Fatty-acid Profiles and Fingerprints of Seven Types of Fish Roes as Determined by Chemometric Methods

Shuang Ma^{1,2}, Lai Hao Li^{1*}, Shu Xian Hao^{1*}, Xian Qing Yang¹, Hui Huang¹, Jian Wei Cen¹, and Yue Qi Wang¹

¹ Key Laboratory of Aquatic Product Processing, Ministry of Agriculture and Rural Affairs; Guangdong Provincial Key Laboratory of Fishery Ecology Environment; South China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Guangzhou 510300, CHINA

² College of Food Science and Technology, Shanghai Ocean University, Shanghai 201306, CHINA

Abstract: The fatty acids in seven species of fish roes were determined by GC-MS in combination with principal component and cluster analyses in order to derive their fatty-acid profiles and fingerprints. Twenty-three common chromatography peaks were identified in the fatty-acid fingerprints of the seven fish roes. A total of 19 typical fatty acids were identified in the fish roes studied. The fatty acid contents of the roes were significantly different, with saturated-fatty-acid contents in the seven roes ranging from 26.69% to 41.81%, and the unsaturated-fatty-acid contents ranging from 57.65% to 72.21%, the total EPA and DHA content (37.20%) is high in *E. cypselurus* roe, especially. The seven roe species were clearly distinguished according to fatty-acid composition and content by principal component analysis (PCA) and divided into two groups by cluster analysis (CA). PCA of the fatty acid data yielded three significant PCs, which together account for 94% of the total variance; with PC1 contributing 54% of the total.

Key words: fish roe, fatty acid, principal component analysis, cluster analysis, fingerprint

1 Introduction

Fish roe is an important nutritional resource for fish embryo development, which have traditionally been discarded as waste during procession except sturgeon roes¹⁾. As the contribution of roe to the whole body weight of fish may be as high as 30% during spawning seasons²⁾, finding ways to use roes would increase the value of fish and reduce waste. Caviar is the most famous and expensive roe product by salting. The increased consumption of caviar in western countries has led to an increase in the number of sturgeon-breeding farms^{3–5)} and increased interests in fish roe usage, accordingly. Worldwide sturgeon caviar production in 2017 amounted to approximately 364 t⁶⁾.

Studies have revealed that fish roe is rich in nutrients, such as lipids (12.4–14.7 g/100 g roe) and proteins (21.6–28.7 g/100 g roe)^{7,8)}. The nutritional value of fish roe arises from its high content of polyunsaturated fatty acids (PUFAs)⁹⁾, especially those of n-3 polyunsaturated fatty acids (30%–50% of total fatty acids)^{10–12)}. For instance, white sea bream roe contains 4.91%–6.73% eicosapentaenoic acid (EPA) and 23.09%–27.50% docosahexaenoic acid (DHA)¹³⁾, Pacific herring roe contains 13.72% EPA and 21.65% DHA¹⁴⁾, and red salmon roe contains 12.21% EPA

and 17.87% DHA¹⁵⁾.

Owing to the expense of sturgeon caviar, cheaper substitutes prepared using roes from cheaper and more available fish species are becoming increasingly popular. Thus, there is a growing demand for properly labeled commodity groups¹⁶⁾, and roe species identification is a matter of concern. Fatty-acid profiling has proven to be as an effective approach to distinguish fat-rich foods by geographical origin or species identity^{17,18)}. To the best of our knowledge, numerous studies on the classification of the origin of foods according to their fatty-acid profiles have been published^{19,20)}. Furthermore, fatty-acids profiling can provide information on the relationship between fish parasites and their host fatty acid composition patterns²¹⁾, and it can also be used to identify different geographical origins of palm oil²²⁾. However, the majority of research into the fatty acids in fish is mainly concerned with nutritional analysis²³⁾ and the effect of fatty acid on gonadal development and fertility²⁴⁾. Few reports on the application of gas-chromatography-based fatty-acid profiling for the authentication and species identification of roe products have been published.

Fingerprinting is a kind of modern instrument analysis and statistical processing method²⁵⁾ that is becoming in-

*Correspondence to: Lai Hao Li, Key Laboratory of Aquatic Product Processing, Ministry of Agriculture and Rural Affairs; South China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Guangzhou 510300, CHINA. Shu Xian Hao, *ibid*. E-mail: laihaoli@163.com (LHL), susanhao2001@163.com (SXH)

Accepted June 23, 2020 (received for review March 7, 2020)

Journal of Oleo Science ISSN 1345-8957 print / ISSN 1347-3352 online

<http://www.jstage.jst.go.jp/browse/jos/> <http://mc.manuscriptcentral.com/jjocs>

creasingly applied to seafood. For example, Chapman *et al.*²⁶⁾ employed transcriptomic fingerprinting to reveal that developmental ovary dysfunction is highly predictive of egg quality. Furthermore, metabolomic fingerprinting has been used to evaluate the impacts of dietary plant proteins and oils on the serum metabolome of two-year-old gilthead sea bream, *Sparus aurata*²⁷⁾. A complete NMR-based lipid fingerprinting protocol has also been developed for the classification of farmed and wild adult gilthead sea bream²⁸⁾.

Multiple studies on the authentication of food products using fingerprinting techniques have been performed²⁹⁾. However, there are very few recent studies on fatty-acid-composition analysis as a means to identify and authenticate commercially available caviars.

Accordingly, the aim of this study was to establish the fatty-acid profiles and fingerprints of seven kinds of caviar using Gas chromatography-mass spectrometry (GC-MS) in combination with principal component analysis (PCA) and cluster analysis (CA). The method may be used to provide fatty acid standards and as a means monitor the quality and authenticity of fish roes.

2 Materials and Methods

2.1 Sample collection

Roe samples were collected from fish farms and supermarkets in China during the spawning seasons in 2018. The roes of *Salmon solar*, *Exocoetidae cypselurus*, and *Pseudosciaena crocea* were obtained from Fujian province, sturgeon roes (*Acipenser baerii*, *Acipenser gueldenstaedti*, *Acipenser schrenckii*, and *Huso dauricus*) were obtained from Zhejiang province. The maturity of the roes is between in the end of III and in the beginning of IV.

Samples were transported to the laboratory on ice and stored at -18°C until use. Boron trifluoride-methanol (14%, *w/w*) and *n*-hexane were HPLC grade, and were obtained from ANPEL Laboratory Technologies Inc., Shanghai, China. A thirty-seven fatty acid methyl ester (FAME) mix (NU-CHEK, USA) was used for standardization.

2.2 Fatty-acid profiling

Lipids were extracted from the samples using a modified Folch's chloroform methanol extraction method³⁰⁾. First, 5.0 g of the fish roe was homogenized with 15 mL chloroform/methanol (2:1, *v/v*). After 1 h of filtration, 0.2-times volume of normal saline was added to the filtrate. The lower layer was isolated then centrifuged and the concentrated lipid was obtained by evaporation under a gentle nitrogen stream. Then, 2 mL of 14% boron trifluoride-methanol solution was added and the solution was heated at 60°C on a water bath for 30 min. Subsequently, 1 mL of *n*-hexane and distilled water, respectively, were added and, after static stratification, the upper liquid was analyzed by

GC-MS (2010 PLUS GC-MS, Shimadzu, Japan).

The GC conditions were: DB-5MS column (30 m $\times$ 0.25 mm, 0.25 μm); injection port temperature 230°C , temperature program of 110°C for 4 min, increasing to 160°C at $10^{\circ}\text{C min}^{-1}$, maintained for 1 min, then increasing to 240°C at $5^{\circ}\text{C min}^{-1}$ and maintained for 15 min. The carrier gas was helium at a flow rate of $1.52 \text{ mL} \cdot \text{min}^{-1}$ (constant linear velocity). The split ratio was 1:30, and the injection volume was 1 μL .

The precision of the GC fingerprint method was determined by analyzing the same sample three times within a day. Sample stability was determined by analyzing one sample three times on one day at different time intervals (0, 6, 12, and 24 h).

The identification and quantification of the individual fatty acids were performed based on the NIST 08 mass spectra library and by comparing their retention times with those of the standards (37 FAME mix). Relative concentrations were calculated and expressed as percentages of the total identified fatty acids.

2.3 Establishment of GC fingerprints

To obtain GC fingerprints, the chromatographic data were imported into the professional software package Similarity Evaluation System for Chromatographic Fingerprint of Traditional Chinese Medicine (Version 2004A). The chromatogram map used as the reference map was S7, and the time window width was set at 0.1 min. Then, three parallel chromatographic data set for each sample were automatically matched (multi-point correction was adopted) to establish a common pattern for the GC fingerprints, which was a reference fingerprint of one sample. Using the same method, seven batches of reference fingerprints were automatically matched to establish a common pattern for the GC reference fingerprints of fish roes and their similarities which is useful information for quality control.

2.4 Statistical analysis

Seven kinds of fish roe were analyzed using the established GC-MS method. Every sample analysis was carried out in triplicate. The data are presented as mean value $\pm$ standard deviation (Mean $\pm$ SD). Statistical analysis of comparisons of the groups was performed by Duncan's test in IBM SPSS 20. CA and PCA were used to distinguish the different fatty-acid compositions of the fish roes. CA and PCA were performed using IBM SPSS 20 software and Unscrambler 9.7.

3 Results and Discussion

3.1 Fatty acid composition of fish roes

The fatty-acid profiles, saturated fatty acids (SFAs), monounsaturated fatty acids (MUFAs) and PUFAs of the

Table 1 Fatty-acid compositions and relative contents for the seven roes (%).

Fatty acid	<i>S. solar</i>	<i>E. cypselurus</i>	<i>P. crocea</i>	<i>A. baerii</i>	<i>A. schrenckii</i>	<i>A. gueldenstaedti</i>	<i>H. dauricus</i>
myristic acid (C14:0)	2.71 ± 0.26 ^d	1.54 ± 0.01 ^b	1.96 ± 0.16 ^c	0.48 ± 0.01 ^a	0.45 ± 0.01 ^a	0.55 ± 0.02 ^a	0.56 ± 0.04 ^a
pentadecanoic acid (C15:0)	0.54 ± 0.03 ^d	0.31 ± 0.02 ^c	1.14 ± 0.13 ^c	0.20 ± 0.01 ^{ab}	0.18 ± 0.00 ^a	0.21 ± 0.01 ^{ab}	0.25 ± 0.01 ^b
palmitic acid (C16:0)	12.65 ± 0.74 ^a	15.26 ± 0.33 ^b	25.54 ± 1.76 ^c	18.40 ± 0.91 ^c	19.19 ± 0.18 ^{cd}	18.83 ± 0.56 ^c	20.34 ± 0.16 ^d
heptadecanoic acid (C17:0)	1.08 ± 0.04 ^d	0.64 ± 0.03 ^a	1.74 ± 0.04 ^c	0.78 ± 0.07 ^b	0.90 ± 0.02 ^c	0.93 ± 0.04 ^c	0.92 ± 0.05 ^c
stearic acid (C18:0)	9.43 ± 0.44 ^c	8.82 ± 0.20 ^b	11.01 ± 0.43 ^c	9.10 ± 0.05 ^{bc}	10.28 ± 0.03 ^d	9.97 ± 0.39 ^d	8.25 ± 0.01 ^a
nonadecanoic acid (C19:0)	0.08 ± 0.01 ^a	0.17 ± 0.01 ^c	0.28 ± 0.03 ^c	0.20 ± 0.00 ^c	0.12 ± 0.03 ^b	0.20 ± 0.02 ^d	0.10 ± 0.01 ^{ab}
arachidic acid (C20:0)	0.07 ± 0.01 ^{ab}	0.25 ± 0.01 ^b	0.15 ± 0.04 ^{ab}	0.13 ± 0.01 ^{ab}	0.11 ± 0.03 ^{ab}	0.13 ± 0.01 ^{ab}	0.09 ± 0.00 ^a
SFAs	26.69 ± 1.10 ^a	26.99 ± 0.28 ^a	41.81 ± 1.63 ^d	29.27 ± 0.72 ^b	31.23 ± 0.07 ^c	30.73 ± 0.69 ^{bc}	30.51 ± 0.19 ^{bc}
palmitoleic acid (C16:1)	4.33 ± 0.15 ^d	8.02 ± 0.24 ^c	3.80 ± 0.32 ^c	2.56 ± 0.05 ^b	1.94 ± 0.13 ^a	2.49 ± 0.10 ^b	2.47 ± 0.09 ^b
heptadecenoic acid (C17:1)	0.76 ± 0.06 ^d	0.76 ± 0.03 ^d	0.66 ± 0.02 ^c	0.38 ± 0.02 ^a	0.42 ± 0.04 ^a	0.43 ± 0.01 ^a	0.51 ± 0.03 ^b
oleic acid (C18:1)	16.64 ± 0.04 ^c	14.78 ± 0.57 ^b	12.75 ± 0.54 ^a	22.39 ± 0.22 ^c	21.46 ± 0.05 ^d	21.50 ± 0.05 ^d	21.96 ± 0.27 ^{cd}
eicosenoic acid (C20:1)	2.66 ± 0.14 ^c	1.61 ± 0.22 ^{ab}	1.57 ± 0.64 ^{ab}	1.44 ± 0.08 ^{ab}	1.17 ± 0.05 ^a	1.40 ± 0.03 ^{ab}	1.70 ± 0.02 ^{ab}
erucic acid (C22:1)	0.18 ± 0.06 ^b	nd	0.21 ± 0.01 ^a	0.20 ± 0.05 ^a	nd	nd	0.08 ± 0.04 ^a
MUFAs	24.57 ± 0.38 ^b	25.16 ± 0.63 ^{bc}	19.00 ± 1.31 ^a	26.97 ± 0.05 ^c	24.98 ± 0.14 ^{bc}	25.82 ± 0.11 ^{cd}	26.72 ± 0.30 ^d
linoleic acid (C18:2)	1.10 ± 0.08 ^a	1.94 ± 0.17 ^b	1.43 ± 0.18 ^a	6.80 ± 0.22 ^d	8.62 ± 0.72 ^c	6.47 ± 0.25 ^d	4.71 ± 0.17 ^c
linolenic acid (C18:3)	0.05 ± 0.01 ^a	0.09 ± 0.00 ^b	0.10 ± 0.00 ^b	0.29 ± 0.00 ^d	0.54 ± 0.01 ^c	0.25 ± 0.01 ^c	0.32 ± 0.01 ^c
eicosatrienoic acid (C20:3)	0.24 ± 0.02 ^b	0.22 ± 0.03 ^{ab}	0.18 ± 0.02 ^a	0.37 ± 0.01 ^c	0.78 ± 0.03 ^d	0.38 ± 0.02 ^c	2.05 ± 0.05 ^c
arachidonic acid (C20:4)	3.34 ± 0.25 ^a	3.55 ± 0.19 ^a	3.65 ± 0.17 ^a	4.71 ± 0.14 ^b	6.82 ± 0.27 ^d	4.69 ± 0.06 ^b	5.47 ± 0.02 ^c
eicosapentaenoic acid (C20:5)	14.90 ± 0.14 ^d	10.88 ± 0.42 ^c	7.20 ± 0.21 ^a	8.91 ± 0.26 ^b	7.10 ± 0.36 ^a	9.02 ± 0.08 ^b	8.67 ± 0.29 ^b
docosapentaenoic acid (C22:5)	7.06 ± 0.59 ^d	3.33 ± 0.11 ^c	3.29 ± 0.56 ^c	1.91 ± 0.25 ^{ab}	1.49 ± 0.09 ^{ab}	1.95 ± 0.14 ^a	2.20 ± 0.03 ^b
docosahexaenoic acid (C22:6)	20.92 ± 0.49 ^b	26.05 ± 0.22 ^d	22.79 ± 0.60 ^c	20.75 ± 0.10 ^b	18.37 ± 0.27 ^a	20.03 ± 0.19 ^b	20.94 ± 0.40 ^b
EPA+DHA	35.83 ± 0.60 ^c	37.20 ± 0.62 ^c	29.99 ± 0.53 ^b	29.67 ± 0.18 ^b	25.50 ± 0.17 ^a	29.08 ± 0.22 ^b	29.62 ± 0.68 ^b
PUFAs	47.64 ± 1.49 ^d	46.33 ± 0.72 ^{cd}	38.65 ± 2.90 ^a	43.73 ± 0.67 ^b	43.78 ± 0.11 ^b	42.81 ± 0.58 ^b	44.38 ± 0.50 ^{bc}

Note: Different superscript letters within the same row indicate significant differences among samples ($p < 0.05$).

nd: undetected.

seven fish roes are presented in **Table 1**. Nineteen fatty acids (7 SFAs, 5 MUFAs, and 7 PUFAs) were identified in the roes of *S. solar*, *P. crocea*, *A. baerii*, and *H. dauricus*. Eighteen fatty acids were identified in the roes of *E. cypselurus*, *A. schrenckii*, and *A. gueldenstaedti*. The MUFA contents of the fish roes are significantly higher than the SFA and PUFA contents ($p < 0.05$). The fatty-acid profiles of the different fish roes are very similar, although differences exist in the relative amounts of individual components³¹.

The range of total SFA contents of the fish roes are 26.69%–41.81%, with that of the *P. crocea* being 41.81%, which is significantly higher than that of the other roes ($p < 0.05$). There is no significant difference in the relative SFA contents of *A. schrenckii*, *A. gueldenstaedti*, and *H. dauricus* roes ($p > 0.05$). *S. solar* roe has the lowest relative SFA content (26.69%). The SFAs in the roes are mostly 14–20 carbon species. The majority of the SFAs are C16:0 and C18:0, and there are large amounts of these SFAs in *P. crocea* roe (25.54% and 11.01%, respectively). Mohamadi Monavar *et al.*³² reported that the relative content of C16:0 in Beluga caviar (24.27%) is lower than that of *P. crocea* roe.

The carbon numbers of the MUFAs in the roes range

from 16 to 22. The most abundant MUFA is C18:1, which is present at relative contents from 12.75% to 22.39%. The second most abundant is C16:1, the relative contents of which range from 1.94% to 8.02%. C22:1 is present in *S. solar*, *P. crocea*, *A. baerii*, and *H. dauricus* roes. Heptadecenoic acid (C17:1) and Eicosenoic acid (C20:1) are quantified at concentrations of below 3%. The relative content of MUFAs in *A. baerii* roe (26.97%) is significantly higher than that of the other roes ($p < 0.05$). Except for that of *P. crocea*, the range of MUFA contents for the other six roes is 24.57%–26.97%.

The PUFAs in roe mainly include C18:2, C18:3, C20:3, C20:4, C20:5, C22:5, and C22:6. EPA (C20:5) and DHA (C22:6), which are widely recognized as helping to reduce incidences of coronary heart diseases, inflammatory and autoimmune disorders, and cancer, are abundant in roes^{13, 33, 34}.

The EPA level of *S. solar* roe (14.90%) is significantly higher ($p < 0.05$) than those of the other types of roes. *E. cypselurus* roe is the second most rich in EPA. The DHA content of the *E. cypselurus* roe is 26.05%; those of *P. crocea* and *S. solar* roes are 22.79% and 20.92%, respectively; and those of *A. baerii*, *A. gueldenstaedti*, and *H. dauricus* roes are 20.75%, 20.03%, and 20.94%, respectively. The lowest level of C22:6 measured in our study is

that for *A. schrenckii* roe. However, this is higher than that for Imperial black caviar (10.50%)⁸⁾. In general, the PUFA contents are slightly different between the seven roes ($p < 0.05$).

The relative contents of total PUFAs ranges from 38.65% to 47.64%. Table 1 shows the PUFA contents of *S. solar* roe (47.64%), *E. cypselurus* roe (46.33%), *H. dauricus* roe (44.38%), *A. schrenckii* roe (43.78%), *A. baerii* roe (43.73%), *A. gueldenstaedti* roe (42.81%), and *P. crocea* roe (38.65%). However, whereas lower values (around 26.70%–29.10%) have been previously reported for the roe of the five specimens of *Silurus glanis*³⁵⁾. This phenomenon may be due to by the different of roe species used in that study.

3.2 Principal component analysis

PCA is the most popular technique for elucidating the pertinent information hidden in a data matrices and remarkably improves the classification of the various samples by grouping them into distinguishable clusters and sub-clusters^{36–38)} that they seem to have the greatest similarity. The principal components (PCs) are new variables by statistical transformations involving the diagonalization of the variable's correlation or variance–covariance matrix, which provide better interpretations of data³⁹⁾.

The fatty acid data matrix consists of 21 rows (samples) and 18 variables (common fatty acids). PCA on the fatty acid data revealed the presence of some natural clustering (Fig. 1). This clustering is in agreement with the fatty acids of all seven roes. Figure 1 shows the three dimensional principal component analysis (3D-PCA) score plot, which explains 94% of the total data variance and clearly distinguishes the presence of the seven clusters, i.e., *S. solar* roe (S1), *E. cypselurus* roe (V2), *P. crocea* roe (Y3), *A. baerii* roe (S4), *A. schrenckii* roe (A5), *A. gueldenstaedti* roe (R6), and *H. dauricus* roe (D7). The first prin-

cipal component (PC1) represents 54% of the variance, and the next two principal components represent 32% (PC2) and 8% (PC3) of the variance.

As can be seen from Fig. 1, the fatty-acid compositions of the seven roes are significantly distinguished in the principal component score plot. *S. solar* roe is located at the left rear and lower part of the whole area; *E. cypselurus* roe is located at the right rear and upper part of the whole area; *P. crocea* roe is located at the left front and lower part of the whole area. Roies of the three sturgeons and *H. dauricus* are relatively close to the origin, but still well distinguished. *P. crocea* roe shows more variability, while the other samples tend to form narrower clusters.

The three principal component scores for *S. solar* roe are negative, indicating that its fatty acid contents are negatively correlated with the three principal components. *E. cypselurus* roe has a negative correlation with PC1, a positive correlation with PC2 and PC3, and higher positive correlation with PC3 than PC2. *P. crocea* roe has a negative correlation with PC1 and PC3 and a highly positive correlation with PC2.

The three sturgeon roes and *H. dauricus* roe are very close to each other in the 3D-PCA score plot, suggesting that their fatty acid compositions are similar. One of the reasons for this is that the three sturgeon and *H. dauricus* belong to the sturgeon genus. Furthermore, it has been speculated that fatty acid profiles are related to the similar living environments of the female fish⁴⁰⁾. There is a relatively high similarity between Siberian roe and *H. dauricus* roe, but a relatively lower negative correlation is observed for PC2 of the *H. dauricus* roe. The similarity between *A. gueldenstaedti* roe and *H. dauricus* roe is relatively high, but there is a relatively negative correlation in *A. gueldenstaedti* roe for PC3, allowing it to be distinguished from *H. dauricus* roe.

The loadings plots of PC2 versus PC1 and PC3 versus PC1, are shown in Fig. 2a and Fig. 2b, respectively. The contributions of the variables with respect to the three PCs in two-dimensional space and their corresponding correlations are clearly displayed⁴¹⁾. PC1 is strongly associated with the concentrations of C14:0, C16:1, C17:1, C18:1, C18:2, C18:3, C20:4, C20:5, C22:5 and C22:6. The dominant variables in PC2 are C15:0 and C16:0. PC3 may be defined by C20:4 and C22:6.

By comparing the PCA scores plot (Fig. 1) with the PCA loadings plots (Figs. 2a and 2b) for each sample, *S. solar* and *P. crocea* fish roe can be easily distinguished since they are inversely correlated along the PC2 axis, which is contributed to by C15:0 and C16:0. C22:6 and C20:4 have higher contributions to *E. cypselurus* roe. Other fatty acids, such as C14:0, C18:1, and C20:4, make an important contribution to discriminating all the different roes because ca. 54% of the variance is effectively described by PC1.

A positive factor loading indicates that the factor will be

Fig. 1 Principal component analysis plot for fatty-acids in different species of fish roe.

x-axis: PC1; y-axis: PC2; z-axis: PC3.

S1: *S. solar*; V2: *E. cypselurus*; Y3: *P. crocea*; S4: *A. baerii*; A5: *A. schrenckii*; R6: *A. gueldenstaedti*; D7: *H. dauricus*.

Fig. 2 Principal component analysis loadings plot for the fatty-acids in different species of fish roes (a) and (b).

higher in the positive axis of that PC. For example, for C22:6, a factor loading of 0.59 is obtained with PC3, which means that the samples located in the superjacent side (e.g., *E. cypselurus* fish roe) of the graph have higher C22:6 values than the samples located in the inferior side (e.g., *A. schrenckii* fish roe). Similarly, a negative factor loading indicates that the factor will be higher in the positive axis of that PC. For example, for C17:1, a factor loading of -0.95 in PC1, meaning that the samples located in the forward side (e.g., *P. crocea* fish roe) of the graph have lower mean concentrations than the samples located in the backward side (e.g., *S. solar* fish roe), the content of fatty acids as the same with the result of Table 1.

Clear segregation among roes at different species is observed, demonstrating that PCA can differentiate roe species through fatty-acid compositions and contents. It has been well documented that the fatty-acid compositions of fish roes reflect the fatty-acid contents of the lipid in the diet supplied to the broodstock^{13, 42–44}. *S. solar*, *E. cypselurus*, *P. crocea* are all mariculture fish and their growing environment is different from that of the four species of freshwater sturgeon. Therefore, the PCA scores for the three kinds of fish roe above are significantly different from those of the four kinds of sturgeon fish roe.

3.3 Cluster analysis

CA is a way of identifying and expressing data in exploratory data analysis (EDA) so as to emphasize sample similarities and differences^{45, 46}. Data standardization was performed out on 18 common fatty-acid types and contents of the seven fish roes using IBM SPSS 20. The relative content of each common fatty acid was analyzed using a hierarchical clustering method and Pearson correlation, and the hierarchical tree diagram shown in Fig. 3 was generated among the average connecting groups.

Figure 3 reveals the similarity of fatty acids between different species of roe. It is possible to identify the presence of two major clusters. The three parallel roe samples form a cluster at the minimum distance, suggesting that the similarity between the parallel samples is very high.

The results indicate that the *A. baerii* and *A. gueldenstaedti* roes are similar and belong to one cluster with a distance of ca. 4, which suggests that their similarity is extremely high. When the distance increases to ca. 7 and 13, *A. schrenckii* and *H. dauricus* roes respectively group with *A. baerii* and *A. gueldenstaedti* roes, showing that Acipenser roes exhibit high similarity in fatty-acid compositions. With the increase of distance to 12, *S. solar* and *E. cypselurus* are grouped, while *P. crocea* roe is grouped into a class at a distance of ca. 17, which explains why the

Fig. 3 Cluster analysis dendrogram for the fatty-acid profiles of the seven fish roes.

Note: 1, 2, and 3: *S. solar*; 4, 5, and 6: *E. cypselurus*; 7, 8, and 9: *P. crocea*; 10, 11, and 12: *A. baerii*; 13, 14, and 15: *A. schrenckii*; 16, 17, and 18: *A. gueldenstaedti*; 19, 20, and 21: *H. dauricus*.

three roes have a certain similarity in fatty-acid composition.

4 Analysis of GC Fingerprints

4.1 GC fingerprint method validation

The precision of the GC fingerprint method was evaluated by analyzing the same sample three times within a day. The sample stability test was performed using one sample three times on one day at different time intervals (0, 6, 12, and 24 h). The repeatability was assessed by analyzing five independently different samples. According to our results, the precision, repeatability, and 24 h stability of our method all conform to the requirements of methodology validation. The relative standard deviations (RSDs) of retention time for each common fingerprint peak are less than 0.20%, 0.10%, and 1% and the RSDs of the relative peak area are less than 3%, demonstrating the instrument precision and stability of the test sample solutions, as well as the repeatability of the method.

4.2 Establishment of common peaks and similarity analysis

The fatty-acid profiles obtained for the seven fish roes are shown in Fig. 4. Based on the chromatographic fingerprint, a total of 23 common peaks were identified by the

GC × NIST Library (Figs. 4 and 5, Table 2), which cover more than 90% of the total area.

It should be noted that some single compounds might produce multiple signals or isomers. These are represented by letters in Table 2, and corresponding isomers are represented by "[number x]". All common peaks were identified by comparing their retention times and standard structures with the NIST database and the 37-FAME standard. There are isomers for three fatty acids, i.e., C17:0 (f), C18:1 (j), and C20:4 (m). Peak 4 (rt = 18.267 s) has a considerably high content, making up more than 10% of the total area, and it also has a moderate retention time, stable peak area, and good shape. Therefore, it was chosen as the reference peak (S). Then, the retention times and peak areas of the 23 common peaks were measured, relative retention times (RRTs) and relative peak areas (RPAs) of all the characteristic common peaks were calculated with respect to this reference peak (Table 2).

The similarities of the roes range from 0.945 to 0.994, which demonstrates that the seven fish roes share coefficients of similarities. Therefore, this common pattern could be applied as a reference GC-MS fingerprint to identify and assess fish roe.

The detection method used in this experiment has good precision and stability, and the GC fingerprint obtained has obvious variation characteristics, which can be distinguished by PCA and CA, and the variation rules remain

Fig. 4 GC-MS fatty acid fingerprints for the seven fish roes.

s1: *S. solar*; s2: *E. cypselurus*; s3: *P. crocea*; s4: *A. baerii*; s5: *A. schrenckii*; s6: *A. gueldenstaedti*; s7: *H. dauricus*; R: reference fingerprint.

Fig. 5 GC-MS reference fingerprint for the seven fish roes.

stable and exclusive in 7 kinds of roe. Therefore, it is suggested to take the characteristic peak of GC fingerprint as the chromatographic identification index of roes. This study fills the gap of fish roe fingerprint information, further improves the database of fatty acids of aquatic products, and lays a foundation for the establishment of fish roe identification technology system.

5 Conclusions

The significant differences in the fatty-acid compositions and contents of seven fish roe samples were identified using GC-MS. The seven roes could be clearly distinguished

by PCA. On the basis of CA, seven roes were divided into two groups: group 1 contained *S. solar*, *E. cypselurus* and *P. crocea* roes, while group 2 contained *A. baerii*, *A. schrenckii*, *A. gueldenstaedti*, and *H. dauricus* roes. These results indicate that the chemometric methods (PCA and CA) performed on fatty-acid profiles are viable tools for identifying the species of fish roe. Based on a reference fingerprint, which represents comprehensive fatty-acid information from seven species of roe, a fingerprint comprising 23 common chromatography peaks was identified for the seven kinds of fish roe.

A "fingerprint" of the process involved in fish roe, PCA and CA together can make a precise distinction to identify species in roes. This may be a very promising method to be

Table 2 Information for common peaks in the GC-MS fingerprints of the seven roes.

Number	Fatty acid	Retention time	Relative peak areas (RPAs)							Reference fingerprint	Relative retention time RSD (%)	Relative peak areas RSD (%)
			<i>S. solar</i>	<i>E. cypselurus</i>	<i>P. crocea</i>	<i>A. baerii</i>	<i>A. schrenckii</i>	<i>A. gueldenstaedtii</i>	<i>H. dauricus</i>			
a	myristic acid (C14:0)	0.760	0.290	0.144	0.100	0.035	0.030	0.032	0.034	0.091	0.110	1.016
b	pentadecanoic acid (C15:0)	0.878	0.049	0.021	0.052	0.012	0.010	0.012	0.012	0.022	0.080	0.773
c	palmitoleic acid (C16:1)	0.973	0.303	0.531	0.127	0.125	0.079	0.114	0.095	0.196	0.040	0.843
d	palmitic acid (C16:0)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	0.000
e	heptadecenoic acid (C17:1)	1.059	0.024	0.049	0.010	0.028	0.036	0.034	0.037	0.032	0.330	0.382
f	heptadecanoic acid (C17:0)	1.071	0.027	0.027	0.011	0.015	0.017	0.017	0.009	0.018	0.450	0.408
f ^[1]	heptadecanoic acid (C17:0)	1.082	0.023	0.013	0.004	0.005	0.002	0.005	0.015	0.009	0.470	0.811
g	heptadecenoic acid (C17:1)	1.090	0.041	0.025	0.015	0.013	0.010	0.014	0.013	0.019	0.040	0.574
f ^[2]	heptadecanoic acid (C17:0)	1.119	0.051	0.031	0.062	0.023	0.024	0.026	0.026	0.033	0.030	0.441
h	linolenic (C18:3)	1.181	0.003	0.006	0.004	0.014	0.026	0.011	0.014	0.011	0.040	0.732
i	linoleic acid (C18:2)	1.201	0.083	0.127	0.052	0.328	0.399	0.330	0.228	0.231	0.050	0.618
j	oleic acid (C18:1)	1.209	0.865	0.785	0.388	0.963	0.911	0.917	0.885	0.834	0.060	0.241
j ^[1]	oleic acid (C18:1)	1.215	0.356	0.207	0.093	0.210	0.161	0.216	0.171	0.204	0.040	0.396
k	stearic acid (C18:0)	1.240	0.750	0.601	0.435	0.512	0.534	0.546	0.410	0.542	0.030	0.209
l	nonadecanoic acid (C19:0)	1.379	0.005	0.008	0.008	0.008	0.005	0.008	0.004	0.007	0.040	0.311
m	arachidonic acid (C20:4)	1.427	0.081	0.124	0.087	0.164	0.211	0.155	0.142	0.141	0.040	0.331
n	eicosapentaenoic acid (C20:5)	1.439	0.788	0.538	0.199	0.325	0.263	0.349	0.279	0.393	0.040	0.522
o	eicosatrienoic acid (C20:3)	1.456	0.013	0.009	0.004	0.016	0.026	0.014	0.015	0.014	0.040	0.479
m ^[1]	arachidonic acid (C20:4)	1.467	0.106	0.035	0.010	0.015	0.013	0.014	0.013	0.029	0.030	1.182
p	eicosenoic acid (C20:1)	1.494	0.070	0.056	0.018	0.047	0.035	0.044	0.048	0.046	0.040	
q	arachidic acid (C20:0)	1.534	0.004	0.012	0.004	0.005	0.004	0.005	0.003	0.005	0.040	0.590
r	docosahexaenoic acid (C22:6)	1.705	0.563	1.020	0.596	0.759	0.621	0.702	0.663	0.711	0.060	0.220
s	docosapentaenoic acid (C22:5)	1.722	0.421	0.140	0.060	0.062	0.042	0.067	0.065	0.119	0.020	1.104
Similarity			0.945	0.970	0.955	0.991	0.984	0.994	0.991	1.000		

Note: corresponding isomers were represented by "number x]",

utilized in the future. But more researches should be carried out to ensure good reliability of this approach.

Acknowledgements

This research was supported by supported by National Natural Science Foundation of China (31772051), China Agriculture Research System (CARS-46) and Special Scientific Research Funds for Central Non-profit Institutes and Chinese Academy of Fishery Sciences (2020TD69, 2020TD73).

References

- Wang, Y.; Yu, J.X.; Zhang, C.L. *et al.* Influence of flavonoids from *Phellinus igniarius* on sturgeon caviar: Antioxidant effects and sensory characteristics. *Food Chem.* **131**, 206-210 (2012).
- Binsi, P.K.; Nayak, N.; Sarkar, P.C.; Sahu, U.; Lalitha, K.V.; Ninan, G.; Ravishankar, C.N. Conversion of carp roe mass to caviar substitutes: Stabilization with oregano extract. *LWT-Food Sci. Technol.* **108**, 446-455 (2019).
- Jelodar, A.; Safari, R. Microbial and chemical quality evaluation of caviar in Iranian processing plants in line with the European Community code. *J. Appl. Ichthyol.* **22**, 411-415 (2006).
- Safari, R.; Yosefian, M. Changes in TVN (Total Volatile Nitrogen) and psychrotrophic bacteria in Persian sturgeon Caviar (*Acipenser persicus*) during processing and cold storage. *J. Appl. Ichthyol.* **22**, 416-418 (2006).
- Hao, S.X.; Wei, Y.; Li, L.H.; Yang, X.Q.; Cen, J.W.; Huang, H.; Lin, W.L.; Yuan, X.M. The effects of different extraction methods on composition and storage stability of sturgeon oil. *Food Chem.* **173**, 274-282 (2015).
- Bronzi, P.; Chebanov, M.; Michaels, J.T.; Wei, Q.; Rosenthal, H.; Gessner, J. Sturgeon meat and caviar production: Global update 2017. *J. Appl. Ichthyol.* **35**, 257-266 (2019).
- Rehbein, H. Caviar, proximate composition, amino acid content and identification of fish species. *Z. Lebensm. Unters. Forsch.* **180**, 457-462 (1985).
- Ma, S.; Hao, S.X.; Li, L.H.; Yang, X.Q.; Huang, H.; Wei, Y.; Cen, J.W.; Zhai, H.L. Comparative analysis of nutritional components of several roes. *South China Fish Sci.* **15** (04), 113-121 (2019) (in Chinese).
- Binsi, P.K.; Viji, P.; Panda, S.K.; Mathew, S.; Zynudheen, A.A.; Ravishankar, C.N. Characterisation of hydrolysates prepared from engraved catfish (*Nemapteryx caelata*) roe by serial hydrolysis. *J. Food Sci. Tec.* **53**, 158-170 (2016).
- Bledsoe, G.E.; Bledsoe, C.D.; Rasco, B. Caviars and fish roe products. *Crit. Rev. Food Sci.* **43**, 317-356 (2003).
- Moriya, H.; Kuniminato, T.; Hosokawa, M.; Fukunaga, K.; Nishiyama, T.; Miyashita, K. Oxidative stability of salmon and herring roe lipids and their dietary effect on plasma cholesterol levels of rats. *Fish Sci.* **73**, 668-674 (2007).
- Shirai, N.; Higuchi, T.; Suzuki, H. Analysis of lipid classes and the fatty acid composition of the salted fish roe food products, Ikura, Tarako, Tobiko and Kazunoko. *Food Chem.* **94**, 61-67 (2006).
- Cejas, J.R.; Almansa, E.; Villamandos, J.E.; Badía, P.; Bolaños, A.; Lorenzo, A. Lipid and fatty acid composition of ovaries from wild fish and ovaries and eggs from captive fish of white sea bream (*Diplodus sargus*). *Aquaculture* **216** (1-4), 299-313 (2003).
- Huynh, M.D.; Kitts, D.D.; Hu, C.; Trites, A.W. Comparison of fatty acid profiles of spawning and non-spawning pacific herring, *Clupea harengus pallasii*. *Comp. Biochem. Physiol. B: Biochem. Mol. Biol.* **146**, 504-511 (2007).
- Mol, S.; Turan, S. Comparison of proximate, fatty acid and amino acid compositions of various types of fish roes. *Int. J. Food Prop.* **11**, 669-677 (2008).
- Consonni, R.; Cagliani, L. Chapter 4-Nuclear magnetic resonance and chemometrics to assess geographical origin and quality of traditional food products. *Advances in Food and Nutrition Res.* **59**, 87-165 (2010).
- Destailats, F.; Cruz-Hernandez, C.; Giuffrida, F.; Dionisi, F. Identification of the botanical origin of pine nuts found in food products by Gas-Liquid chromatography analysis of fatty acid profile. *J. Agric. Food Chem.* **58**, 2082-2087 (2010).
- Esteki, M.; Farajmand, B.; Amanifar, S.; Barkhordari, R.; Ahadiyan, Z.; Dashtaki, E.; Mohammadlou, M.; Heyden, Y.V. Classification and authentication of Iranian walnuts according to their geographical origin based on gas chromatographic fatty acid fingerprint analysis using pattern recognition methods. *Chemom. Intell. Lab. Syst.* **171**, 251-258 (2017).
- Yi, L.; Wu, H.; Liang, Y.; Chromatographic fingerprint and quality control of traditional Chinese medicines. *Chin. J. Chromatogr.* **26**, 166-171 (2008).
- Rutkowska, J.; Bialek, M.; Adamska, A.; Zbikowska, A. Differentiation of geographical origin of cream products in Poland according to their fatty acid profile. *Food Chem.* **178**, 26-31 (2015).
- Taraschewski, H.; Aitzetmüller, K.; Werner, G.; Kühs, K. Lipids of fish parasites and their hosts: Fatty acid fingerprints of four species of acanthocephalans and of their hosts' intestinal tissues. *Parasitol. Res.* **81**, 522 (1995).

- 22) Tres, A.; Ruiz-Samblas, C.; Van, D.V.G.; Van Ruth, S.M. Geographical provenance of palm oil by fatty acid and volatile compound fingerprinting techniques. *Food Chem.* **137**(1-4), 142-150(2013).
- 23) Turchini, G.M.; Hermon, K.M.; Francis, D.S. Fatty acids and beyond: fillet nutritional characterisation of rainbow trout (*Oncorhynchus mykiss*) fed different dietary oil sources. *Aquaculture* **491**, 391-397(2018).
- 24) Ji, H.; Li, J.; Liu, P. Regulation of growth performance and lipid metabolism by dietary n-3 highly unsaturated fatty acids in juvenile grass carp. *Ctenopharyngodon idellus*. *Comp. Biochem. Physiol. B: Biochem. Mol. Biol.* **159**, 49-56(2011).
- 25) Meng, J.; Leung, S.Y.; Jiang, Z.H.; Dong, X.P.; Zhao, Z.Z.; Xu, L.J. Establishment of HPLC-DAD-MS fingerprint of fresh *Houttuynia cordata*. *Chem. Pharm. Bull.* **53**, 1604-1609(2005).
- 26) Chapman, R.W.; Reading, B.J.; Sullivan, C.V. Ovary transcriptome profiling via artificial intelligence reveals a transcriptomic fingerprint predicting egg quality in striped bass. *Morone saxatilis*. *Plos One* **9**(5), e96818(2014).
- 27) Gil-Solsona, R.; Caldach-Giner, J.A.; Nácher-Mestre, J.; Lacalle-Bergeron, L.; Sancho, J.V.; Hernández, F.; Pérez-Sánchez, J. Contributions of MS metabolomics to gilthead sea bream (*Sparus aurata*) nutrition. Serum fingerprinting of fish fed low fish meal and fish oil diets. *Aquaculture* **498**, 503-512(2019).
- 28) Riccardo, M.; Roberto, C.; Tonina, R.; Roberto, A. Addressing marketplace gilthead sea bream (*Sparus aurata* L.) differentiation by ¹H NMR-based lipid fingerprinting. *Food Res. Int.* **63**, 258-264(2014).
- 29) Medina, S.; Pereira, J.A.; Silva, P.; Perestrelo, R.; Câmara, J.S. Food fingerprints-A valuable tool to monitor food authenticity and safety. *Food Chem.* **278**, 144-162(2019).
- 30) Folch, J.; Lees, M.; Sloane-stanley, G. H. A simple method for the isolation and purification of total lipids from animal tissues. *J. Biol. Chem.* **226**, 497-509(1957).
- 31) Kaitaranta, J.K.; Linko, R.R. Fatty acids in the roe lipids of common food fishes. *Comp. Biochem. Physiol. B: Biochem. Mol. Biol.* **79**, 331-334(1984).
- 32) Mohamadi Monavar, H.; Afseth, N.K.; Lozano, J.; Alimardani, R.; Omid, M.; Wold, J.P. Determining quality of caviar from Caspian Sea based on Raman spectroscopy and using artificial neural networks. *Talanta* **111**, 98-104(2013).
- 33) Liang, P.; Li, R.f.; Sun, H.; Zhang, M.; Cheng, W.J.; Chen, L.J.; Cheng, X.W.; Akoh, C.C. Phospholipids composition and molecular species of large yellow croaker (*Pseudosciaena crocea*) roe. *Food Chem.* **245**, 806-811(2018).
- 34) Madhusree, M.; Sandhya, M. A comparative analysis of different extraction solvent systems on the extractability of eicosapentaenoic acid from the marine eustigmatophyte *Nannochloropsis oceanica*. *Algal Res.* **38**, 101387(2019).
- 35) Saliu, F.; Leoni, B.; Pergola, R. Lipid classes and fatty acids composition of the roe of wild *Silurus glanis* from subalpine freshwater. *Food Chem.* **232**, 163-168(2017).
- 36) Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. *Chemom. Intell. Lab.* **2**(1), 37-52(1987).
- 37) Viaene, J.; Vander Heyden, Y. Introduction to herbal fingerprinting by chromatography. *Chemometrics in Chromatography* **2018**, 351-370(2018).
- 38) Kharbach, M.; Kamal, R.; Marmouzi, I.; Barra, I.; Cherah, Y.; Alaoui, K.; Vander Heyden, Y.; Bouklouze, A. Fatty-acid profiling vs UV-Visible fingerprints for geographical classification of Moroccan Argan oils. *Food Control.* **95**, 95-105(2019).
- 39) Zhao, J.; Du, X.J.; Cheng, N.; Chen, L.Z.; Xue, X.F.; Zhao, J.; Wu, L.M.; Cao, W. Identification of monofloral honeys using HPLC-ECD and chemometrics. *Food Chem.* **194**, 167-174(2016).
- 40) Harel, M.; Tandler, A.; Kissil, G.W. The kinetics on nutrient incorporation into body tissues of gilthead seabream (*Sparus aurata*) females and the subsequent effects on egg composition and egg quality. *Br. J. Nutr.* **72**, 45-48(1994).
- 41) Shin, E.C.; Craft, B.D.; Pegg, R.B.; Phillips, R.D.; Eitenmiller, R. Chemometric approach to fatty acid profiles in Runner-type peanut cultivars by principal component analysis (PCA). *Food Chem.* **119**, 1262-1270(2010).
- 42) Fernández-Palacios, H.; Izquierdo, M.S.; Robaina, L.; Valencia, A.; Salhi, M.; Vergara, J.M. Effect of n-3 HUFA level in broodstock diets on egg quality of gilthead seabream (*Sparus aurata* L.). *Aquaculture* **132**, 325-337(1995).
- 43) Rodríguez, C.; Cejas, J.R.; Martín, V.M.; Badía, P.; Samper, M.; Lorenzo, A. Influence of n-3 highly unsaturated fatty acid deficiency on the lipid composition of broodstock gilthead seabream (*Sparus aurata* L.) and on egg quality fish. *Physiol. Biochem.* **18**, 177-187(1998).
- 44) Almansa, E.; Pérez, M.J.; Cejas, J.R.; Badía, P.; Villamandos, J.E.; Lorenzo, A. Influence of broodstock gilthead seabream (*Sparus aurata* L.) dietary fatty acids on egg quality and egg fatty acid composition throughout the spawning season. *Aquaculture* **170**, 323-336(1999).
- 45) Johnson, R.A.; Wichern, D.W. *Applied multivariate statistical analysis*. Prentice Hall, London(2002).
- 46) Berrueta, L.A.; Alonso-Salces, R.M.; Héberger, K. Supervised pattern recognition in food analysis. *J. Chromatogr. A* **1158**, 196-214(2007).