences observed.

Potential next steps are proposed as follows:

1. Identification of the tentatively identified markers in Tables 2 and 3 where analytical standards can be purchased to undertake confirmation analysis.

2. A follow-on study to establish the resolution of this technique in other pressing areas of interest. i.e. in fishing waters closer together, e.g. English Channel vs The North Sea to assess if the same promising results are determined.

3. A follow-on study with greater sample numbers to validate the approach before it can be used for enforcement purposes. This is a proof of principle study with a moderate statistical power. To obtain a high statistical power ($\beta > 0.8$) it is suggested that ≥ 200 sample replicates per region should be analysed.

4. To undertake all three of the recommendations above would be a further study of approximately 6–12 months costing in the region of £50,000.

7. References

Christie, W. (1993). Preparation of ester derivatives of fatty acids for chromatographic analysis. *Advances in Lipid Methodology*, 2:69-111.

Garrido D, Varo I, Morales AE, Hidalgo MC, Navarro JC, Hontoria F, Monroig O, Iglesias J, Otero JJ, Estevez A, Perez J, Martin MV, Rodriguez C, Almansa E and Cardenete G (2017). Assessment of stress and nutritional biomarkers in cultured *Octopus vulgaris* paralarvae: Effects of geographical origin and dietary regime, *Aquaculture*, (468(1):558-568.

Levin, N, Salek, RM and Steinbeck, C. (2016). From “Databases to Big Data” in Holmes, E., Nicholson., J.K., Darzi, A.W. and Lindon, J.C. (Eds). *Metabolic Phenotyping in Personalized and Public Healthcare,* Academic Press (pp 317-331)

Murzina M, Nefedova ZA, Falk-Petersen S and Ripatt PO (2013). Lipid Status of the Two High Latitude Fish Species, *Leptoclinus maculatus* and *Lumpenus fabricii*. *International Journal of Molecular Sciences*.;14(4):7048-7060.