



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

Thesis for the Degree of Doctor of Engineering

Development of Food Authenticity Markers

Using a Chemometrics Approach:

A Case Study on Origin Discrimination of

Cultured Red Seabream (*Pagrus major*)

from Korea and Japan Based on Food

Compositional and Metabolomic Analysis

by

Junho Yang

Department of Food Science and Technology

The Graduate School

Pukyong National University

August 2025

Development of Food Authenticity Markers

Using a Chemometrics Approach:

A Case Study on Origin Discrimination of Cultured Red  
Seabream (*Pagrus major*) from Korea and Japan Based  
on Food Compositional and Metabolomic Profiles

(화학계량학 접근을 활용한 식품 진위판별 마커

개발: 한국산 및 일본산 양식 참돔의 식품성분 및  
대사체 분석을 통한 원산지 판별사례연구)

Advisor: Prof. Ji-Young Yang

by

Junho Yang

A thesis submitted in partial fulfillment of the requirements for the  
degree of

Doctor of Engineering

in Department of Food Science and Technology, The Graduate  
School, Pukyong National University

August 2025

Development of Food Authenticity Markers Using a  
Chemometrics Approach:  
A Case Study on Origin Discrimination of Cultured Red  
Seabream (*Pagrus major*) from Korea and Japan Based  
on Food Compositional and Metabolomic Analysis

A dissertation

by

Junho Yang

Approved by:

\_\_\_\_\_  
(Chairman) Young-Mog Kim. Ph.D.

\_\_\_\_\_  
(Member) Gun-Do Kim. Ph.D.

\_\_\_\_\_  
(Member) Hyeyoung Lee. Ph.D.

\_\_\_\_\_  
(Member) Kil-Bo Shim. Ph.D.

\_\_\_\_\_  
(Member) Ji-Young Yang. Ph.D.

August 22, 2025

# CONTENTS

요 약 .....	xiii
<b>Chapter 1 . General Introduction.....</b>	<b>1</b>
1.1. Background of red seabream ( <i>Pagrus major</i> ).....	1
1.2. Fish origin labeling.....	6
1.3. Food authentication of the geographical origin.....	9
1.4. Compositional analysis.....	11
1.4.1. Amino acid analysis for nutritional profiling.....	11
1.4.2. Fatty acid analysis for origin differentiation.....	16
1.4.3. Mineral analysis reflecting geographical variability .....	18
1.5. Metabolomics for comprehensive discovery biomarker.....	22
1.6. Chemometrics approach for food authenticity .....	25
1.7. Objective of the thesis .....	27
<b>Chapter 2 . Development of Biomarkers to Distinguish Different     Origins of Red Seabreams (<i>Pagrus major</i>) from</b>	

<b>Korea and Japan by Fatty Acid, Amino Acid, and Mineral Profiling .....</b>	<b>29</b>
2.1. Introduction .....	31
2.2. Materials and methods .....	33
2.2.1. Sample preparation for Korean and Japanese cultured red seabream.....	33
2.2.2. Fatty acid analysis for Korean and Japanese cultured red seabream.....	35
2.2.3. Amino acid analysis for Korean and Japanese cultured red seabream.....	36
2.2.4. Minerals analysis for Korean and Japanese cultured red seabream.....	37
2.2.5. Statistical analysis .....	38
2.3. Results and discussion .....	39
2.3.1. Differences in fatty acids, amino acids, and minerals profiling between Korean and Japanese red seabreams	40

2.3.2. Comparison of red seabream origin by fatty acid analysis of seasonal red seabream.....	46
2.3.3. Selection of candidate biomarker for discriminating different origins of red seabream.....	51
2.3.4. Identification of red seabream origin using linoleic acid analysis.....	59
2.4. Conclusion .....	62
<b>Chapter 3 . Analysis of Metabolites of Red Seabream (<i>Pagrus major</i>) from Different Geographical Origins by Capillary Electrophoresis Time-of-Flight Mass Spectrometry .....</b>	<b>66</b>
3.1. Introduction.....	68
3.2. Material and methods .....	71
3.2.1. Sample preparation for Korean and Japanese red seabream .....	71
3.2.2. Instrument condition of CE-TOF/MS.....	73

3.2.3. Data processing for metabolite profiles .....	75
3.2.4. Statistical analysis .....	75
3.3. Results and discussion .....	76
3.3.1. Comparison of red seabreams from different regions by metabolite profiles .....	76
3.3.2. Identification of metabolic difference between Korean and Japanese red seabream .....	90
3.4. Conclusion .....	97
<b>Chapter 4 . Machine Learning-Based Biomarker Discovery for Origin Discrimination of Cultured Red Seabream Using Metabolomics .....</b>	<b>99</b>
4.1. Introduction .....	102
4.2. Materials and methods .....	104
4.2.1. Data preprocessing for metabolite profiles normalization .....	105
4.2.2. OPLS-DA modeling and overfitting assessment.....	105

4.2.3. Random forest classification modeling .....	106
4.2.4. Feature selection via statistical criteria.....	107
4.2.5. Biomarker validation using statistical analysis .....	107
4.3. Results and discussion .....	108
4.3.1. Evaluation of the OPLS-DA model to distinguish Korean and Japanese cultured red seabream .....	108
4.3.2. Classification performance of random forest using full CE- TOF/MS dataset from Korean and Japanese red seabream .....	111
4.3.3. Impact of feature selection on classification model stability from metabolic data of Korean and Japanese red seabream.....	114
4.3.4. Identification of candidate biomarkers to distinguish Korean and Japanese cultured red seabream.....	122
4.4. Conclusion .....	126
<b>Summary .....</b>	<b>129</b>

**Acknowledgment .....134**

**References .....136**



## List of Figures

- Fig. 1. 1. The morphology of red seabream (A) and distributed in the Northwest Pacific, including South Korea, Japan, China, and Taiwan (B).....2**
- Fig. 2. 1. Heatmap with clustering of chemical profiles for Korean and Japanese red seabreams.....41**
- Fig. 2. 2. Principle component analysis (PCA) score plot of fatty acid compounds from Korean and Japanese red seabreams.....47**
- Fig. 2. 3. OPLS-DA model of fatty acid analysis (A) and heatmap of the top 10 fatty acids (B) from red seabreams collected during spring, summer, and winter.....49**
- Fig. 2. 4. Biomarker analysis to identify candidate biomarkers to discriminate between Korean and Japanese red**

<p>seabreams (A, autumn, n = 60; B, spring, n = 10; C, summer, n = 10; D, winter, n = 10). .....</p>	52
<p><b>Fig. 2. 5. The ROC curves for linoleic acid across all red seabream samples including standard and seasonal groups showed an AUC of 0.937 for concentration (g/100 g) (A) and 1.0 for its percentage of total fatty acids (B). .....</b></p>	58
<p><b>Fig. 3. 1. PCA of metabolome data from Japanese (JPM, n = 10) and South Korean (KPM, n = 10) red seabream muscle samples.....</b></p>	78
<p><b>Fig. 3. 2. The heatmap with hierarchical clustering analysis of red seabream using 233 putative metabolite profiles between the different origins of Korea and Japan.</b></p>	79
<p><b>Fig. 3. 3. A heatmap with hierarchical clustering of the top 103</b></p>	

<p><b>metabolites in the different origins of red seabreams (A) and the metabolism enrichment analysis of different metabolites in red seabreams by matching the KEEG database (B).</b>.....</p>	81
<p><b>Fig. 3. 4. Difference of histidine metabolism between Korean and Japanese red sea breams.</b>.....</p>	92
<p><b>Fig. 3. 5. Difference of arginine biosynthesis pathway between Korean and Japanese red seabreams.</b>.....</p>	95
<p><b>Fig. 4. 1. Evaluation of a predictive model of OPLS-DA model for discrimination of Korean and Japanese red seabreams.</b>.....</p>	110
<p><b>Fig. 4. 2. Random forest classification of full dataset to distinguish different origins of red seabream.</b>.....</p>	113
<p><b>Fig. 4. 3. Volcano plot for feature selection from 233 metabolic</b></p>	

**data to approach biomarker discovery. .... 116**

**Fig. 4. 4. Machin learning-based biomarker to distinguish  
different origins of red seabream validation..... 117**

**Fig. 4. 5. Comparison of top 15 features selected by random  
forest before and after feature selection..... 120**



## List of Tables

<b>Table 1. 1. The research area concerns the geographical origin discrimination of marine products, with studies dating back to 2000. ....</b>	<b>12</b>
<b>Table 2. 1. The list of fatty acids with variable importance for Korean and Japanese red seabream .....</b>	<b>42</b>
<b>Table 2. 2. Comparative analysis of linoleic acid content (g/100 g) and cut-off value in fish samples collected across four seasons.....</b>	<b>54</b>
<b>Table 2. 3. Comparative analysis of linoleic acid content (% of total fatty acids) and cut-off value in fish samples collected across four seasons. ....</b>	<b>55</b>
<b>Table 2. 4. Identification of red seabream origin using linoleic acid percentage of total fatty acid by fatty acid</b>	

<b>analysis.....</b>	<b>60</b>
<b>Table 3.1. CE-TOF/MS conditions for obtaining the metabolome profiles from red seabream muscle samples.....</b>	<b>74</b>
<b>Table 3. 2. Key differential metabolites and their associated pathways between Korean and Japanese red seabream.....</b>	<b>83</b>
<b>Table 4. 1. Evaluation of statistical metrics for biomarker selection to distinguish the origins of Korean and Japanese red seabreams .....</b>	<b>124</b>

화학계량학 접근을 활용한 식품 진위판별 마커 개발:  
한국산 및 일본산 양식 참돔의 식품성분 및 대사체 분석을  
통한 원산지 판별 사례 연구

양 준 호

부경대학교 식품공학과

요 약

참돔(*Pagrus major*)은 대표적인 수산 단백질원으로서 한국, 일본, 중국 등지에서 오랫동안 소비되어 온 주요 양식 어종이다. 현재 우리나라에 유통되는 수입 참돔의 대부분(약 99%)은 일본산 양식 참돔이며, 이는 국내 수산물 시장의 상당 부분을 차지하고 있다. 그러나 최근 일본의 원자력 발전소 폐수 해양 방류 등의 해양 환경 오염 문제로 인해, 국내 소비자들

사이에서 수입 수산물에 대한 기피 현상이 증가하고 있다. 이에 따라 일부 시장에서는 원산지 미표기 또는 허위표시 등 수산물 원산지 위반 사례가 보고되고 있으며, 정부에서는 현장 단속을 강화하고 있으나 농축산물과 달리 수산물에 대한 과학적 원산지 판별 기법은 아직 부족한 실정이다. 본 연구는 참돔의 원산지에 따른 생물학적 차이를 이화학적 성분 분석과 대사체학 기반 분석을 통해 규명하고, 이를 바탕으로 참돔 원산지 판별법을 개발하고자 하였다.

우선, 한국산과 일본산 참돔의 성분 조성 차이를 비교하기 위해 지방산, 아미노산, 미네랄 조성을 분석하였다. 총 29 종의 지방산, 17 종의 구성 아미노산, 4 종의 미네랄을 분석하였으며, 계층적 군집 분석(HCA)과 직교 부분 최소 제곱 판별 분석(OPLS-DA)을 통해 지방산 조성에서 명확한 군집 차이를 확인하였다. 계절별 참돔을 분석한 결과,  $p$ -value 및 변수 중요도(VIP score)를 기준으로 상위 10 개의 지방산을 선별하였고, 이 중 리놀레산(linoleic acid)은 ROC curve

분석에서 높은 민감도와 특이도를 보여 바이오마커 후보로 제시되었다. 특히 리놀레산 함량이 총 지방산 대비 6.45%를 기준으로 한국산과 일본산 참돔을 효과적으로 구분할 수 있는 가능성이 확인되었다.

또한, CE-TOF/MS 기반의 비표적 대사체 분석을 통해 한국산과 일본산 양식 참돔 간의 생리학적 차이를 평가하였다. 총 233 개의 대사산물을 분석하였고, PCA 와 HCA 분석을 통해 원산지에 따른 대사체 패턴의 차이를 확인하였다. 이 중 11 개의 대사산물이 유의한 차이를 보였으며, KEGG pathway 분석 결과 히스티딘 대사와 아르기닌 생합성과 밀접한 관련이 있음을 확인하였다. 이를 통해 대사체 기반의 원산지 판별 가능성을 제시하였다.

한국산과 일본산 양식 참돔의 대사체 데이터를 기반으로, 랜덤 포레스트(Random Forest, RF)를 활용한 바이오마커 발굴 전략을 제시하였다. 고차원 저샘플 대사체 데이터에서 발생 가능한 과적합 문제를 해결하기 위해, FDR 보정  $p$  값( $<0.05$ )과

$|\log_2 \text{ fold change}| (>2)$ 를 기준으로 한 univariate 기반 feature selection 을 수행하였다. 차원 축소 후 구축된 RF 모델은 OOB 오류율이 5%에서 0%로 감소하고 변수 중요도(MDA)가 증가하여 모델의 안정성과 예측력을 향상시켰다. 통계적 유의성과 함께 효과 크기(Cohen's  $d$ )의 증가 및 bootstrap 기반 검증( $p < 0.001$ )을 통해 변수의 신뢰도가 강화되었으며,  $N$ ,  $N$ -dimethylglycine, creatine, hypotaurin 은 생물학적 타당성까지 고려된 유망한 바이오마커로 확인되었다. 본 연구는 소규모 대사체 데이터에서 신뢰도 높은 바이오마커 탐색을 위한 차원 축소 기반 머신러닝 접근의 가능성을 제시하였다.

본 연구에서는 참돔의 원산지 판별을 위해 이화학적 분석과 대사체학 기반 접근법을 활용하여, 한국산과 일본산 참돔 간의 대사적 차이를 분석하고 원산지 판별을 위한 바이오마커를 발굴하였다. 이러한 통합 분석을 통해 참돔의 원산지를 구별할 수 있는 대사체 기반 바이오마커 개발의 가능성을 확인할 수

있었다. 하지만, 본 연구에 사용된 시료는 계절적, 지리적 범위가 제한적이었기 때문에, 제시된 원산지 판별 기법의 적용성은 제한될 수 있다. 참돔의 대사체는 양식 환경, 수온, 사료 등의 다양한 해양 환경 변수에 영향을 받을 수 있으므로, 향후 연구에서는 장기적이고 지속적인 샘플링과 모니터링을 통해 이러한 변수들을 반영한 대사적 변화 차이 탐색이 필요하다. 결론적으로, 본 연구는 수산물의 원산지를 판별하기 위한 새로운 접근법으로써 이화학 분석 및 대사체 정보를 통합하고, 이를 통계적 분석 기법과 결합하여 실질적인 판별력을 갖는 바이오마커를 발굴하는 전략을 제시할 수 있음을 시사한다.

# Chapter 1. General Introduction

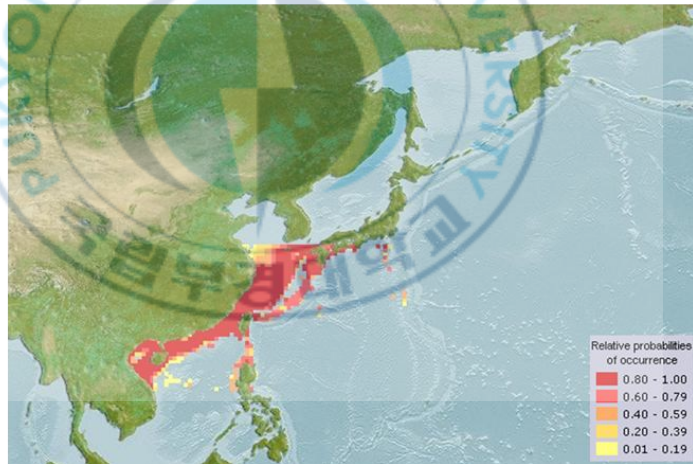
## 1.1. Background of red seabream (*Pagrus major*)

Red seabream (*Pagrus major*), a species in the Sparidae family, has long been farmed and consumed in Korea, Japan, and China. The red sea bream is a fish species that grows to more than 1 m in length and has a bright red color with blue spots (Myoung, 2009) (Fig.1.1 A). It is widely distributed in the Northwest Pacific, including the southern sea of Korea, Jeju Island, Japan, China, and Taiwan ([www.fishbase.se/search.php](http://www.fishbase.se/search.php)) (Fig.1.1 B). It lives at a depth of 30 to 150 meters off the coast, and the water temperature is 15 to 28 degrees Celsius. It is difficult to survive in low temperatures below 10 degrees Celsius, so it requires a suitable habitat (Kang & Hwang, 2003). According to the statistics of production by fisheries in Korea (KOSIS, 2024), the amount of fish produced by culture is more than that of naturally harvested fisheries, such as pelagic and offshore fisheries, and relies on aquaculture production. In Korea, red seabream, blackhead seabream (*Acantopagrus schlrgeii*), and rock bream

**A**



**B**



**Fig. 1. 1. The morphology of red seabream (A) and distribution in the Northwest Pacific, including South Korea, Japan, China, and Taiwan (B).**

*Oplegnathus fasciatus*) are the main species cultivated in aquaculture (NIFS, 2020). Among them, the aquaculture production of red seabream accounted for 84.27% as of 2022, representing the most significant proportion of domestic aquaculture production (MOF, 2024). Additionally, red seabream is imported from Japan and China, with Japanese red seabream accounting for approximately 99% of imports (TRASS, 2024). Although both Korea and Japan produce red seabream through aquaculture, the systems used in each country differ significantly in terms of feed types and culture environments. The marine culture method is used due to the environmental conditions required by red seabream, such as appropriate water temperature (NIFS, 2020; Shin et al., 2008). In Korea, red seabream aquaculture is mainly located on the southern coast (Tongyeong, Geoje, Goseong-gun, Namhae-gun, Gyeongsangnam-do, and Yeosu, Jeollanam-do) with seawater temperature as 11 to 25°C, In Japan, red seabream farms have been established off the coast of Ehime Prefecture (seawater temperature 15-27°C), including the Seto Inland Sea, located west of Shikoku Island (Sawayama et al., 2019; Zenitani et al., 2009). According to the Ministry of Oceans and Fisheries, raw feed (91%)

and compound feed (9%) are fed to Korean red seabream farms, while in Japan, 100% of the feed fed to Japanese red seabream farms is compound feed, showing the difference in feed supply (NIFS, 2014).

Seafood is increasingly consumed globally as a source of protein due to its high protein content and nutritional value, which includes unsaturated fatty acids such as eicosapentaenoic acid (C20:5) and docosahexaenoic acid (C22:6) (Kawarazuka, 2010; Lund, 2013). Among these, red seabream is a representative white fish produced in Korea and is widely used as a source of fish protein because of its high protein and low-fat content. The nutritional value of red sea bream has been studied extensively. Glutamic acid, aspartic acid, and lysine have been reported as the main amino acids in red seabream (Shin et al., 2008; Yoon et al., 2015), and it is believed that these amino acids play a significant role in the flavor of red seabream. In addition, lysine is a grain-restricted amino acid, which increases the value of red seabream as a protein source in Asian countries where rice is a staple food (Rharrabti et al., 2001). A study investigating the amino acid composition of red seabream according to origin and growth conditions reported no significant differences in the content of

each analyzed amino acid. However, amino acids such as taurine, alanine, and proline, which are responsible for taste, were reported to increase seasonally (Kim et al., 2000). In addition, a comparison of the amino acid composition of farmed and wild-caught red seabream from the same region reported differences in taurine content but no significant differences in other amino acids. The fatty acid composition of red seabream was 14.5% saturated fatty acids, 15.7% monounsaturated fatty acids, and 69.8% polyunsaturated fatty acids, with unsaturated fatty acids identified as the major fatty acids, with eicosapentaenoic acid at 7.1% and docosahexaenoic acid at 15.2% (Yoon et al., 2015). These are n-3 fatty acids that have been reported to lower serum triglyceride and cholesterol concentrations and inhibit the development of coronary artery disease and thrombosis (Calder & Yaqoob, 2009).

As such, red sea bream has long been a highly marketable fish species for its flavorful and nutritious source of fish protein. Advances in aquaculture technology have increased production, allowing Japan to import farmed red seabream, making it one of the most readily available fish species to consumers.

## **1.2. Fish origin labeling**

In South Korea, a system has been implemented to ensure the labeling of the country of origin for seafood and seafood products, guaranteeing consumers' right to know, promoting fair trade, and protecting both producers and consumers (MOF, 2021). The labeling of the country of origin for seafood is mandatory for the following parties: 1) those who produce and process seafood and seafood products for shipment; 2) those who sell seafood and seafood products at department stores, discount stores, wholesale markets, and traditional markets; 3) those who sell seafood and seafood products through TV home shopping, the internet, newspapers, and delivery apps; 4) those who manufacture and distribute edible salt. Additionally, the origin of seafood must be indicated for domestic seafood, deep-sea seafood, processed seafood, and imported seafood and its products, with the standards as follows: 1) Domestic seafood: labeled as domestic or locally produced; 2) Deep-sea seafood: labeled as deep-sea or deep-sea (specific sea area); 3) Processed seafood:

indicating the origin of the used raw materials; 4) Imported seafood and its products: indicating the country of import (according to the country of origin at the time of customs clearance under the 'Foreign Trade Act'). The methods of indicating the origin are divided into three types: 1) For packaged seafood, print the origin information on the package or attach it using stickers or labels produced by electronic scales; 2) For unpackaged seafood, attach tags to the seafood, or indicate the origin on signs or containers used for sales; 3) For live seafood, such as live fish, ensure that domestic and imported products of the same type are not mixed in storage facilities like aquariums and indicate the origin using signs or placards (NFQS, 2024).

Currently, 322 types of seafood are managed under the origin labeling system, including 225 domestic and deep-sea seafood, 73 types of processed seafood, and 24 types of imported seafood (NFQS, 2024). Although the South Korean government strictly monitors the labeling of seafood origins, including traceability and post-traceability, violations still occur in some seafood markets where the origin of seafood is falsely labeled or not labeled at all (MOF, 2022). Cases of origin labeling

violations include falsely labeling Japanese red seabream as Korean red seabream. Following the discharge of contaminated water from the Fukushima nuclear power plant in 2011, South Korean consumers have been reluctant to purchase Japanese seafood due to safety concerns. This preference for safe food due to environmental pollution and food safety concerns is supported by research findings (Kim & Chong, 2023). The Ministry of Food and Drug Safety (MFDS) conducts radiation tests on imported seafood to ensure safety. It has banned the import of agricultural and seafood products from certain high-risk areas in Japan after the nuclear accident (MFDS, 2024).

Despite efforts to secure consumers' right to know and the safety of food consumption in South Korea, violations of origin labeling for seafood still occur in some markets. This is believed to be due to consumers' persistent concerns about the safety of imported seafood. Therefore, measures are needed to ensure the safety of imported seafood for domestic consumers.

### **1.3. Food authentication of the geographical origin**

Food authenticity encompasses the reliability of food in terms of its quality, geographical origin, and labeling (Deng et al., 2024). In recent years, the authentication of geographical origin has become a critical issue in food science, driven by the growing prevalence of food fraud and the need to ensure traceability, consumer trust, and regulatory compliance (FAO, 2021). Various analytical techniques have been developed to address this issue, each based on distinct scientific principles with differing degrees of applicability. Physicochemical analysis provides a cost-effective and rapid means of assessing compositional differences between samples, although it may lack precision for complex samples or subtle differences (Tsagkaris et al., 2021). Stable isotope ratio analysis (SIRA) offers high accuracy by reflecting environmental and climatic conditions, but it requires expensive instrumentation and extensive reference databases (Camin et al., 2007; Gonzalvez et al., 2009). DNA-based methods are highly accurate for species identification and particularly useful in detecting substitution fraud in processed products; however, their utility is limited when differentiating geographic origin

within the same species (Hellberg & Morrissey, 2011). Metabolomics, which profiles endogenous small molecules that reflect physiological and environmental influences, enables high-resolution discrimination and biomarker discovery, although it demands advanced instrumentation and statistical expertise (Selamat et al., 2021). Among these, physicochemical analysis, which identifies general compositional trends, and metabolomics, which reflects environmental and physiological influences, stand out for their practical applicability and ability to reflect both compositional and regional biological variation.

In Korea, the system for labeling the country of origin of agricultural and marine products has been implemented to promote fair trade and ensure consumers' right to know. Although a monitoring framework is in place to support transparency in the agricultural and fisheries markets, origin mislabeling incidents continue to be reported. The National Agricultural Products Quality Management Service (NAQS) provides origin authentication information for agricultural and livestock products and has developed field testing kits, such as those for detecting swine fever antibodies in domestic pigs (NAQS, 2021; 2024). In contrast, the origin

authentication of marine products remains limited. For example, the National Institute of Fisheries Science (NIFS) has developed a method to trace the origin of imported eel using interspecies genetic differences (NIFS, 2018). However, this method does not apply to intra-species differentiation, which is often required for marine species with similar genetic backgrounds. Although previous studies have reported genetic and physicochemical differences among fish from different regions (Table 1.1), current origin-tracing technologies are only applicable to a few species and face limitations in resolving subtle geographic differences. Therefore, to expand the range of applicable species and improve classification accuracy, the integration of chemical and metabolomic analyses presents a promising solution, offering both practicality and scientific depth for the origin authentication of marine products.

## **1.4.Compositional analysis**

### **1.4.1. Amino acid analysis for nutritional profiling**

Seafood, like livestock, serves as a source of protein. Fish protein

**Table 1. 1. The research area concerns the geographical origin discrimination of marine products, with studies dating back to 2000.**

No.	Marine products	Analysis type	Methods	Ref.
1	Red seabream ( <i>Pagurus major</i> ); Rockfish ( <i>Sebastes pachycephalus</i> ); Flounder ( <i>Paralichthys dentatus</i> )	Compositional analysis	Taste compounds	Kim et al., 2000
2	European seabass ( <i>Dicentrarchus labrax</i> L.)	Compositional analysis	NIRS	Xiccato et al., 2004
3	Olive flounders ( <i>Paralichthys olivaceus</i> )	Genetic	DNA sequence analysis	Song et al., 2004
4	Red seabream ( <i>Pagurus major</i> )	Compositional analysis	Heavy metal contents and antioxidant activity	Hwang et al., 2015

*Continued*

---

5	Yellow fin tuna ( <i>Thunnus albacares</i> )	Genetic	Genomics	Percoraro et al., 2018
6	Perch ( <i>Perca fluviatilis</i> ); pumpkinseed sunfish ( <i>Lepomis gibbosus</i> )	Metabolomics	UHPLC-HRMS/MS	Marie & Gallet., 2022
7	Sea cucumber ( <i>Apostichopus japonicus</i> )	Metabolomics	UPLC-Q-TOF/MS	Zhao et al., 2022
8	Tropical tuna	Metabolomics	NMR	Bodin et al., 2022
9	Ark clam ( <i>Scapharca subcrenata</i> )	Genetic	SNP	Choi et al., 2023

---

contributes to the formation of protein structures through the biosynthesis of amino acids (Mohanty et al., 2014). Moreover, amino acids play essential roles in gene expression, nutrient transport, neurotransmission, and metabolism within the organism, acting as crucial biomolecules. Amino acids in human nutrition are generally classified into essential amino acids (arginine, histidine, cysteine, leucine, lysine, methionine, threonine, tryptophan, tyrosine, and valine), non-essential amino acids (aspartic acid, serine, and alanine), and conditionally essential amino acids (glutamic acid, glutamine, glycine, proline, and taurine) (Wu, 2010, 2013).

Amino acid profiling is conducted to investigate the nutritional composition of seafood, focusing on both compositional and free amino acids. Compositional amino acids form protein structures, which are analyzed through acid hydrolysis to break down the protein into individual amino acids for profiling. Studies that have performed compositional amino acid profiling in fish have shown that the distribution of amino acids is consistent across various fish species, with no significant differences observed within the same species. On the other hand, free amino acids are recognized for their role in protein metabolism within the

cells of organisms (Christensen, 1964). Profiling these compounds can reveal region-specific metabolic patterns, providing potential markers for discriminating geographical origins.

The released amino acids are separated using High-Performance Liquid Chromatography (HPLC), most commonly through ion-exchange or reversed-phase columns (Bayer et al., 1976). Separation is based on differences in charge, hydrophobicity, or polarity among the amino acids, allowing for precise resolution of individual components. To enhance detectability, derivatization of amino acids is often employed, either prior to (pre-column) or after (post-column) chromatographic separation (Bidlemeier et al., 1984; Le Boucher et al., 1997). Reagents such as o-phthalaldehyde (OPA), ninhydrin, or 9-fluorenylmethyl chloroformate (FMOC-Cl) are widely used for this purpose, providing strong chromophoric or fluorophoric properties that enable sensitive and selective detection via UV or fluorescence detectors (Acquaviva et al., 2016).

### **1.4.2. Fatty acid analysis for origin differentiation**

Fatty acid analysis is a critical technique for characterizing the lipid composition of biological and food samples (Adam et al., 2003). It provides essential information about the types and relative abundances of individual fatty acids, which is valuable for nutritional assessment, product quality control, and functional evaluation of biomaterials (Lands et al., 1992; Kratz et al., 2002).

The most widely employed method for fatty acid analysis is Gas Chromatography (GC), particularly when analyzing fatty acid methyl esters (FAMES) (Seppänen-Laakso et al., 2002). Conversion of free fatty acids or lipid-bound fatty acids into their methyl ester derivatives enhances volatility and thermal stability, allowing for efficient separation and quantification by GC. In many applications, GC is coupled with mass spectrometry (GC-MS) to improve sensitivity and structural identification of complex fatty acid profiles (Ecker et al., 2012; Quehenberger et al., 2011). However, for routine quantitative analysis, gas chromatography with flame ionization detection (GC-FID) is most commonly employed due to its high sensitivity, reproducibility, and ease of use for detecting

FAMEs (Carvalho et al., 2012; Danish & Nizami, 2019). The analytical workflow typically begins with lipid extraction, often performed using organic solvent systems such as chloroform–methanol (2:1, v/v), followed by derivatization, where fatty acids are esterified to form FAMEs using boron trifluoride (BF<sub>3</sub>) or hydrochloric acid (HCl) (Bligh et al., 1959; Morrison et al., 1964; Stoffel et al., 1959). This derivatization step is crucial for ensuring accurate and reproducible chromatographic performance.

Through this combination of extraction, derivatization, and chromatographic separation, fatty acid analysis provides a reliable approach to profiling lipid composition in diverse sample types, supporting applications across food science, biomedical research, and industrial product development (Chiu & Kuo, 2020; Ecker et al., 2012; Frostegård & Bååth, 1996; Quehenberger et al., 2011). In addition to being a rich source of protein, seafood is also recognized for its nutritional value in terms of lipids and fatty acids (Hardy & Lee, 2010). Although the composition of polyunsaturated fatty acids, such as omega-3, varies among different species, seafood consumption offers considerable health

benefits (Jabeen & Chaudhry, 2011). As a result, strategies to enhance seafood sustainability, such as dietary supplementation and seasonal harvesting, are of significant importance in industrial systems (Bianchi et al., 2022). Lipid supply is crucial not only for human health but also in aquaculture diets. Lipids play a vital role in forming cell membranes and contribute to lipid metabolism and growth (Hixon, 2014; Ogata et al., 2002). Also, the seasonal variation of fatty acid can be attributed to differences in fish (Zlatanov & Laskaridis, 2007). Given that fatty acid profiles differ among fish from different geographical origins due to aquaculture practices or oceanographic conditions, this analysis is particularly valuable for origin authentication

### **1.4.3. Mineral analysis reflecting geographical variability**

Minerals are essential nutrients involved in numerous physiological functions and are broadly classified into macrominerals and microminerals according to their required dietary intake levels (Lozano Muñoz & Díaz, 2020). Accurate quantification of these elements in food

is critical for evaluating nutritional value, ensuring product quality, and supporting public health monitoring.

Macrominerals, such as calcium (Ca), phosphorus (P), magnesium (Mg), sodium (Na), potassium (K), chloride (Cl), and sulfur (S), are needed in relatively large amounts and play vital roles in bone formation, fluid balance, nerve conduction, and muscle contraction (Ali, 2023). In contrast, microminerals, also known as trace elements, are required in smaller quantities but are equally important. Key examples include iron (Fe), zinc (Zn), copper (Cu), iodine (I), selenium (Se), manganese (Mn), and fluoride (F). These elements contribute to enzymatic reactions, immune regulation, antioxidant defense, and cellular metabolism (Mehraj et al., 2017; Chongtham et al., 2021; Farag et al., 2023). Due to their low abundance in food matrices, highly sensitive analytical techniques are necessary for the accurate detection and quantification of these compounds.

To assess mineral content in food, a range of instrumental techniques are employed. Atomic Absorption Spectroscopy (AAS) is widely used for its specificity and cost-effectiveness in determining the concentrations of

individual minerals (Nielsen et al., 2010). Atomic Emission Spectroscopy (AES), which measures the light emitted by excited atoms, facilitates multi-element detection (Thirumdas et al., 2019). More advanced methods, such as Inductively Coupled Plasma Optical Emission Spectroscopy (ICP-OES) and Inductively Coupled Plasma Mass Spectrometry (ICP-MS), offer high sensitivity and the ability to quantify multiple elements at trace levels simultaneously (Ferreira et al., 2023; Muller et al., 2016).

The analysis of minerals in food serves multiple key purposes. In nutritional science, it enables the evaluation of dietary adequacy and supports the formulation of balanced diets (Bouzari et al., 2015; Mehraj et al., 2017; Singla et al., 2023). In the food industry and regulatory sectors, it ensures product safety, consistency, and compliance with labeling standards (Mir-Marques et al., 2016).

Minerals such as calcium, iron, magnesium, phosphorus, selenium, and zinc are essential for human health and are commonly found in fish. These minerals support critical biological processes, including maintaining bone health, promoting enzyme functions, regulating fluid balance, and protecting cells from oxidative stress (Wu, 2010, 2013). The concentration

and availability of these minerals in fish are influenced by environmental factors, including water quality, salinity, temperature, and the fish's diet, all of which can vary across geographical regions (Hardy & Lee, 2010).

Marine fish, for example, are generally richer in iodine and sodium due to the higher salinity of ocean waters, while freshwater species may exhibit differing levels of minerals based on the composition of their aquatic environment (Jabeen & Chaudhry, 2011). Furthermore, variations in seasonal conditions and biological processes also impact the mineral content in fish, with differences observed in the mineral profiles of fish harvested at different times of the year (Zlatanov & Laskaridis, 2007). Understanding regional and environmental influences on the mineral composition of fish is essential for assessing their nutritional value and for optimizing aquaculture practices to improve mineral content. Such regional and seasonal variations in mineral profiles offer a valuable basis for distinguishing the geographical origin of environmentally sensitive fish species.

## **1.5. Metabolomics for comprehensive discovery biomarker**

Metabolomics is the study of identifying and quantifying small molecules or chemicals found in cells, organs, and organisms (German et al., 2005). The small molecules can include various endogenous and exogenous chemicals, such as peptides, amino acids, nucleic acids, carbohydrates, organic acids, minerals, and substances that can be ingested or synthesized in a cell or an organism as a metabolic response (Johnson et al., 2012). Metabolomics in organisms reveals variations in metabolic systems resulting from environmental factors such as temperature and water pollution (Johnson et al., 2012; Marie & Gallet, 2022; Otify et al., 2023). Mass spectrometry (MS) has become a cornerstone of metabolomics due to its high sensitivity, specificity, and ability to identify compounds based on the mass-to-charge ratio ( $m/z$ ) (Lei et al., 2011). This analytical capability facilitates the detection of subtle metabolic differences across complex biological matrices, which may not be discernible through conventional compositional or targeted analyses (Di Minno et al., 2021). Such enhanced resolution is achieved by coupling high-resolution mass spectrometry with chromatographic or

electrophoretic separation systems, allowing for a more comprehensive metabolic overview (Guillarme et al., 2010). To effectively separate and reduce matrix complexity prior to MS detection, various chromatographic and electrophoretic techniques are employed, including capillary electrophoresis (CE), liquid chromatography (LC), and gas chromatography (GC) (Moco et al., 2007; Munjal et al., 2022). Capillary Electrophoresis (CE) separates metabolites primarily based on their charge-to-size ratio under the influence of an electric field. It is especially suited for analyzing polar and charged small molecules, including amino acids, nucleotides, and organic acids. CE-MS offers high-resolution separation with minimal sample requirements and is particularly valuable for targeted or niche metabolomics applications (Boizard et al., 2016; Drouin & Ramautar, 2021). Liquid chromatography (LC), especially in its reversed-phase and hydrophilic interaction liquid chromatography (HILIC) modes, is one of the most commonly used platforms for non-volatile, thermally labile, and moderately polar to non-polar metabolites. LC-MS is a highly versatile technique that is widely applied in both targeted and untargeted metabolomics, including proteomics, secondary metabolite

profiling, and biomarker discovery (Höcker et al., 2021; Saito-Shida et al., 2018; Týčová et al., 2017; Zhou & Zhong, 2022). This method is widely used in various fields, including compound identification, proteomics, drug metabolomics, and food composition analysis (Makarov & Scigelova, 2010). Gas Chromatography-Mass Spectrometry (GC-MS) is well-suited for analyzing volatile and semi-volatile metabolites, including organic acids, alcohols, fatty acids (after derivatization), and sugars (Garcia & Barbas, 2010). GC generally provides high chromatographic resolution and reproducibility, and when coupled with MS (especially with electron ionization), it allows for confident metabolite identification based on established spectral libraries (Fiehn, 2016). MS-based metabolomics platforms have thus become indispensable tools in comprehensive metabolic profiling, enabling the identification of potential biomarkers for physiological status, disease diagnostics, food authentication, and environmental assessment (Gil-Solsona et al., 2016; Hoffmann et al., 2017; Johnson & Gonzalez, 2012; Klockmann et al., 2016; Lim et al., 2018).

## **1.6. Chemometrics approach for food authenticity**

In food science, food analysis involves a high-dimensional data set composed of various chemical components, including fatty acids, amino acids, stable isotope ratios, and metabolomic analysis, etc (Ye et al., 2023). Unlike traditional univariate approaches that rely on the presence or concentration of a single compound, food authenticity research requires the integration of multivariate data to detect correlations between variables in a food group.

Multivariate analyses incorporating unsupervised and supervised statistical techniques were performed to discriminate and predict metabolomes. These approaches are useful for identifying differences among groups, obtaining information about biological systems, and recognizing potential biomarkers through prediction models such as principal component analysis (PCA), hierarchical cluster analysis (HCA), and partial least squares discriminant analysis (PLS-DA) (Kim et al., 2009). As metabolomics generates large and complex datasets, applying chemometrics is essential for efficiently processing and interpreting the

data. Chemometrics, which combines statistical, mathematical, and computational tools, plays a key role in identifying relevant patterns and discriminative features from high-dimensional data for food authentication (Rodionova et al., 2024). PCA and orthogonal partial least squares discriminant analysis (OPLS-DA) are widely used for exploratory visualization, group separation, and biomarker identification among chemometric techniques. Furthermore, machine learning approaches such as support vector machines (SVM), Random forest, and other classification algorithms are increasingly integrated with chemometric pipelines to enhance predictive performance and model robustness (Huang et al., 2023; Liu et al., 2022). In omics, which encompasses complex datasets such as metabolomics, various feature selection methods are employed to facilitate effective biomarker discovery and identify specific biomarkers (Grissa et al., 2016). These methods are widely used in both targeted and untargeted profiling. In targeted approaches, models based on stable isotope ratios, fatty acids, and metabolite concentrations have been successfully used for the origin authentication of milk (Wang et al., 2021), honey (Chen et al., 2017), and wolfberry (Gong et al., 2022).

In untargeted approaches, omics-based profiling has identified molecular markers for distinguishing between farmed and wild salmon (Florino et al., 2019; Gu et al., 2020), as well as extra-virgin olive oils using metabolomic signatures (Ghisoni et al., 2019). Collectively, these chemometric strategies not only enhance the ability to discriminate the origin of food products but also contribute to the discovery of reliable biomarkers, laying the foundation for robust food traceability systems.

### **1.7.Objective of the thesis**

A country-of-origin labeling system has been established to ensure fair trade and consumer trust. However, violations of origin labeling still occur in some seafood markets in Korea. These violations are partly due to consumer concerns regarding food safety, particularly following incidents such as the Fukushima nuclear accident and the release of contaminated water. While strict monitoring and enforcement of origin labeling regulations are in place, research on methods for origin discrimination in seafood remains insufficient compared to similar methods developed for agricultural and livestock products.

This study aims to establish a robust analytical framework for discriminating the geographical origin of red seabream from Korea and Japan through a chemometrics-based approach. Specifically, it integrates compositional profiling, including fatty acids, amino acids, and minerals, with untargeted metabolomic analysis and chemometric modeling to identify and validate potential biomarkers. The specific objectives of this research are as follows:

- Identifying biomarkers to distinguish different origins of red seabreams using compositional analysis
- Comparison of different metabolisms in Korean and Japanese red seabreams
- A metabolomics-based strategy for the discovery of food authenticity markers for discrimination between cultured red seabream from Korea and Japan

**Chapter 2. Development of Biomarkers to Distinguish  
Different Origins of Red Seabreams (*Pagrus major*)  
from Korea and Japan by Fatty Acid, Amino Acid,  
and Mineral Profiling**

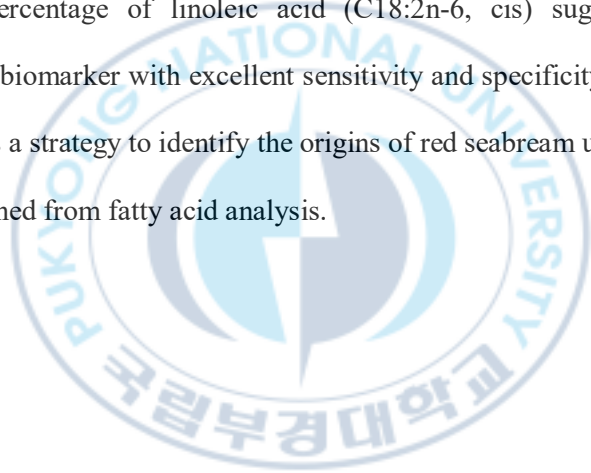
**Abstract**

The aim of the present study was to determine the origin of fish based on fatty acid, amino acid, and mineral analyses, and to develop biomarkers that can discriminate between Japanese and Korean red seabream. To identify the differences between the two groups, 29 fatty acid families, 17 amino acids, and 4 minerals were analyzed in 60 fish samples (standard sample collected in autumn), and fatty acid profiles were analyzed using

---

This chapter was based on the previously published manuscript: Development of biomarkers to distinguish different origins of red seabreams (*Pagrus major*) from Korea and Japan by fatty acid, amino acid, and mineral profiling. Food Research International. (2024, Vol. 180).

heatmap with hierarchical clustering analysis and orthogonal projections to latent structures discriminant analysis. The top 10 fatty acids that were different between the two groups were selected from all seasonal fish samples by combining variable importance in projection scores and p-values. According to the receiver operating characteristic curve analysis results, percentage of linoleic acid (C18:2n-6, cis) suggested as a candidate biomarker with excellent sensitivity and specificity. This study introduces a strategy to identify the origins of red seabream using linoleic acid obtained from fatty acid analysis.



## 2.1. Introduction

Red seabream (*Pagrus major*) is a migratory fish of the Sparidae family that has long been used as an aquatic protein source in Korea, Japan, and China (Kim et al., 2000). In particular, red seabream is actively cultivated in Korea and Japan, with 8,313 tons of red seabream annually produced by fish farms in Korea, accounting for 77.64% of the total production (MOF, 2022). Japanese and Chinese red seabream are the two main types imported into South Korea, with Japanese red seabream accounting for most imports at 97.71% (TRASS, 2023). However, the Fukushima nuclear disaster in Japan has led to the contamination of Japanese waters with radioactive wastewater, raising concerns about the radioactive contamination of seafood (BBC, 2021). As a result, the Korean government has enforced a strict quarantine on imported fish, while also requiring the labeling of fish according to their origins (MFDS, 2023, NFQS, 2023). Violations have occurred several times in some fish markets in South Korea, failing to label or mislabeling Japanese fish, including red seabream (MOF, 2022). Although guidelines for determining the chemical origin of agricultural and livestock products are available, no such

guidelines exist for marine products (NAPQ, 2021, 2022).

Previous studies have investigated regional and origin differences in fishery; for example, DNA sequence analysis was used to determine genetic differences among olive flounders (*Paralichthys olivaceus*) (Song et al., 2004), the nutritional value and safety of red seabream was evaluated (Hwang et al., 2015), genomic analysis was used to determine global geographical differences in yellow fin tuna (*Thunnus albacares*) (Pecoraro et al., 2018), and metabolomic analysis was used to investigate the chemistry of red seabream (Yang et al., 2022), yellowtail (*Seriola quinqueradiata*) (Shin et al., 2021), and yellow goosfish (*Lophius litulon*) (Yang et al., 2023). Such analyses are commonly reported in studies based on differences in geographic environments. In addition, the nutritional composition of fish can vary by size, sex, temperature, climate, seasonal variations, and diet (Bianchi et al., 2022, Hixson, 2014, Chaouch et al., 2003, Jang et al., 2009). Several studies have compared the nutritional content of fish from different regions (Kim et al., 2000), but even within the same species, nutritional content can vary by size (Lee et al., 2022). Therefore, identifying a candidate biomarker for determining the origin of

a fish can pose challenges. The present study was conducted in conjunction with a previous study that examined the metabolome of red seabream (Yang et al., 2022). Although some studies have previously examined differences in fatty acids and amino acids between farmed and wild-caught red seabreams (Kim et al., 2000), none have examined definitive differences in amino acids, fatty acids, and minerals in red seabreams from different countries.

In this study, candidate biomarkers were investigated by profiling fatty acids, amino acids, and minerals in Korean and Japanese red seabream, and differences in origin were preliminarily analyzed using multivariate statistical analysis. Seasonal fish samples were examined, and potential biomarkers for determining the origin of red seabream were identified through statistical approaches. Finally, a method for origin discrimination of red seabream based on fatty acid profiling was proposed.

## **2.2. Materials and methods**

### **2.2.1. Sample preparation for Korean and Japanese cultured**

### **red seabream**

Korean and Japanese red seabreams were sampled during the four seasons, and the initial samples collected during autumn were considered as standards. The samples during the other seasons were used for comparison with the standard sample. The standard sampling ( $n = 60$ ) was performed during autumn, and the seasonal sampling ( $n = 30$ ) was performed during winter, spring, and summer. The standard fish samples (Korean fish cultured by  $34^{\circ}44'19''\text{N}$   $128^{\circ}23'47''\text{E}$ ; Japanese fish cultured by  $34^{\circ}4'08''\text{N}$ ,  $136^{\circ}13'54''\text{E}$ ) were obtained from the Jagalchi Seafood Market in Busan, South Korea. For Korean cultured red seabream, the total body length and weight of standard samples ( $n = 30$ ) were  $44.82 \pm 2.90$  cm and  $1.28 \pm 0.24$  kg, respectively. The total body length and weight of Japanese cultured red seabream ( $n = 30$ ) were  $49.05 \pm 3.30$  cm and  $1.81 \pm 0.47$  kg, respectively. For comparisons with seasonal fish, the fish samples were purchased from the same market from winter to summer ( $n = 30$  per origin). Additionally, 20 fish samples from farms of unknown locations were mixed with Korean and Japanese red seabreams and used to evaluate the efficacy of the determination method identified in this study.

The fish muscles were used as the analytical samples. All fish samples used were edible fish, and the appropriate ethical procedures were followed. Fish were immediately euthanized by neck dislocation, the inedible fins, guts, and blood were removed, and the fish fillets were prepared. The animal study protocol was reviewed and approved by the Pukyong National University-Institutional Animal Care and Use Committee on ethical procedures and scientific care (approval number: PKNUIACUC-2022–30). The analytical samples were freeze-dried and homogenized using Mixer Mill MM400 (Retsch GmbH & Co., Haan, Germany) at 25 Hz for 60 s, and stored at -80 °C until analysis.

### **2.2.2. Fatty acid analysis for Korean and Japanese cultured red seabream**

The fatty acid content in samples containing 50 to 100 mg of fat was analyzed according to the Korean Food Code (MFDS, 2023). The fat and fatty acids in the samples were acid-digested with 8.3 M HCl and extracted with 12.5 mL diethyl ether, followed by 12.5

mL anhydrous petroleum ether. The fatty acids were methyl esterified using a 7% BF<sub>3</sub>-methanol solution and analyzed by gas chromatography (YL 6500 GC, Agilent Technologies, Santa Clara, CA, USA) combined with flame ionization detector (FID) and capillary GC column SP-2560 (100 m × 0.25 mm × 0.2 μm). The injection volume and temperature were 1.0 μL and 225°C, respectively, with a split ratio of 200:1. The oven temperature program was: holding at 100°C for 4 min, increasing the temperature to 240°C at a rate of 3°C/min, and holding for at least 15 min. The carrier gas used was helium with 0.75 mL/min flow. The content of each fatty acid was calculated by obtaining the conversion factor of each fatty acid in the GC-FID system from 37 fatty acid methyl ester standards for carbon numbers 4 to 24 and converting the content of fatty acid ethyl esters to that of fatty acids.

### **2.2.3. Amino acid analysis for Korean and Japanese cultured**

### **red seabream**

For total amino acids analysis, the homogenized fish muscle samples (150 mg) were hydrolyzed using 10 mL 6 N HCL 110°C for 24 h. The digestion solution was pressure dried at 40°C, purified to 50 mL with sodium citrate buffer (pH 2.2), filtered through a 0.20 µm membrane filter, and used as a sample for amino acid analysis (White and Fly, 1986). Constituent amino acids were analyzed using an Amino Acid Sequential Analyzer (L-8900, Hitachi, Tokyo, Chiyoda-ku, Japan) and ion exchange column (#2622SC-PH, Hitachi, Tokyo, Chiyoda-ku, Japan).

#### **2.2.4. Minerals analysis for Korean and Japanese cultured red seabream**

The contents of minerals that indicate differences between origins according to previous studies (Park et al., 2006) were selectively analyzed. the mineral compositions analyzed including magnesium (Mg), calcium (Ca), potassium (K), and sodium (Na) in Korean and Japanese fish to compare their differences. Analytical samples were prepared according to

the Korea Food Code (MFDS, 2023). The muscle samples (approximately 0.5 g) were pre-digested at 140°C for 15 min with 4 mL of 65% nitric acid. The digested samples were treated in a microwave reaction system (Microwave Reaction System Multiwave PRO; Anton Paar, Graz, Austria) at 1250 W, ramp for 20 min, and cooled by air for 15 min to prepare the analytical samples. The four minerals were analyzed by inductively coupled plasma, ICP-OES (Avio20, PerkinElmer., Santa Clara, USA). The plasma and nebulizer gases used were Ar at flow rates of 10 and 0.55 L/min, respectively. The auxiliary gas flow was 0.2 L/min, the radiofrequency (RF) power was 1,300 watts, and the pump flow and speed were 1.0 ml/min and 100 rpm, respectively. The samples were extracted using 20 times the volume of each sample weight.

#### **2.2.5. Statistical analysis**

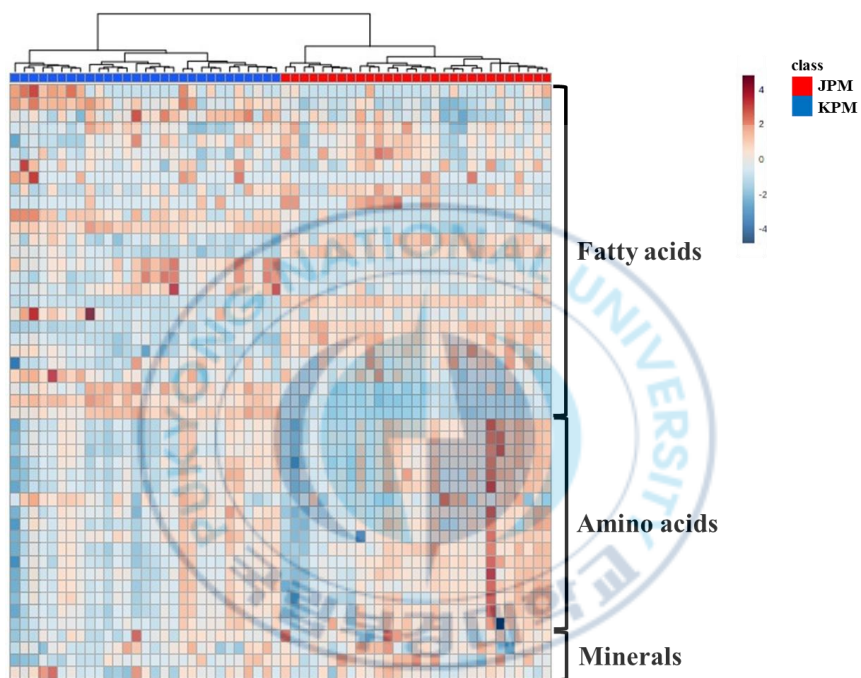
Statistical analyses were performed with univariate and multivariate statistical analyses. Univariate analyses were performed with SPSS software (version 27, SPSS Inc. Chicago, IL, USA) to compare various

components of red seabream using Student's *t*-test. Multivariate analyses were performed with SIMCA® version 18 (Sartorius, Umeå, Sweden) using principal component analysis (PCA) to determine the different origins of fish between Korean and Japanese aquaculture systems. In addition, orthogonal partial least squares discriminant analysis (OPLS-DA) was performed to predict the fish origins. Moreover, a heatmap of the different compounds in Korean and Japanese fish based on MetaboAnalyst 6.0 (<https://www.metaboanalyst.ca>) was constructed and hierarchical cluster analysis was performed. A combination of variable importance (VIP) scores and *p*-values obtained by OPLS-DA and *t*-tests, respectively, were performed to select candidate biomarkers for discrimination between the two groups of fish. Finally, receiver operating characteristic (ROC) curve analysis was performed to select biomarkers based on area under the curve (AUC) values. The fatty acid data were compared using ANOVA with the Bonferroni multiple range test, and statistical significance was set at  $p < 0.05$ .

### **2.3. Results and discussion**

### **2.3.1. Differences in fatty acids, amino acids, and minerals profiling between Korean and Japanese red seabreams**

In this study, 29 fatty acids, 17 amino acids, and four minerals were detected in the fish muscle samples of Korean and Japanese red seabreams collected in autumn. These included 11 saturated fatty acids, seven monounsaturated fatty acids, 10 polyunsaturated fatty acids, nine essential amino acids eight unessential amino acids, mg, Ca, K, and Na. All detected compounds are presented in a heatmap with hierarchical classification analysis (Fig. 2.1). The heat map revealed differences between Korean and Japanese red seabreams in 50 analyzed compounds. To evaluate the major components in the phylogenetic analysis, VIP scores, and *p*-values were compared to the differences between the two groups (Table. 2.1). Fatty acids were the main components that distinguished between Korean and Japanese red seabreams, with 14 fatty acids showing the best performance with VIP scores above 1.0 and *p*-values below 0.001. The VIP score and *p*-value are routinely used in combination to select candidate biomarkers (Santacruz et al., 2020). Thus,



**Fig. 2. 1. Heatmap with clustering of chemical profiles for Korean and Japanese red seabreams. KPM, Korean red seabream (n = 15); JPM, Japanese red seabream (n = 15).**

**Table 2. 1. The list of fatty acids with variable importance for Korean and Japanese red seabream**

Compounds	Quantitative analysis (g/100 g)			Comparative analysis	
	KPM	JPM	<i>p</i> -value <sup>1</sup>	Fold change <sup>2</sup>	VIP score <sup>3</sup>
Linoleic acid (C18:2n-6, Cis)	1.65±1.13	6.51±1.23	<0.001	3.94	1.96
Docosahexaenoic acid (C22:6n-3)	12.15±1.96	7.14±1.65	<0.001	0.59	1.86
Eicosapentaenoic acid (C20:5n-3)	5.28±0.77	3.08±0.55	<0.001	0.58	1.84
Palmitoleic acid (C16:1)	5.02±0.52	3.42±0.53	<0.001	0.68	1.84

Values are expressed as mean with standard deviation

<sup>1</sup> The *p*-value was obtained using t-test.

<sup>2</sup> The fold change is of computed by using averaged detection values.

<sup>3</sup> The VIP score was obtained by OPLS-DA

KPM, Korean red seabream (n = 15); JPM, Japanese red seabream (n = 15)

*Continued*

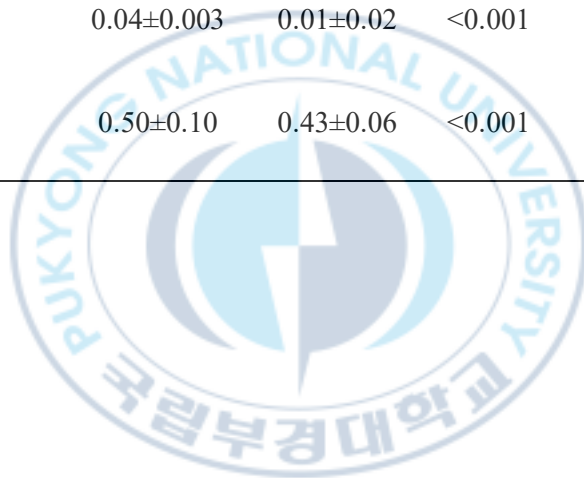
$\alpha$ -Linolenic acid (C18:3n-3)	0.50±0.11	0.87±0.09	<0.001	1.72	1.84
cis-11,14-Eicosadienoic acid (C20:2)	0.14±0.04	0.24±0.04	<0.001	1.72	1.79
Oleic acid (C18:1n-9,Cis)	13.72±1.91	18.18±1.74	<0.001	1.32	1.75
dihomo $\gamma$ -Linolenic acid (C20:3n-6)	0.10±0.03	0.16±0.03	<0.001	1.62	1.61
Docosadienoic acid (C22:2)	0.65±0.07	0.53±0.06	<0.001	0.83	1.54
cis-11-Eicosenoic acid (C20:1)	1.74±0.42	1.24±0.19	<0.001	0.71	1.49
Elaidic acid (C18:1n-9,trans)	0.11±0.02	0.14±0.03	<0.001	1.34	1.47

*Continued*

---

Myristic acid (C14:0)	2.41±0.25	2.04±0.18	<0.001	0.84	1.37
Myristoleic acid (C14:1)	0.04±0.003	0.01±0.02	<0.001	0.24	1.34
Nervonic acid (C24:1)	0.50±0.10	0.43±0.06	<0.001	0.86	1.08

---



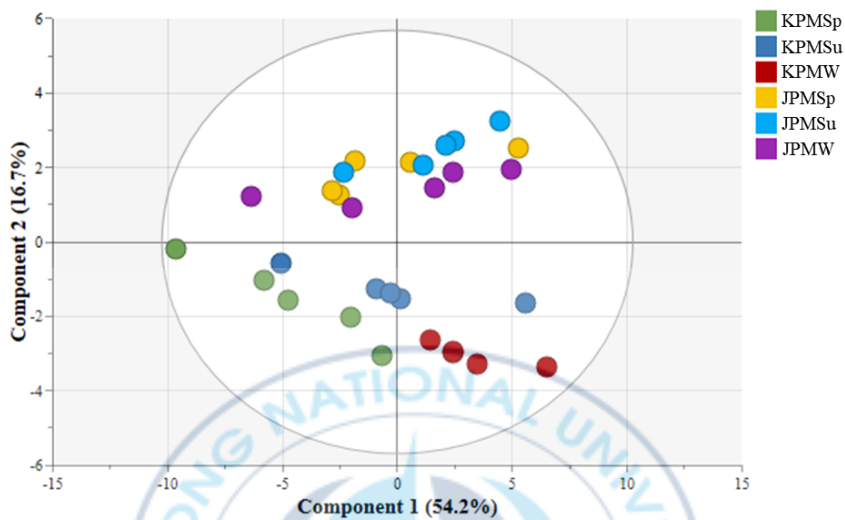
the 14 identified fatty acids are suitable as candidate biomarkers.

Species-specific differences in amino acids are not significantly found in fish (Bianchi et al., 2022). In addition, a previous study reported that the amino acid composition of farmed red seabreams from several regions was not significantly different (Kim et al., 2000). These discrepancies could be attributed to the differences in the analysis of the constituent amino acids in fish muscle tissue and metabolome extraction and analysis (Yang et al., 2022). Minerals can be classified as macro minerals (including calcium, magnesium, potassium, and sodium) and micro minerals (including manganese, copper, iodine, cobalt, fluoride, and selenium). Heavy metals, which are micro minerals, are used as indicators of marine environmental pollution because they can accumulate in aquatic products from polluted rivers and coastal waters (Choi et al., 2012). However, marine contamination is not specific to the country of origin of fish because it can vary according to the region and season. Macro minerals (Na, K, Mg, Ca) were analyzed based on prior findings indicating significant origin-related variations in these elements among commercial mackerel samples (Park et al., 2006). However, the results indicated that

these macro minerals were not effective in differentiating between Korean and Japanese cultured red seabreams. In contrast, variations in the composition of fatty acids in fish can be attributed to differences in their feed in addition to other factors, such as temperature, season, sex, and size (Zlatanos & Laskaridis, 2007, Özogul et al., 2009). Interestingly, Korean and Japanese aquaculture systems differ in seawater temperature and feed composition. Thus, fatty acid analyses were conducted to evaluate their potential for distinguishing fish of different geographical origins.

### **2.3.2. Comparison of red seabream origin by fatty acid analysis of seasonal red seabream**

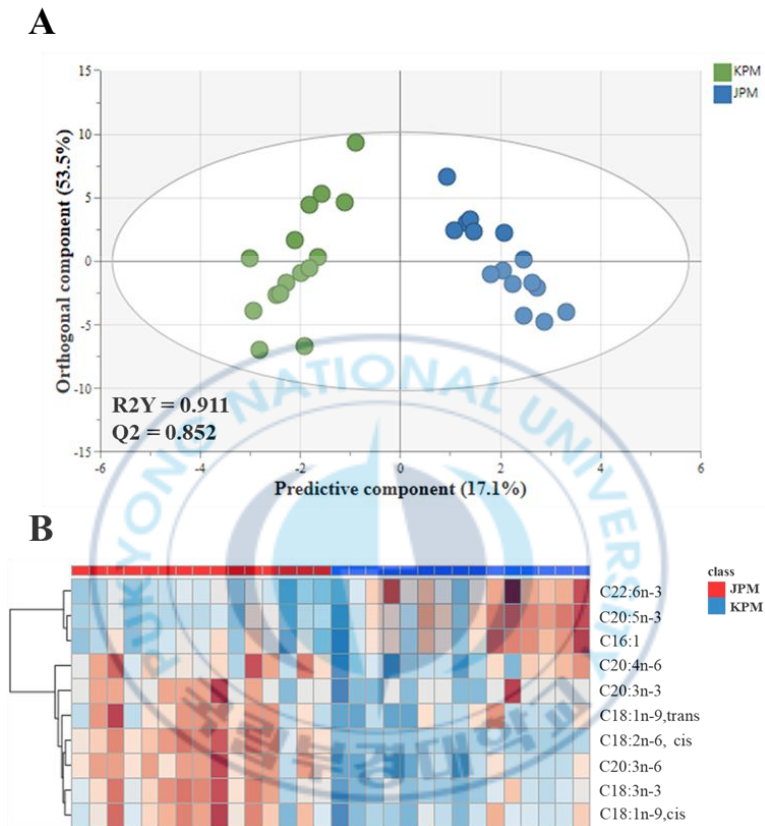
Through amino acid, fatty acid, and mineral analysis, the fatty acid content of Korean and Japanese fish was different. As previously mentioned, fatty acid content can differ by season; fatty acid content can vary by season; therefore, fatty acid profiles were further examined in fish samples collected during winter, spring, and summer. A total of 29 fatty



**Fig. 2. 2. Principle component analysis (PCA) score plot of fatty acid compounds from Korean and Japanese red seabreams.** JPMSp, Japanese red seabreams of spring (n = 5); JPMSu, Japanese red seabreams of summer (n = 5), JPMW, Japanese red seabreams of winter (n = 5); KPMSp, Korean red sea breams of spring (n = 5); KPMSu, Korean red seabreams of summer (n = 5); KPMW, Korean red seabreams of winter (n = 5).

acids were detected in each season, and the results are presented in the PCA plot shown in Fig. 2.2. The PCA plot represents the fatty acid content in Korean and Japanese fish collected in spring, summer, and winter with 54.2% and 16.7% of component 1 and 2, respectively. The fatty acids in fish from the same country showed overlap between seasons, and PCA plots revealed distinct differences in fatty acid content between Korean and Japanese fish.

PCA revealed that fatty acids were correlated between Korean and Japanese red seabream regardless of the season. PCA also shows differences between groups, OPLS-DA can reduce the complexity of PCA and improve predictions (Zhang et al., 2019). Thus, OPLS-DA was performed to identify a compound that could represent the differences between Korean and Japanese fish in all seasons (Fig. 2.3A). The OPLS-DA distinguished the fatty acids in red seabreams, with  $R^2Y$  and  $Q^2$  values of 0.911 and 0.852, respectively. A heatmap was constructed, and hierarchical analysis was performed using seasonal fatty acid data to determine the top 10 fatty acids that could differentiate between samples from different geographical locations (Fig. 2.3B). The top 10 compounds



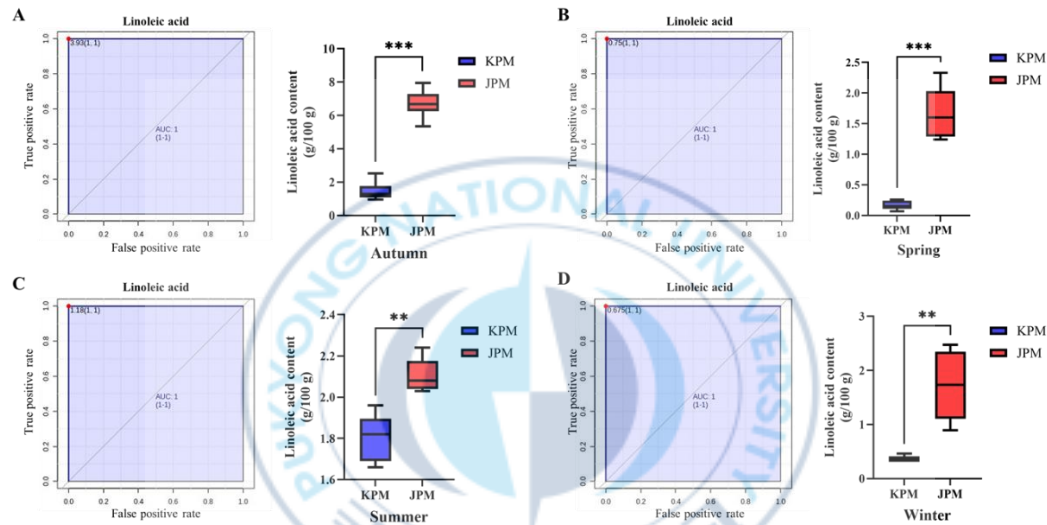
**Fig. 2. 3. OPLS-DA model of fatty acid analysis (A) and heatmap of the top 10 fatty acids (B) from red seabreams collected during spring, summer, and winter. KPM, Korean red seabream (n = 15); JPM, Japanese red seabream (n = 15).**

showed significant differences based on VIP scores  $>1.5$  obtained by OPLS-DA and  $p$ -value  $<0.001$ . Unsaturated fatty acids dominated the top10 compounds that were significantly different between Korean and Japanese red seabream. In general, marine products contain more unsaturated fatty acids than saturated fat and are, therefore, recommended in healthy diets (Taşbozan & Gökçe, 2017). The fatty acids of fish and their prey are similar, reflecting the effect of dietary intake on the fatty acid profile (Hardy & Lee, 2010). The feeds used in Korean and Japanese red seabream aquacultures in this study were different. In Korea, 91% of melt pellets (MP) and 9% of extrusion pellets (EP) were used at red seabream farms, whereas dried and extrusion feed were used in Japanese red seabream (NIFS, 2014; Song, 2012). Anchovy (*Engraulis japonicus*), pacific sand lance (*Ammodytes personatus*), and mackerel (*Scomber japonicus*) were used as ingredients in MP pellets (NIFS, 2018), which are known to be rich in unsaturated fatty acids such as EPA and DHA (Lee et al., 2012; Okumura et al., 2055; Shulgina et al., 2019). Interestingly, EPA and DHA levels were higher in Korean than in Japanese red seabream samples (Fig. 2.3B). A combination of VIP score  $> 1$  as multivariate

statistical analysis and  $p$ -value  $< 0.05$  as univariate statistical analysis is used to identify a candidate biomarker (Abdullah et al., 2017). Therefore, the top 10 compounds are valid candidate biomarkers for distinguishing between the two red seabreams in this study.

### **2.3.3. Selection of candidate biomarker for discriminating different origins of red seabream**

VIP scores and  $p$ -values obtained using OPLS-DA and t-test were used to identify compounds that could indicate differences between the two groups. However, although this model can be used to identify biomarkers that discriminate between the two groups, it is limited in providing transparency. Therefore, a ROC curve was used to screen for robust biomarkers (Xia et al., 2013). Linoleic acid (C18:2n-6, cis) was identified as a promising biomarker by ROC curve analysis using four seasonal fish samples (Fig. 2.4). All AUC values were 1.0 with both sensitivities and specificities of 1.0. Since a model with an AUC of 0.9–1.0 is suggested to have excellent performance (Xia et al., 2013), linoleic acid was selected



**Fig. 2. 4. Biomarker analysis to identify candidate biomarkers to discriminate between Korean and Japanese red seabreams (A, autumn, n = 60; B, spring, n = 10; C, summer, n = 10; D, winter, n = 10). Linoleic acid was identified as a biomarker. All AUC values were 1.0. Significant differences were performed by t-test (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ).**

as a biomarker to differentiate between Korean and Japanese red seabream during all four seasons.

The linoleic acid (C18:2n-6, cis) content is higher in EP than in MP diets, and a study showed that linoleic acid levels in olive flounders that were fed EP diets were higher than in those fed MP diets (Kim et al., 2010). In general, the difference in the fatty acid composition of fish is attributed to the composition of the feed (Taşbozan & Gökçe, 2017). This study showed higher levels of linoleic acid in Japanese red seabream, primarily fed with EP diet, allowing us to compare the differences between the Korean and Japanese fish. Tables 2.2 and 2.3 compare Korean and Japanese red seabreams in the concentration of linoleic acid and its percentage of the total fatty acids. The cut-off values were presented based on 1.0 sensitivity and 1.0 specificity from the ROC curve. The linoleic acid content was found to be  $1.65 \pm 1.13$ ,  $0.18 \pm 0.08$ ,  $0.41 \pm 0.18$ , and  $0.36 \pm 0.06$  g/100 g for Korean red seabreams and  $6.51 \pm 1.23$ ,  $1.65 \pm 0.43$ ,  $2.27 \pm 0.42$ ,  $1.73 \pm 0.64$  g/100 g for Japanese red seabreams in the autumn (standard), spring, summer, and winter samples, respectively.

Therefore, the linoleic acid content in each season was higher in

**Table 2. 2. Comparative analysis of linoleic acid content (g/100 g) and cut-off value in fish samples collected across four seasons.**

Linoleic acid content (g/100 g)	Standard	Spring	Summer	Winter	RSD, % <sup>1</sup>	RE, % <sup>2</sup>
Korea	1.65±1.13 <sup>a</sup>	0.18±0.08 <sup>b</sup>	0.41±0.18 <sup>b</sup>	0.36±0.06 <sup>b</sup>	92.79	-60.61
Japan	6.51±1.23 <sup>a</sup>	1.65±0.43 <sup>b</sup>	2.27±0.42 <sup>b</sup>	1.73±0.64 <sup>b</sup>	49.14	-53.30
Cut-off <sup>3</sup>	3.93	0.75	1.18	0.68	94.55	-58.40

Mean and standard deviation with different letters indicate significant differences ( $p < 0.05$ ), whereas the same letters indicate no significant difference ( $p > 0.05$ ).

<sup>1</sup> Precision was evaluated using relative standard deviation (RSD, %).

<sup>2</sup> Accuracy is assessed by calculating the relative error (RE, %).

<sup>3</sup> The ROC curve of each content obtained the cut-off values according to linoleic acid concentration (g/100 g) and percentage of total fatty acids.

KPM, Korean red seabream (n = 45); JPM, Japanese red seabream (n = 45).

**Table 2. 3. Comparative analysis of linoleic acid content (% of total fatty acids) and cut-off value in fish samples collected across four seasons.**

Linoleic acid content (% of total fatty acids)	Standard	Spring	Summer	Winter	RSD, % <sup>1</sup>	RE, % <sup>2</sup>
Korea	2.70 ±1.72 <sup>a</sup>	1.55 ±0.09 <sup>b</sup>	2.57 ±0.42 <sup>a</sup>	1.83 ±0.21 <sup>a</sup>	59.56	-19.91
Japan	10.60 ±1.99 <sup>a</sup>	10.72 ±0.30 <sup>b</sup>	12.25 ±0.17 <sup>a</sup>	10.87 ±0.91 <sup>a</sup>	15.92	4.81
Cut-off <sup>3</sup>	6.45	5.98	7.58	5.94	11.78	0.58

Mean and standard deviation with different letters indicate significant differences ( $p < 0.05$ ), whereas the same letters indicate no significant difference ( $p > 0.05$ ).

<sup>1</sup> Precision was evaluated using relative standard deviation (RSD, %).

<sup>2</sup> Accuracy is assessed by calculating the relative error (RE, %).

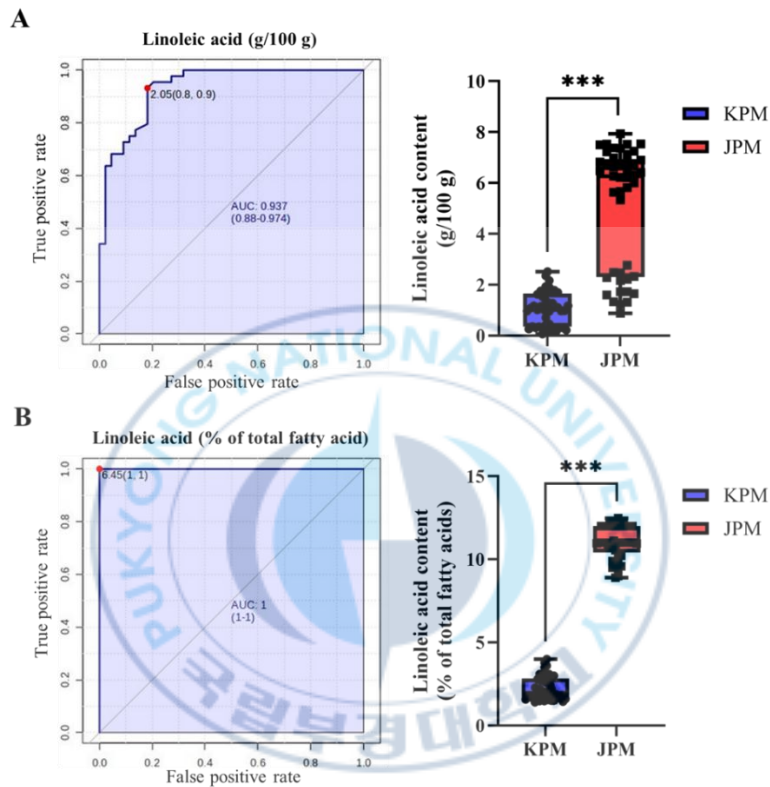
<sup>3</sup> The ROC curve of each content obtained the cut-off values according to linoleic acid concentration (g/100 g) and percentage of total fatty acids.

KPM, Korean red seabream (n = 45); JPM, Japanese red seabream (n = 45).

Japanese red seabreams. However, when comparing the linoleic acid content for each season, the results of the standard and the other seasonal samples were significantly different ( $p < 0.05$ ). In contrast, when the linoleic acid content was converted to percentages of the total fatty acids, the values were similar to those in standard fish samples ( $p > 0.05$ ), except for those collected during spring. The percentages of linoleic acid of total fatty acids were  $2.70\% \pm 1.72\%$ ,  $1.55\% \pm 0.09\%$ ,  $2.57\% \pm 0.42\%$ , and  $1.83\% \pm 0.21\%$  for Korean red seabream, and  $10.60\% \pm 1.99\%$ ,  $10.72\% \pm 0.30\%$ ,  $12.25\% \pm 0.17\%$ , and  $10.87\% \pm 0.91\%$  for Japanese red seabream in the autumn, spring, summer, and winter samples, respectively. The percentages of linoleic acid were also higher in Japanese red seabreams, similar to their concentrations. To evaluate analytical reliability, both the concentration and percentage of linoleic acid were assessed using a standard sample. The findings demonstrated that percentage-based expression provided superior accuracy and precision compared to concentration-based measurements (Table 2.2 and 2.3).

Although the seasonal concentrations of linoleic acid exhibited variability, a consistent pattern was observed when expressed as a

percentage. To assess the discriminatory performance, ROC curve analyses were conducted for both concentration and percentage values using aggregated seasonal data, comparing their effectiveness in distinguishing between the two red seabream groups across seasons (Fig. 2.5). The ROC curve is used to evaluate the sensitivity (true positive rate), specificity (false positive rate), and accuracy (AUC) of a model (Florkowski, 2008). In this study, the ROC curve for linoleic acid concentration showed a sensitivity of 0.974, a specificity of 0.877, and an AUC value of 0.938 (Fig. 2.5A). This model suggested that linoleic acid concentration is a good biomarker for discriminating between different fish species. However, the ROC curve of the percentage of linoleic acid of total fatty acids showed a sensitivity of 1.00, a specificity of 1.00, and an AUC of 1.00, suggesting that this model is more accurate for discriminating between Korean and Japanese red seabream (Fig. 2.5B). Further, the sensitivity and specificity may suggest a higher cut-off value for the percentage ROC than for the concentration ROC. Thus, linoleic acid was selected as a biomarker for distinguishing Korean and Japanese red seabream through fatty acid analysis. A value of 6.45% was determined



**Fig. 2. 5.** The ROC curves for linoleic acid across all red seabream samples including standard and seasonal groups showed an AUC of 0.937 for concentration (g/100 g) (A) and 1.0 for its percentage of total fatty acids (B). KPM, Korean red seabream (n = 45); JPM, Japanese red seabream (n = 45).

as a threshold for linoleic acid for identifying Korean (<6.45%) and Japanese red seabream (>6.45%) (Fig. 2. 5B).

#### **2.3.4. Identification of red seabream origin using linoleic acid analysis**

The proportion of linoleic acid within the total fatty acid composition was subsequently applied to an independent set of 20 red seabream samples for further validation. The additional samples were obtained as described in section 2.1. Using the 20 fish samples (10 samples for each origin), 29 fatty acids were detected using the same procedure as in section 2.2. Linoleic acid was detected in all fish samples of two different origins and identified by percentage of total fatty acids (Table 2. 4). The level of linoleic acid was higher in Japanese red seabream than in Korean fish, relatively. In Korean and Japanese red seabreams, the linoleic acid percentages were 1.28%–2.89% and 6.52%–13.67%, respectively (Table 2. 3). Using a cut-off value of 6.45 (section 2.3). The origin of all Korean and Japanese red seabream samples was accurately identified, and the

**Table 2. 4. Identification of red seabream origin using linoleic acid percentage of total fatty acid by fatty acid analysis**

Sample		Discrimination		Sample		Discrimination	
no	origin	linoleic acid (% of total fatty acid)	Identify	no	origin	linoleic acid (% of total fatty acid)	Identify
1	Korea	1.47	Korea	7	Japan	7.67	Japan
2	Korea	2.51	Korea	8	Japan	11.55	Japan
3	Korea	2.66	Korea	9	Japan	12.30	Japan
4	Korea	1.28	Korea	10	Japan	13.67	Japan
5	Japan	11.53	Japan	11	Korea	1.96	Korea
6	Japan	11.60	Japan	12	Korea	1.42	Korea

*Continued*

---

13	Korea	2.43	Korea	17	Korea	2.89	Japan
14	Japan	6.52	Japan	18	Korea	2.53	Japan
15	Japan	11.05	Japan	19	Japan	13.04	Japan
<b>16</b>	Korea	2.72	Japan	20	Japan	12.68	Japan

---



results were in full agreement with the verified origin of each sample. This result suggests that the percentage of total fatty acids of linoleic acid can be used to distinguish between Korean and Japanese red seabreams with high sensitivity and specificity.

The present study suggests a strategy for distinguishing fish origins based on fatty acid analyses of fresh fish meat samples. However, lipid oxidation during food processing is a potential limitation for the method's application to determine fish origins. Therefore, further research is required to investigate differences in lipid oxidation during the processing of seafood or fish with different origins.

#### **2.4. Conclusion**

In this study, amino acids, fatty acids, and minerals analysis were performed on cultured red seabream from Korea and Japan. Among the 17 amino acids, 29 fatty acids, and 4 minerals used for evaluation, the fatty acid analysis showed the best ability to discriminate between Korean and Japanese fish. Using PCA and OPLS-DA, this study also showed that fatty

acid profiling could discriminate between fish samples of different origins during different seasons. In particular, the top 10 fatty acids that were significantly different between Korean and Japanese fish were determined using a heatmap with hierarchical clustering analysis and were found to be unsaturated fatty acids, with VIP score  $>1.5$  and  $p < 0.05$ . Distinguished models were evaluated in each seasonal fish sample using ROC curve analysis to select a strong candidate biomarker. The linoleic acid concentration was identified as a good biomarker candidate with an ROC curve resulting in an AUC of 1.0 obtained from all seasonal samples. Moreover, the percentage of linoleic acid of total fatty acids presented a better model for the candidate biomarker. The linoleic acid percentage ROC curve suggested a higher sensitivity and specificity than concentration. A cut-off value of 6.45% was obtained from the percentage ROC curve to distinguish between Korean and Japanese red seabreams. This cut-off value accurately identified the origins of the fish when applied to 20 additional samples.

This study has confirmed the potential of linoleic acid as a biomarker for determining the country of origin of red sea bream. However, it needs

to be further validated before it can be applied in practice. Validation is recommended for developing biomarkers in clinical and non-clinical studies (Lee et al., 2006; Whitmire et al., 2011). Therefore, using fish fatty acid biomarkers to determine the origins of red seabream should be validated in future studies. In this study, linoleic acid was nominated as a candidate origin marker for distinguishing between Korean and Japanese red seabream. However, the reliance on a single fatty acid biomarker is limited by its susceptibility to environmental variation. Fatty acid levels in fish tissue can fluctuate due to several extrinsic factors, including aquaculture feed composition, which is known to vary between Korea and Japan and directly influences lipid metabolism. Seasonal differences and other environmental parameters, such as water temperature and prey availability, may further alter metabolic responses. While this study was conducted under current aquaculture conditions, these findings may not be directly generalizable to different rearing environments. Therefore, the reliability of origin discrimination based solely on linoleic acid remains constrained. To enhance robustness and accuracy, further research is needed to identify additional or complementary markers. In this context,

a non-targeted metabolomics approach simultaneously measures a broad spectrum of metabolites and provides a more comprehensive and stable biochemical fingerprint. Unlike single-compound analysis, this multi-dimensional profiling offers greater resilience to environmental variability and enhances the discriminatory power for geographic origin authentication.



# **Chapter 3. Analysis of Metabolites of Red Seabream (*Pagrus major*) from Different Geographical Origins by Capillary Electrophoresis Time-of-Flight Mass Spectrometry**

## **Abstract**

Capillary electrophoresis time-of-flight mass spectrometry (CE-TOF/MS) was utilized to profile the muscle metabolome of red seabream, aiming to explore metabolomic-based strategies for discriminating the geographical origin of the species. The metabolites were extracted using 50% acetonitrile in water. Chromatographic separation was successfully used to classify Japanese and South Korean red seabream metabolite profiles. Principal component analysis and hierarchical cluster analysis

---

This chapter was based on the previously published manuscript: Analysis of metabolites of red seabream (*Pagrus major*) from different geographical origins by capillary electrophoresis time-of-flight mass spectrometry. *Plos one*. (2022, Vol. 17).

showed a good ability to categorize the samples according to their origin. Amino acids showed the most significant quantitative difference in South Korean and Japanese muscle samples. Specifically, the levels of ornithine, anserine, histidine, L-alanine, L-glutamic acid, L-isoleucine, dimethylglycine, and L-isoleucine in Japanese red seabream samples were significantly higher than in South Korean samples. In contrast, 2'-Deoxyguanosine and inosine monophosphate levels in South Korean muscle samples were significantly higher than in Japanese red muscle samples. The compounds also presented the differences in histidine metabolism and arginine biosynthesis pathways between Korean and Japanese red sea bream. The monitored metabolite profiles suggest that South Korean and Japanese red seabreams can be identified based on amino acid levels.

### **3.1. Introduction**

Red seabream belongs to the family Sparidae; its appearance is characterized by red coloring on the back and tail along with small blue spots. Red seabream is rich in essential amino acids, thereby representing a valuable source of aquatic protein (Kim et al., 2000), so it is widely consumed in South Korea, Japan, and China, particularly in the former two countries owing to the greater development of their aqua-culture technologies. Recently, the Japanese Government proposed using ocean water to treat wastewater from the Fukushima Nuclear Power Plant. Among the radionuclides in the wastewater, including cesium 134, cesium 137, and strontium 90, tritium has raised the most significant concern (BBC news, 2021), considering the potential exposure of humans to the element through seafood consumption. Exposure to tritium may increase the risk of cancer and mutagenesis in humans (Nile et al., 2021; Tauchi et al., 2011), which inevitably raises safe-ty concerns over seafood consumption. In addition, there have been cases of misrepresentation of the origins of aquatic products. Therefore, it is important to identify the

origins of sea bream between South Korea and Japan. An earlier study elucidated the origins of the olive flounder (*Paralichthys olivaceus*) using gene polymorphism analyses (Song et al., 2004), and near-infrared reflectance spectroscopy has been used to identify the origin of sea bass (*Dicentrarchus labrax* L.) (Xiccato et al., 2004). While previous studies have compared the nutritional composition of several regional fishes, including red sea bass (Kim et al., 2002), it has been difficult to identify sea bass by geographic origin. However, nutritional composition may vary even within the same species depending on size, posing challenges to the identification of reliable origin biomarkers (Lee et al., 2022).

Metabolomics is the study of identifying and quantifying small molecules or chemicals found in cells, organs, and organisms (German et al., 2005). The small molecules can include a variety of endogenous and exogenous chemicals such as peptides, amino acids, nucleic acids, carbohydrates, organic acids, minerals, and chemicals that can be ingested or synthesized in a cell or an organism as a metabolic response. Metabolomics has become an important tool in food science, including food composition analysis, food quality assessment, and food

consumption monitoring. In particular, food component analysis consists of various steps, such as the determination of individual protein, fat, carbohydrate, dietary fiber, and ash levels; however, metabolite analysis enables the collection of more chemically detailed information (Wishart, 2008). In food authenticity research, metabolomics allows for the identification and quantification of numerous low-molecular-weight compounds, which can be statistically modeled to distinguish between different sample groups (Cubero-Leon et al., 2014). For example, metabolomic techniques have been successfully applied to differentiate sea cucumbers (*Apostichopus japonicus*) from various geographic origins (Zhao et al., 2020), and to identify regional signatures in Bokbunja fruit and beef from Wagyu and Holstein cattle (Heo et al., 2011; Yamada et al., 2020).

Numerous metabolomic technologies are currently available, including capillary electrophoresis (CE), nuclear magnetic resonance, gas chromatography, and liquid chromatography-mass spectrometry, each with advantages and disadvantages (Zhao et al., 2016). CE can be used to analyze low- and high-molecular-weight polar and non-polar compounds

and time-of-flight mass spectrometry (TOF-MS) is widely used with CE owing to its advantages of speed, wide mass range, sensitivity, and high ionization transfer efficiency (Lazar et al., 1997). In particular, CE-TOF coupling combines the efficiency and low sample consumption of CE with the high accuracy and resolution of TOF-MS. While TOF-MS provides a broad theoretical mass range, its application strength in CE-TOF/MS arises primarily from its ability to detect polar and charged small molecules with high sensitivity. Therefore, it has particular strengths in analyzing ionic and polar substances; CE-TOF/MS is considered a superior technique in biomolecular analysis (Staub et al., 2009).

Therefore, in this study, capillary electrophoresis time-of-flight mass spectrometry (CE-TOF/MS) was used to monitor the metabolomic profiles of South Korean and Japanese red seabreams in order to evaluate the feasibility of origin discrimination.

## **3.2. Material and methods**

### **3.2.1. Sample preparation for Korean and Japanese red**

## **seabream**

Japanese (n = 10) and Korean (n = 10) red seabream were used in this study. All red seabream samples were collected as fresh edible fish at the Jagalchi seafood market in Busan in October 2020. Korean red seabreams (n = 10, cultured at Korea (34°44'19" N 128°23'47" E)'s total body length and weight were measured ( $45.20 \pm 1.20$  cm,  $1.30 \pm 0.30$  kg). Japanese red seabreams (n = 10, cultured at Japan (34°4'08" N 136°13'54" E)'s total body length and weight were measured at  $46.70 \pm 1.1$  cm,  $1.40 \pm 0.20$  kg, respectively. This study was carried out in strict accordance with the guidelines on animal experimental ethics of the Korean Animal and Plant Quarantine Agency (Institutional Animal Care and Use Committee, 2020). If a part of a dead body is collected, it does not have to be deliberated by the Ethics Committee. Therefore, this experiment did not require deliberation by the Ethics Committee. All muscle samples were harvested between the dorsal and pectoral fins on the sideline and were then stored at  $-80^{\circ}\text{C}$  for the subsequent metabolite analysis. The muscle tissue samples (approximately 30 mg) were collected and mixed with 750  $\mu\text{l}$  of 50% acetonitrile in water (v/v), containing methionine sulfone and

camphor-10-sulfonic acid as internal standards (20  $\mu\text{M}$ ), and homogenized using a homogenizer (3,500 rpm, 60 s  $\times$  5 times), followed by the addition of 50% acetonitrile in water (v/v). The supernatant (400  $\mu\text{L}$ ) was then filtered through a 5-kDa cut-off filter (ULTRAFREE-MC-PLHCC; Human Metabolome Technologies, Yamagata, Japan) to remove macromolecules. The filtrate was centrifugally concentrated and resuspended in 50  $\mu\text{L}$  of ultrapure water immediately before measurement.

### **3.2.2. Instrument condition of CE-TOF/MS**

Separations were performed using a fused silica capillary column (50  $\mu\text{m} \times 80 \text{ cm}$ ) filled with cation buffer solution to detect cations and anion buffer solutions to detect anions. The extracted metabolites were analyzed using an Agilent CE-TOF/MS system (Agilent Technologies Inc., Santa Clara, CA, USA) in two modes: cationic and anionic metabolites (Das et al., 2017). The operating conditions for CE-TOF/MS analysis of metabolites are presented in Table 3.1.

**Table 3.1. CE-TOF/MS conditions for obtaining the metabolome profiles from red seabream muscle samples.**

Device	Agilent CE-TOFMS system (Agilent Technologies Inc.)	
Column	Fused silica capillary i.d. 50 $\mu\text{m}$ $\times$ 80 cm	
Condition	Cationic Metabolite	Anionic Metabolite
Run buffer	Cation Buffer Solution (p/n: H3301-1001)	Anion Buffer Solution (p/n: I3302-1023)
Rinse buffer	Cation Buffer Solution (p/n: H3301-1001)	Anion Buffer Solution (p/n: I3302-1023)
Sample injection	Pressure injection 50 mbar, 10 sec	Pressure injection 50 mbar, 22 sec
CE voltage	Positive, 30 kV	Positive, 30 kV
MS ionization	ESI Positive	ESI Negative
MS capillary voltage	4,000 V	3,500 V
MS scan range	m/z 50–1,000	m/z 50–1,000
Sheath liquid	HMT Sheath Liquid (p/n: H3301-1020)	HMT Sheath Liquid (p/n: H3301-1020)

### **3.2.3. Data processing for metabolite profiles**

Peaks detected in CE-TOFMS analyses were extracted using automatic integration software (Master Hands ver. 2.17.1.11 developed at Keio University) to obtain peak information, including m/z, migration time (MT), and peak area. Detected metabolites were based on the Human Metabolome Technologies (HMT) standard library (HMT, Yamagata, Japan). The peak area was converted to a relative peak area based on the metabolite peak area, internal standard peak area, and sample amount. The peak detection limit was determined based on the signal-to-noise ratio (S/N) of 3. Putative metabolites were then assigned to the HMT standard and known-unknown peak libraries based on m/z and MT. The tolerance was  $\pm 0.5$  min for MT and  $\pm 10$  ppm for m/z.

### **3.2.4. Statistical analysis**

Principal component analysis (PCA) was performed using SIMCA® version 18 (Sartorius, Umeå, Sweden) to explore differences in the overall metabolic profiles between Korean and Japanese red seabream groups.

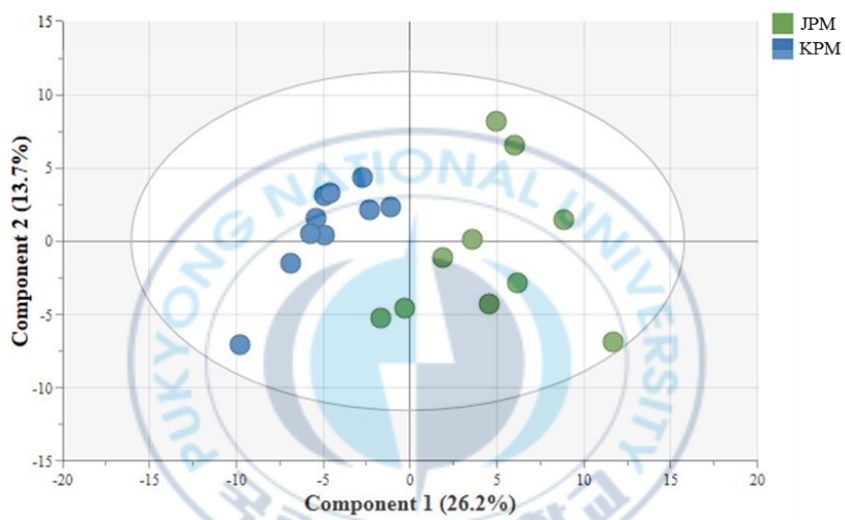
Hierarchical cluster analysis (HCA) was conducted based on a heatmap of 233 putative metabolites, and pathway enrichment analysis was performed using MetaboAnalyst 6.0 (<https://www.metaboanalyst.ca>) based on the KEGG database. To identify differential metabolites between the two groups, univariate statistical analysis (t-test,  $p < 0.05$ ) and multivariate analysis using orthogonal partial least squares discriminant analysis (OPLS-DA) were applied. Variables with high importance for projection (VIP > 1.2) and significant p-values were considered discriminative. The identified metabolites were further mapped to metabolic pathways to elucidate functional differences between the Korean and Japanese fish groups using KEGG pathway analysis.

### **3.3. Results and discussion**

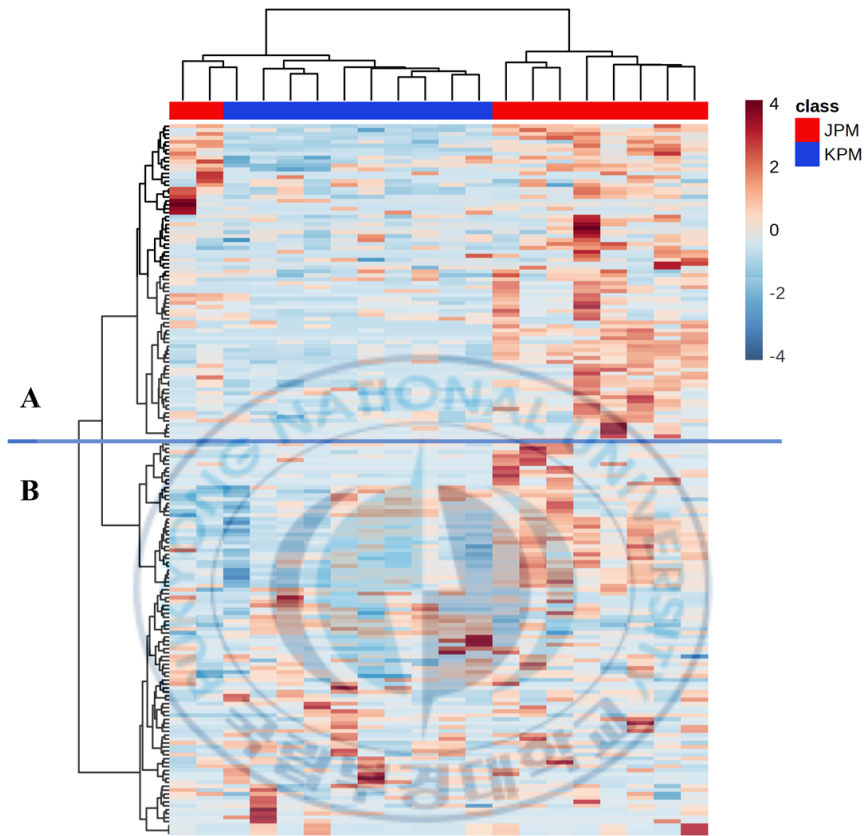
#### **3.3.1. Comparison of red seabreams from different regions by metabolite profiles**

CE-TOF/MS presented 233 putative metabolites (94 amino acids; 28 fatty acids, lipid metabolites; 17 carbohydrate metabolites; 25 nucleoside

metabolites; 21 organic acids; 19 organoheterocyclic compounds; 19 organic nitrogen compounds; 10 unknown) in South Korean and Japanese red seabream muscle samples with 171 in the cation mode and 62 in the anion mode. Principal component analysis (PCA) revealed distinct metabolic profiles associated with regional variations in South Korean and Japanese red seabream (Fig. 3.1), with the metabolite phenotype represented by components 1 and 2, accounting for 26.2% and 13.7%, respectively. By component 1, the PCA score plot of Korean red seabream samples showed distinct metabolite phenotypes, and the plot of some Japanese fish presented a slight overlap with the Korean fish group. The heatmap indicates that the Japanese fish samples were slightly mixed with Korean fish (Fig. 3.2). For Part B, the analyzed substances do not exhibit specific characteristics based on their origin. However, some compounds exhibit different characteristics based on differences in origin, as shown in Part A of the heatmap. Part A suggests that the diverse metabolome composition of fish is related to differences in metabolic processes due to differences in origin, matching those of yellowfin seabream in the fish metabolite database of KEGG. Among them, the top 103 metabolites

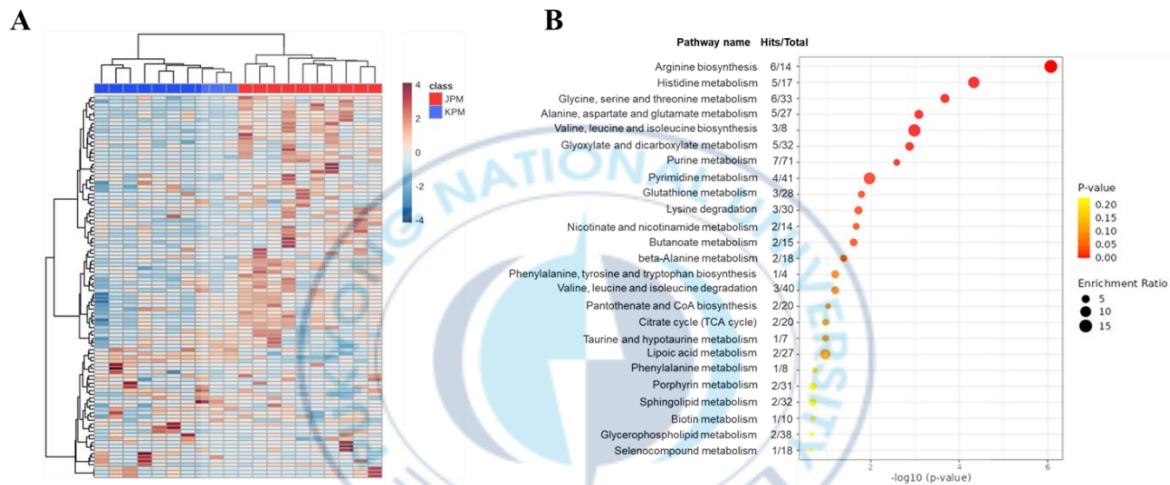


**Fig. 3. 1. PCA of metabolome data from Japanese (JPM, n = 10) and South Korean (KPM, n = 10) red seabream muscle samples.**



**Fig. 3. 2.** The heatmap with hierarchical clustering analysis of red seabream using 233 putative metabolite profiles between the different origins of Korea and Japan. KPM, Korean red seabream (n = 10); JPM, Japanese red seabream (n = 10).

exhibited different hierarchical clustering between Korean and Japanese fish, as shown in the heatmap (Fig. 3.3A). The top 25 pathways enriched in red seabream are related to 103 different metabolites (Fig. 3.3B). This result suggests that the pathways related to amino acids and peptides differ in the metabolism of red seabream of different origins. Almost pathways were influenced by dietary supplementation; 1) Arginine biosynthesis (Clark et al., 2020; Wang et al., 2021); 2) Histidine metabolism; 3) Alanine, aspartate, and glutamate metabolism (Albert et al., 2015); 4) Glycine, serine and threonine metabolism; 5) Valine, leucine and isoleucine biosynthesis; 6) Purine metabolism (Cai, e al., 2023; Yang et al., 2022). Furthermore, pathways related to glycine, serine, and threonine metabolism (Bai et al., 2016; Rahimi et al., 2020) and glyoxylate and dicarboxylate metabolism (Hu et al., 2022) have been linked to external stress factors, such as oxidative stress and cold stress, respectively. Especially, histidine metabolism is known to be influenced by various factors, such as sex, age, muscle type, dietary feeding method, and management (Boldrev, 2006; Haris et al., 2012; Tomonaga et al., 2005). The 11 relative metabolites with matched fish pathways showed difference



**Fig. 3. 3. A heatmap with hierarchical clustering of the top 103 metabolites in the different origins of red seabreams (A) and the metabolism enrichment analysis of different metabolites in red seabreams by matching the KEEG database (B). KPM, Korean red seabream (n = 10); JPM, Japanese red seabream (n = 10).**

with  $p < 0.05$  and VIPs  $> 1.3$  between fish of different origins (Table 3.2). The levels of ornithine, L-glutamic acid, anserine, histidine, dimethylglycine, L-glycine, L-alanine, L-isoleucine, L-valine, and 2'-deoxyguanosine in Japanese red seabream were significantly higher than those in Korean red seabream. In contrast, the levels of inosine monophosphate in Korean fish were significantly higher than those in Japanese red seabream.

Arginine is an essential nutrient for fish growth, but since fish cannot synthesize arginine endogenously, it must be supplied through their diet fish (Clark et al., 2020; Wang et al., 2021). Also, arginine can be synthesized in the fish urea cycle, as arginine is broken down into ornithine and urea by arginase. Ornithine reacts with carbamoyl phosphate, which is generated from glutamate, through the action of ornithine trans-carbamylase, forming citrulline. Citrulline is then involved in arginine synthesis (Wright et al., 1995). Alanine is synthesized by metabolic processes in the human body and is produced by reductive amination of pyruvic acid. similar to threonine, serine, and glycine, it is also known to be related to saccarinity (Jin et al., 2006). Furthermore, glutamic acid is

**Table 3. 2. Key differential metabolites and their associated pathways between Korean and Japanese red seabream**

Metabolites	Related pathway	Relative area		Comparative analysis		
		KPM	JPM	Fold change <sup>1</sup>	<i>p</i> -value <sup>2</sup>	VIPs <sup>3</sup>
Ornithine	Arginine biosynthesis	$1.80 \times 10^{-4}$ $\pm 4.80 \times 10^{-5}$	$5.30 \times 10^{-4}$ $\pm 2.40 \times 10^{-4}$	0.33	<0.001	1.49
L-Glutamic acid	Arginine biosynthesis; Histidine metabolism; Alanine, aspartate, and glutamate metabolism; Glyoxylate and dicarboxylate metabolism	$2.90 \times 10^{-2}$ $\pm 8.50 \times 10^{-3}$	$5.00 \times 10^{-2}$ $\pm 1.20 \times 10^{-2}$	0.58	<0.001	1.51

Values are expressed as mean  $\pm$  SD.

<sup>1</sup> The fold change was computed using the average detection values, with the values of the PMK samples as the denominators.

<sup>2</sup> The *p*-value was obtained using Welch's t-test.

<sup>3</sup> VIP score obtained by OPLS-DA

KPM, Korean red seabream (n = 10); JPM, Japanese red seabream (n = 10).

*Continued*

---

Citrulline	Arginine biosynthesis	$1.00 \times 10^{-3}$ $\pm 2.24 \times 10^{-4}$	$3.41 \times 10^{-3}$ $\pm 2.03 \times 10^{-3}$	0.29	0.005	1.49
Anserine	Histidine metabolism	$2.50 \times 10^{-3}$ $\pm 5.00 \times 10^{-4}$	$3.40 \times 10^{-3}$ $\pm 7.70 \times 10^{-4}$	0.72	0.005	1.31
Histidine	Histidine metabolism	$1.00 \times 10^{-3}$ $\pm 2.40 \times 10^{-4}$	$1.60 \times 10^{-3}$ $\pm 3.70 \times 10^{-4}$	0.65	0.001	1.37
Dimethylglycine	Glycine, serine and threonine metabolism	$9.00 \times 10^{-5}$ $\pm 2.90 \times 10^{-5}$	$4.50 \times 10^{-4}$ $\pm 1.60 \times 10^{-4}$	0.20	<0.001	1.58
L-Glycine	Glycine, serine and threonine metabolism; Glyoxylate and dicarboxylate metabolism	$5.60 \times 10^{-2}$ $\pm 8.70 \times 10^{-3}$	$8.40 \times 10^{-2}$ $\pm 2.20 \times 10^{-2}$	0.67	0.003	1.39

---

*Continued*

L-Alanine	Alanine, aspartate and glutamate metabolism	$7.80 \times 10^{-2}$ $\pm 2.60 \times 10^{-2}$	$1.10 \times 10^{-1}$ $\pm 1.90 \times 10^{-2}$	0.69	0.003	1.36
L-Isoleucine	Valine, leucine and isoleucine biosynthesis	$4.90 \times 10^{-2}$ $\pm 1.40 \times 10^{-2}$	$7.00 \times 10^{-2}$ $\pm 1.10 \times 10^{-2}$	0.002	1.40	1.57
2'-Deoxyguanosine	Purine metabolism	$5.00 \times 10^{-5}$ $\pm 1.10 \times 10^{-5}$	$7.35 \times 10^{-5}$ $\pm 1.16 \times 10^{-5}$	0.68	<0.001	1.41
Inosine monophosphate	Purine metabolism	$2.10 \times 10^{-1}$ $\pm 1.30 \times 10^{-2}$	$1.79 \times 10^{-1}$ $\pm 2.60 \times 10^{-2}$	1.19	0.002	1.33

abundant in nature and is found in various foods, including fish, seafood, pork, and beef. Glutamic acid is associated with the umami taste, utterly distinct from the four basic tastes—sweet, salty, sour, and bitter (Oruna-Concha et al., 2007; Yamaguchi & Ninomiya, 2000). Additionally, South Korean and Japanese red seabreams showed different alanine and glutamic acid contents. This difference could be attributed to the differences between their aquaculture feeds. In South Korean aquaculture, 91% of melt pellets (MP) and 9% of extrusion pellets (EP) are supplied to red seabreams, whereas in Japanese aquaculture, dry pellets (DP) and EP are mainly used (NIFS, 2014; Song, 2003). Sand eels (*Ammodytes personatus*) are used as MP in South Korean aquaculture (Kim & Kang, 1991).

In contrast, the protein and lipid levels in the compound feed used in Japanese aquaculture are 55% and 10%, respectively (Takeuchi et al., 1991). A previous study showed significant differences in the levels of glutamic acid and alanine (also related to the umami taste) in flounder fed with MP and EP; the alanine level was higher in flounder fed EP (6.64%) than in flounder fed MP (6.05%). Similarly, glutamic acid content was

higher in flounder-fed EP (15.81%) than flounder-fed MP (15.42%) (Jang et al., 2009).

Isoleucine is an essential amino acid for all fish species and is primarily deposited in the skeletal muscle proteins. In a study on the loss of meat quality due to dietary deficiency and excess isoleucine in grass carp (*Ctenopharyngodon idella*), differences in the shear force of grass carp muscle were associated with differences in isoleucine levels in the feed. The pH and shear force of their muscle were improved by increasing the isoleucine levels in the feed (3.8–18.5 g/kg diet). In addition, the isoleucine levels in grass carp muscle samples (3.25%–3.50%) can be increased by increasing the isoleucine level in the diet (Gan et al., 2014).

Valine (Val), similar to isoleucine, is a branched-chain amino acid essential for inhibiting protein synthesis and degradation in fish. Valine deficiency reportedly reduced growth performance in Mrigal carp (*Cirrhinus mrigala*) (Ahmed & Khan, 2006). In a study on the dietary valine requirements of juvenile red seabreams, growth performance was associated with the dietary valine levels (Rahimnejad & Lee, 2013). Another study on the dietary valine requirements of catla (*Catla catla*)

reported a suitable dietary valine level for improved catla growth performance (1.12–1.71% valine in 0.51% of the diet). Furthermore, catla fed a diet containing the optimal valine level showed increased high-grade protein content and valine retention and gain (Zehra & Khan, 2014). Therefore, the current valine supply in optimized mixed feed of red sea bream seems to be a factor that causes differences in growth rate and valine content between South Korean and Japanese red sea bream.

*N, N*-Dimethylglycine (DMG) acts as an antioxidant metabolite by methylation of free radicals and as a source of glycine for glutathione synthesis. It can enhance the body's antioxidant capacity, as well as improve an animal's ability to protect against oxidative stress and promote overall health (Bai et al., 2016; Rahimi et al., 2020). Increased DMG levels in rainbow trout exposed to clove oil be protective stress responses to anesthetics (Rahimi et al., 2020). Other studies have suggested that dietary DMG could improve athletic performance in horses and mice (Bai et al; Greene et al., 1996). To demonstrate that Japanese red sea bream has high DMG levels, additional data is required to determine whether it is added to the diet to improve athletic performance or produced in response

to external stress generated during the distribution process.

In the present study, the analyzed amino acids tended to be more abundant in EP-fed Japanese red seabream than in EP-fed South Korean seabream. Therefore, the differences in amino acids are considered to be due to differences in the compositions of feed supplied by South Korean and Japanese aquacultures, which is consistent with the results of a study on the effect of feed composition on the fat and amino acid composition of red seabream (Kanosu & Watanabe, 1976).

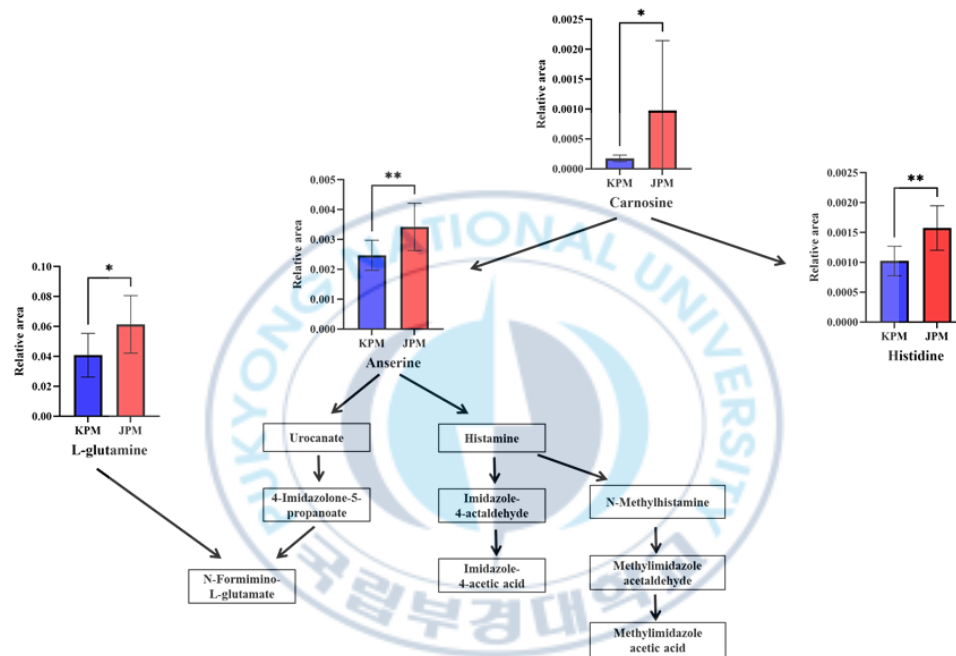
Inosine monophosphate (IMP) is associated with the umami taste, along with glutamic acid, guanosine monophosphate (GMP), and adenosine monophosphate (AMP), which participate in 5'-ribonucleotide synthesis. In particular, IMP is abundant in meat and fish, whereas GMP is more related to plants. 5-AMP, which has a lower taste intensity than IMP or GMP, is another important umami-related substance widely distributed in natural foods (Oruna-concha et al., 2007; Yamaguchi & Ninomiya, 1998; Ninomiya, 2002). A regional study on loach (*Misgurnus mizolepis*) (Kim et al., 2000) and a study on wild and cultured sweetfish (*Plecoglossus altivelis*) (Suyama et al., 1977) also showed differences in

the IMP content, which varied with the aquaculture region. The results of the present study also showed differences in IMP levels according to region. The difference in glutamic acid and IMP levels contributes to the difference in the umami taste of red seabreams obtained from South Korean and Japanese aquaculture; therefore, this taste difference is likely associated with the different metabolites produced in South Korean and Japanese red seabreams, depending on their feed composition.

### **3.3.2. Identification of metabolic difference between Korean and Japanese red seabream**

The main metabolic differences observed between Korean and Japanese red seabream (Table 3.2) involve compounds related to both histidine metabolism and arginine biosynthesis. Specifically, L-histidine, anserine, carnosine, and L-glutamic acid are associated with histidine metabolism, while ornithine, citrulline, and L-glutamic acid are key intermediates in the arginine biosynthesis pathway. These metabolites revealed distinct origin-specific patterns, particularly highlighting

significant differences in histidine-derived intramuscular dipeptides (IDPs) such as anserine and carnosine, as well as shared intermediates like L-glutamic acid (Fig. 3.4). Anserine, carnosine, and L-histidine are abundant intramuscular dipeptides (IDPs) found in both mammals and aquatic species. These compounds play crucial roles in muscle functionality, including pH buffering, antioxidant activity, increased Ca<sup>2+</sup> sensitivity, and the inhibition of protein glycation (Hiemori-Kondo et al., 2021; Wang et al., 2021). While in humans, only carnosine and L-histidine are naturally present in muscle, anserine is typically obtained through the consumption of meat or fish (Brosnan & Brosnan, 2020). The synthesis of IDPs is influenced by various factors, including temperature and feeding systems (González-Ricás et al., 2020). Generally, IDP synthesis is mediated by carnosine synthase 1, while their degradation is carried out by carnosinase 2 (Katafuchi et al., 2024). For example, high ambient temperatures in poultry reduce anserine and carnosine levels, along with a decrease in carnosine synthase 1 mRNA expression and an increase in carnosinase 2 expression (Ayumi et al., 2024). Similar trends were observed in pork under high-temperature conditions (Yang et al., 2014), where oxidative



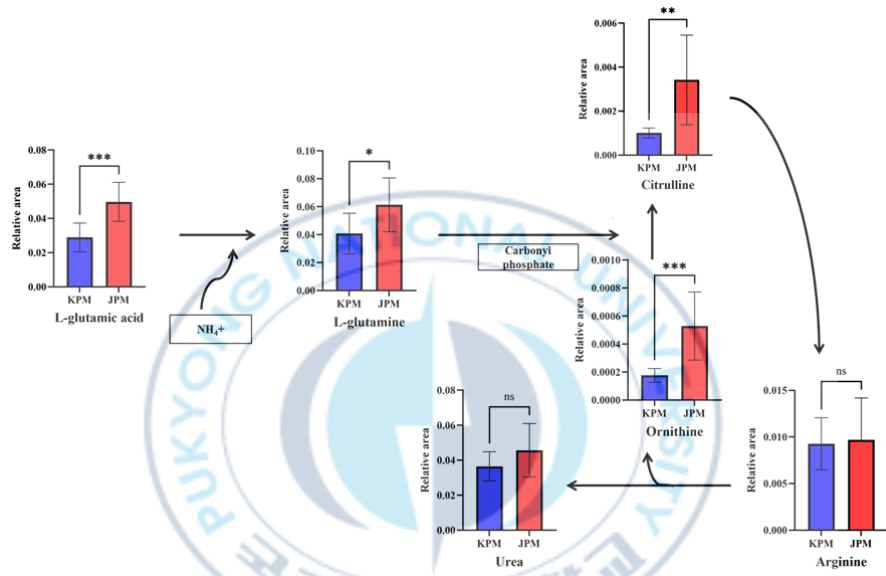
**Fig. 3. 4. Difference of histidine metabolism between Korean and Japanese red sea breams.** Red bar is Japanese red seabream (JPM, n=10) and blue bar is Korean red seabream (KPM, n=10). Significant differences were performed by t-test (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ).

stress generated by reactive oxygen species (ROS) leads to DNA damage and lipid oxidation (Peiretti et al., 2011). Conversely, low-temperature conditions act as a stress factor in red seabream (Min et al., 2017). In smolt salmon, a high-protein diet resulted in higher anserine levels compared to low- or medium-protein diets, and the muscle buffering capacity was most significant in the high-protein diet group (Ogata & Marai, 1994).

In this study, lower levels of IDPs were observed in Korean red seabream compared to Japanese red seabream, with differences in aquaculture conditions, including sea temperature and feeding systems, identified between the two regions. Generally, the optimal habitat temperature for red seabream ranges from 18°C to 25°C (Shin et al., 2018). However, Korean aquaculture farms experience water temperatures ranging from 11°C to 25°C, while Japanese farms have a slightly higher temperature range from 15°C to 27°C, highlighting a clear temperature difference between the two regions.

In teleost fish, the arginine biosynthesis pathway typically involves key precursors, including glutamate, citrulline, and arginine (Wang et al., 2021). During early developmental stages, the activity of urea cycle-

related enzymes is generally high, facilitating efficient arginine biosynthesis. However, in adult teleosts, nitrogen is primarily excreted as ammonia through the gills, reducing the necessity for a complete urea cycle (Terjesen et al., 2001; LeMoine & Walsh, 2013; Wright & Land, 1998). Consequently, enzymes such as CPS III, OTC, ASS, and ASL exhibit lower activity, with some expression detected in muscle tissues rather than in the liver (Buentello & Gatlin, 2000). Additionally, Japanese red seabream showed significantly elevated concentrations of multiple intermediates in the arginine biosynthesis pathway compared to Korean red seabream (Fig.3.5). These findings suggest that dietary differences or environmental adaptations may be influencing nitrogen metabolism. Citrulline supplementation, for instance, has been demonstrated to increase arginine levels more effectively than direct arginine intake (Clark et al., 2020). Additionally, urea cycle enzyme activities have increased in response to stressors, such as high stocking density or high pH, in various fish species (Laberge et al., 2009). Furthermore, the observed differences may be attributed to genetic divergence between brood stocks, which influences the expression of key urea cycle enzymes, including arginine:



**Fig. 3. 5. Difference of arginine biosynthesis pathway between Korean and Japanese red seabreams.** Red bar is Japanese red seabream (JPM, n=10) and blue bar is Korean red seabream (KPM, n=10). Significant differences were performed by t-test (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ).

glycine amidinotransferase, nitric oxide synthases, and arginase, as noted in other marine species (Morris, 2016; Zou et al., 2019).

Japanese aquaculture farms typically supply high-protein diets in their formulated feed (Takeuchi et al., 1991). In general, the aquaculture system for red seabream in Japan utilizes extruded pellet (EP) feed composed of various protein sources, such as fish meal, soy protein, and wheat gluten (Teru, 2006). These feeds are also formulated with a balanced composition of essential amino acids, including arginine, known for its role in growth and immune regulation, and histidine, which contributes to the biosynthesis of intramuscular dipeptides (IDPs) such as anserine and carnosine. The protein sources in the feed significantly influence growth performance, feed efficiency, survival rate, and immune function (Lim et al., 2021). This nutritionally optimized feed composition may directly affect specific metabolic pathways in red seabream, particularly those related to arginine and histidine metabolism. As a result, red seabream cultured in Japan may exhibit unique metabolite profiles reflecting region-specific physiological responses induced by their feeding regime. Therefore, the distinct metabolomic characteristics observed in Japanese-

cultured red seabream are likely associated with dietary influences on physiological metabolism and may serve as biologically relevant indicators for geographical origin discrimination.

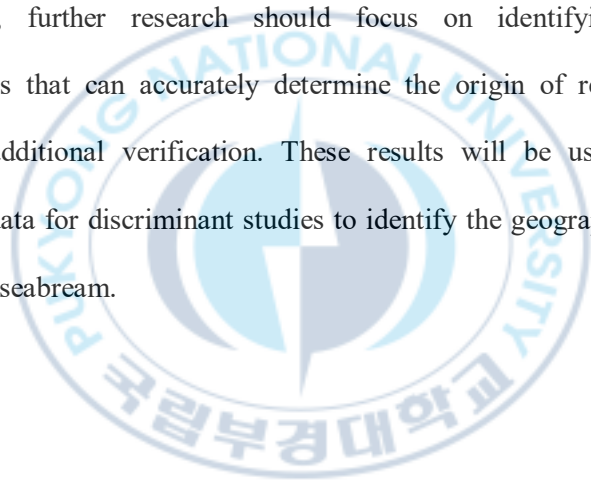
Further studies are needed to confirm these mechanisms, particularly those involving gene expression analyses and enzyme activity assays. These approaches will help determine whether the observed metabolite alterations are due to transcriptional regulation or post-translational metabolic control within the arginine biosynthesis network.

### **3.4. Conclusion**

Using metabolomics and statistical analysis, in this study, the metabolite showed a difference profiles between South Korean and Japanese seabream, especially concerning amino acids (L-alanine, L-glutamic acid, L-isoleucine, dimethylglycine, and L-valine), nucleoside metabolites (IMP), and organic nitrogen compounds (TMAO). These compounds show specific metabolisms, such as histidine synthesis and the arginine biosynthesis pathway. These distinct profiles are due to the

difference in the aquaculture feed, which is formulated to improve fish conditions and the environmental factors associated with South Korean and Japanese aquaculture, respectively.

This study demonstrated the ability of metabolite profile analysis to help identify the geographic origin of red seabream by monitoring. Therefore, further research should focus on identifying specific biomarkers that can accurately determine the origin of red seabream through additional verification. These results will be used as basic research data for discriminant studies to identify the geographical origin of the red seabream.



## **Chapter 4. Machine Learning-Based Biomarker Discovery for Origin Discrimination of Cultured Red Seabream Using Metabolomics**

### **Abstract**

This study aimed to assess the impact of feature selection for biomarkers discovery of red seabream (*Pagrus major*) from Korea and Japan using random forest. The metabolite data were obtained from a previous study using capillary electrophoresis time-of-flight mass spectrometry (CE-TOF/MS), yielding 233 putative metabolites. The data were preprocessed using sum normalization, log transformation, and Pareto scaling. An OPLS-DA model was initially constructed to assess the overfitting model with raw metabolic data. Validation of the model via permutation testing revealed a negative  $Q^2$  slope ( $Q^2 = -0.213$ ), while the  $R^2$  slope remained close to 1 ( $R^2 = 0.972$ ), suggesting possible overfitting.

Random Forest (RF), an ensemble-based learning algorithm, was applied due to its robustness against high-dimensional data to address overfitting observed in multivariate models. When trained on the full metabolomics dataset, the RF model achieved a strong classification performance with an out-of-bag (OOB) error rate of 5%. However, instability in tree formation was observed due to the small sample size and complexity of the data. To reduce dimensionality, feature selection was performed using volcano plot criteria based on FDR-adjusted  $p$ -values ( $< 0.05$ ) and  $|\log_2$  fold change| ( $> 2$ ). The RF model trained on the reduced feature set demonstrated improved variable importance and enhanced predictive performance (OOB error rate of 0%). Notably, the average effect size (Cohen's  $d$ ) increased from 0.32 to 0.42, and all selected variables showed strong statistical reliability with bootstrap  $p$ -values  $< 0.001$ . Among the selected features,  $N,N$ -dimethylglycine, creatine, and hypotaurine showed consistent discriminative potential, supported by both statistical validation and biological relevance. These results support using univariate filtering as an effective dimensionality reduction strategy in small-sample metabolomics studies for biomarker discovery. Further validation with

larger sample sizes and seasonal data is required to ensure generalizability and biomarker robustness.



## 4.1. Introduction

In Chapter 3, a comparative metabolomic analysis was conducted to explore metabolic differences between Korean and Japanese cultured red seabream using CE-TOF/MS. Although 11 statistically significant metabolites were identified, further validation is required to assess their potential as robust biomarkers for geographical origin discrimination

In metabolomics research, chemometric methods, such as multivariate statistical analyses, are commonly employed to handle the complexity of high-dimensional datasets and to identify patterns that distinguish between biological groups. Given the high dimensionality and multicollinearity often present in metabolomic datasets, orthogonal partial least squares discriminant analysis (OPLS-DA) was utilized to reduce dimensionality while maximizing group separation (Liu et al., 2016). This supervised technique not only facilitates the construction of predictive classification models but also yields variable importance in projection (VIP) scores, which indicate the relative influence of each metabolite. Metabolites with VIP scores greater than 1.0 were considered candidate biomarkers (Yang et al., 2019). OPLS-DA has also been widely used in

disease diagnosis and biomarker discovery (Bergholt et al., 2014; Yang et al., 2021; Ildiz et al., 2021). However, overfitting of the predictive model can limit the reliability of the biomarker. In addition to multivariate analysis, random forest (RF), a machine learning technique based on ensemble learning, was employed due to its high classification performance and ability to handle complex, nonlinear relationships within metabolomic data (Chen et al., 2013; Fan et al., 2011; Patterson et al., 2011). This approach reduced the risk of overfitting and improved classification stability in high-dimensional datasets by using multiple decision trees with randomly selected subsets of variables at each node (Breiman, 2011). RF was used to rank feature importance and refine the selection of biomarkers. However, due to the relatively small sample sizes typical in metabolomics studies, there remains a significant risk of model overfitting, limiting the generalizability of the findings (Grissa et al., 2016). To address this, feature selection is a key strategy for reducing data dimensionality by eliminating non-informative variables that may negatively impact the robustness and predictive accuracy of learning-based models (Acharjee et al., 2020). While model-based methods such as

Boruta can be unstable and prone to overfitting when applied to limited datasets (Xu et al., 2022). In contrast, univariate filtering based on FDR-adjusted p-values and fold change provides a simple, interpretable, and robust alternative, particularly well-suited for small-sample studies. This approach prioritizes statistically and biologically meaningful variables, enhancing model stability during downstream analysis (Haury et al., 2011).

Although Chapter 3 demonstrated significant differences in metabolite levels between red seabream originating from South Korea and Japan through comparative metabolomic profiling, this chapter focuses on the development and evaluation of a biomarker discovery strategy. Specifically, this approach integrates machine learning-based classification, statistical feature selection, and validation techniques. Given the high dimensionality and limited sample size ( $p \gg n$ ) inherent in the metabolomics dataset, appropriate feature selection and modeling strategies were employed to enhance model stability and ensure reliable discrimination of red seabream by geographical origin.

## **4.2. Materials and methods**

#### **4.2.1. Data preprocessing for metabolite profiles normalization**

Raw metabolite data obtained from CE-TOF/MS analysis from Korean and Japanese red seabream (n=20) were preprocessed as described in Chapter 3. To reduce measurement variation across the entire sample, metabolite data were preprocessed using a three-step pipeline: sum normalization, log transformation, and Pareto scaling. Sum normalization was applied to correct for differences in total metabolite abundance among samples. Subsequently, log transformation was performed to reduce right-skewness and minimize the influence of extreme values. Finally, Pareto scaling was used to standardize the data by dividing each variable by the square root of its standard deviation, thereby reducing the impact of significant variance while preserving biological information. These preprocessing steps enhanced sample comparability and ensured reliable statistical analysis (Robert et al., 2006).

#### **4.2.2. OPLS-DA modeling and overfitting assessment**

Orthogonal partial least squares discriminant analysis (OPLS-DA) was

conducted using SIMCA ver. 18 (Sartorius Stedim Data Analytics AB, Umeå, Sweden) to evaluate the separation between red seabream samples from South Korea and Japan. Model performance was assessed by  $R^2Y$  (explained variance) and  $Q^2$  (predictive accuracy). The OPLS-DA model was validated by cross-validated ANOVA (CV-ANOVA,  $p < 0.05$ ) test and permutation test ( $n = 200$ ) to assess model robustness and potential overfitting (Eriksson et al., 2008).

#### **4.2.3. Random forest classification modeling**

Random forest classification was conducted using MetaboAnalyst 6.0 (<https://www.metaboanalyst.ca>) to identify key metabolites that distinguish between the geographical origins of red seabream. Random forest classification was performed with a 500-decision-tree algorithm. Variables were ranked according to the mean decrease in accuracy, and model performance was assessed using the out-of-bag (OOB) error rate and a confusion matrix, which provided an estimate of the model's predictive accuracy and robustness.

#### **4.2.4. Feature selection via statistical criteria**

Feature selection was conducted to prevent overfitting caused by the high dimensionality and low sample size ( $p \gg n$ ) nature of the metabolomics dataset. Its primary role was to eliminate noisy or irrelevant variables that could compromise the robustness and predictive performance of downstream classification models. Univariate analysis was applied using volcano plot criteria, combining false discovery rate (FDR)-adjusted  $p$ -values ( $< 0.05$ ) and  $|\log_2 \text{fold-change}|$  thresholds ( $> 1.0$ ), to filter out features with low statistical significance and minimal group separation. Welch's  $t$ -test was used to identify significantly different metabolites between groups, and multiple testing correction was performed using the Benjamini-Hochberg FDR method. Fold change was calculated based on the ratio of average detection values, with Korean samples used as the reference group. Volcano plots were visualized with  $-\log_{10}(p)$  on the  $y$ -axis and  $\log_2$  fold change on the  $x$ -axis (Li, 2012).

#### **4.2.5. Biomarker validation using statistical analysis**

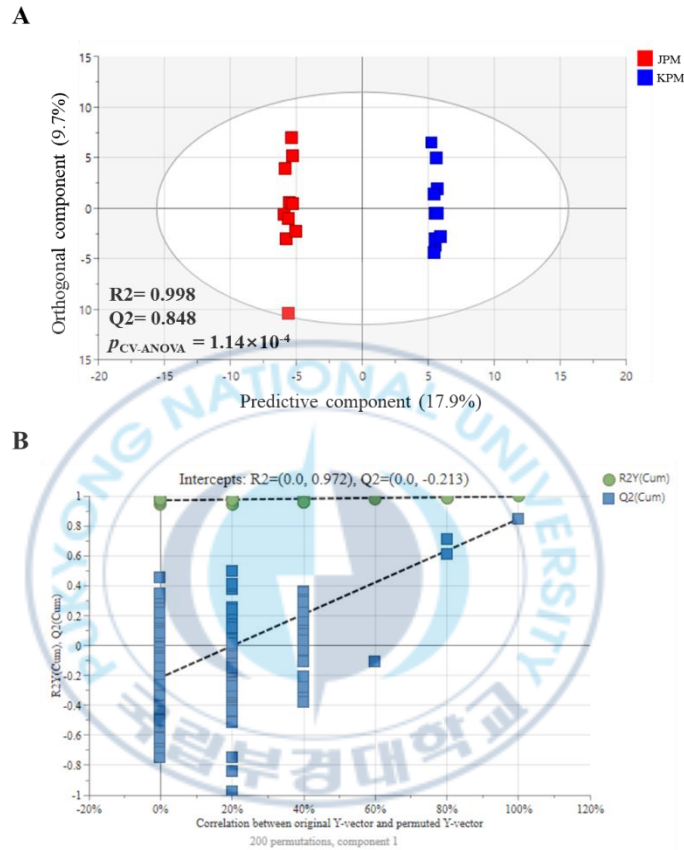
Two statistical analyses, Cohen's  $d$  and bootstrap resampling, were performed with SPSS software (version 27, SPSS Inc. Chicago, IL, USA) to compare and evaluate the reliability and robustness of the identified biomarkers. Cohen's  $d$  was calculated for each metabolite to quantify the effect size between Korean and Japanese groups. A bootstrap resampling approach (5,000 iterations) was used to assess the selection stability of each feature.

### **4.3. Results and discussion**

#### **4.3.1. Evaluation of the OPLS-DA model to distinguish Korean and Japanese cultured red seabream**

Based on the metabolic differences observed in Chapter 3, candidate biomarker identification was performed using multivariate statistical analysis with a normalized dataset of 233 metabolites from Korean and Japanese red seabream. In this Chapter, data preprocessing was carried out using sum normalization, log transformation, and Pareto scaling. This normalization strategy aimed to reduce unwanted variations caused by

experimental and instrumental conditions while preserving the inherent biological information from the raw metabolite profiles (Misra, 2020). This approach differentiates from earlier analyses that utilized unprocessed metabolite data in Chapter 3, thereby enhancing the reliability and interpretability of the subsequent biomarker discovery process. OPLS-DA is a supervised multivariate analysis method suitable for biomarker selection and classification of two groups (Wheelock & Wheelock, 2013). In this study, OPLS-DA revealed a clear separation between red seabream samples from Japan (JPM) and South Korea (KPM), indicating distinct metabolic profiles (Fig. 4. 1A). The OPLS-DA model yielded a predictive component explaining 17.9% of the total variance, with an orthogonal component accounting for 9.7%. The model was statistically significant, as confirmed by CV-ANOVA ( $p = 1.14 \times 10^{-4}$ ), indicating robust discrimination between red seabream samples from different geographical origins. A permutation test ( $n = 200$ ) was conducted to evaluate the robustness and reliability of the OPLS-DA model (Fig. 4. 1B). The permuted model showed a high  $R^2$  value (0.972), indicating a



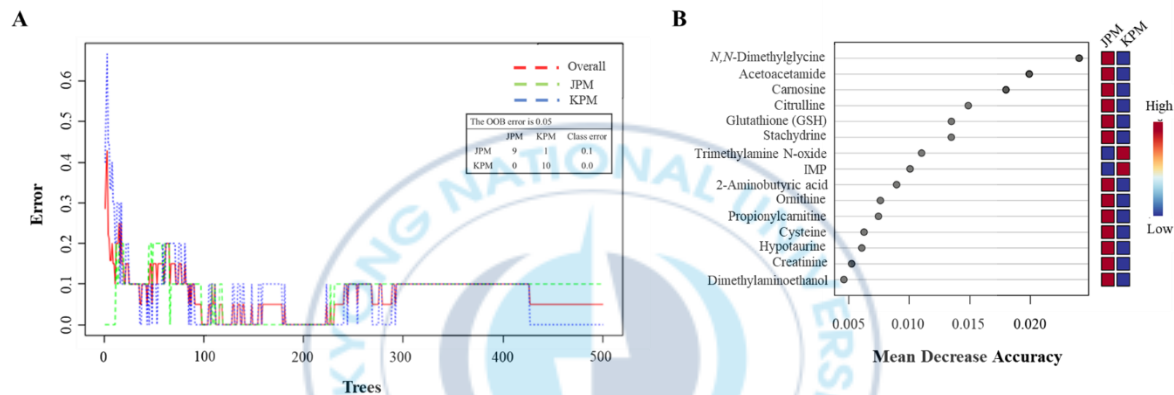
**Fig. 4. 1. Evaluation of a predictive model of OPLS-DA model for discrimination of Korean and Japanese red seabreams.** CV-ANOVA confirmed the model's statistical significance ( $p = 1.14 \times 10^{-4}$ ) (A). The permutation test plot based on 200 iterations was evaluated by overfitting (B). KPM, Korean red seabream ( $n = 10$ ); JPM, Japanese red seabream ( $n = 10$ ).

strong model fit. In contrast, the  $Q^2$  value was negative (-0.213), with a regression line showing a downward slope. For a valid OPLS-DA model, the original model's  $R^2$  and  $Q^2$  values should exceed those of the permuted models (Triba et al.,2015). Although the OPLS-DA model showed a high  $R^2$  value (0.972), indicating a strong fit to the data, this metric alone cannot confirm model validity due to the risk of overfitting. In permutation testing, even randomly assigned class labels can yield high  $R^2$  values if the model is excessively flexible. In contrast,  $Q^2$  values are based on predictive accuracy via cross-validation and tend to decline with increasing label permutation. The original model's  $Q^2$  is significantly higher than that of permuted models, and a negative slope of the  $Q^2$  regression line confirms that the model captures meaningful class differences rather than noise.

#### **4.3.2. Classification performance of random forest using full CE-TOF/MS dataset from Korean and Japanese red seabream**

Although the OPLS-DA model achieved high classification accuracy,

permutation testing revealed signs of overfitting. Therefore, a machine learning-based Random Forest model was subsequently applied to validate the predictive robustness and reduce bias. RF is an ensemble learning method based on classification and regression trees. It constructs multiple decision trees with randomly selected subsets of variables at each node, thus reducing the risk of overfitting and enhancing classification stability in high-dimensional datasets (Breiman, 2011; Degenhardt et al., 2019; Nicodemus & Malley, 2009). A classification model was built using all 233 metabolite features without prior filtering. The RF model, constructed with 500 decision trees, achieved a low out-of-bag (OOB) error rate of 5.0%, indicating strong internal classification performance (Fig. 4.2A). The top 15 features were ranked based on the mean decrease in accuracy (MDA), with importance values ranging from 0.004 to 0.024, including *N*, *N*-Dimethylglycine, acetoacetamide, carnosine, citrulline, glutathione (GSH), stachydrine, trimethylamine *N*-oxide, inosine monophosphate, 2-aminobutyric acid, ornithine, propionylcarnitine cysteine, hypotaurine, creatine, and dimethylaminoethanol (Fig. 4.2B).



**Fig. 4. 2. Random forest classification of full dataset to distinguish different origins of red seabream.** The supervised learning was performed with 500 decision trees with an out-of-bag (OOB) error of 5% (A). Fifteen key features were identified based on mean decrease accuracy rankings (B). KPM, Korean res seabream (n = 10); JPM, Japanese red seabream (n = 10).

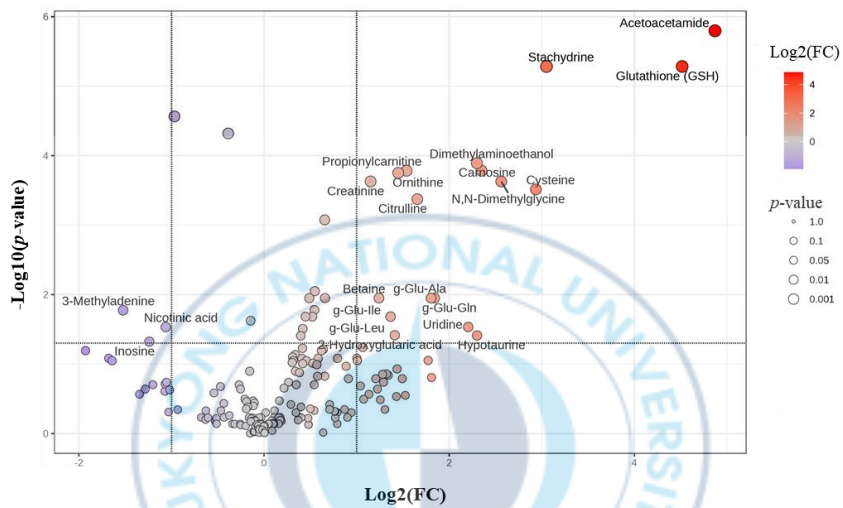
These features were initially considered key variables contributing to the classification model for origin discrimination. As the previous OPLS-DA model exhibited signs of overfitting, the potential risk remains particularly high in datasets with a small number of samples and a large number of features (Acharjee et al., 2020). In machine learning applications involving large and complex datasets, such as metabolomics, feature selection is a crucial strategy for reducing dimensionality and minimizing overfitting (Acharjee et al., 2020; Grissa et al., 2016). Therefore, a univariate feature selection method was applied to reduce data dimensionality and improve model robustness and interpretability.

#### **4.3.3. Impact of feature selection on classification model stability from metabolic data of Korean and Japanese red seabream**

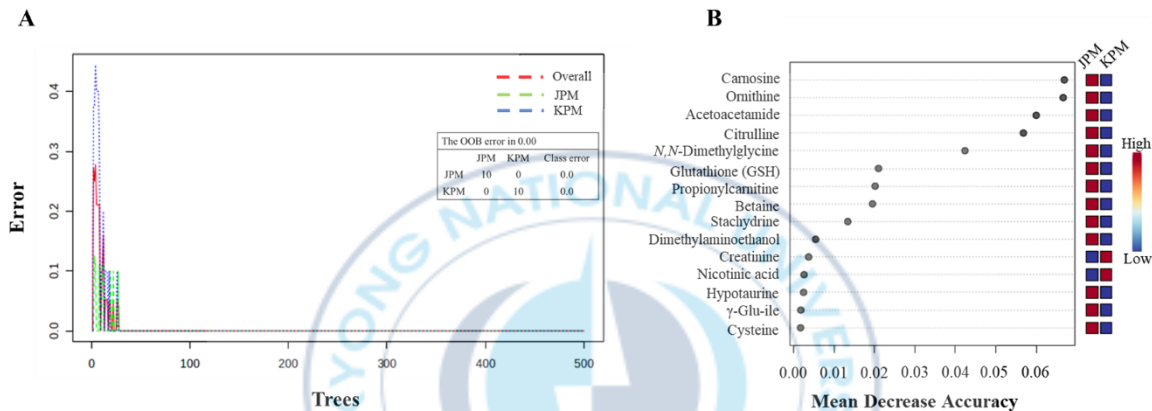
Feature selection is a widely used strategy to prevent overfitting by removing variables that contribute little to model performance, thereby retaining only those that are most informative for classification (Grissa et

al., 2016). To evaluate how feature selection influences model performance, this study compared classification accuracy between a model trained on the full high-dimensional dataset, where the number of variables far exceeds the number of samples ( $p \gg n$ ), and a model built using a reduced set of features obtained through statistical filtering. Features deemed statistically insignificant or potentially detrimental to model stability were filtered using volcano plot criteria that combined fold change and statistical significance (Cumbo et al., 2025; Rabiei et al., 2023). In this study, metabolites were selected based on a fold change  $\geq 2.0$  and an FDR-adjusted  $p$ -value  $< 0.05$ , and volcano plots were generated accordingly (Fig. 4. 3). 18 metabolites showed significant differences between the Korean and Japanese red seabream groups with three specific compounds from Korean red seabream and 15 specific compounds from Japanese red seabream.

In this study, the number of features was reduced to 18 using volcano plot filtering to mitigate the risk of overfitting, despite the small sample size. The RF model was constructed using 500 decision trees with a minimized dataset. As the number of trees increased, the out-of-bag (OOB)



**Fig. 4. 3. Volcano plot for feature selection from 233 metabolic data to approach biomarker discovery.** The volcano plot was conducted by filtration with  $\log_2(\text{fold change } 2.0)$  and  $-\log_{10}(p < 0.05)$  based FDR.

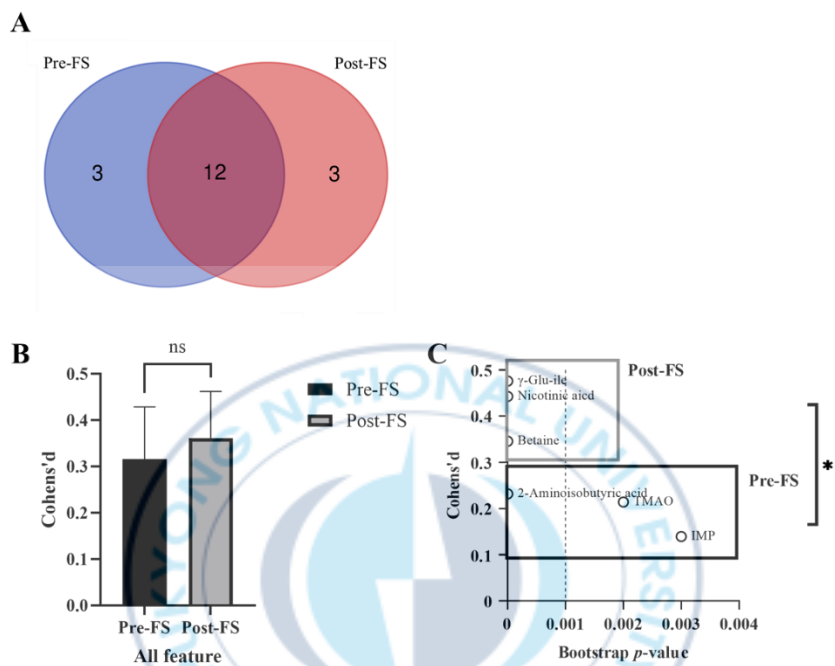


**Fig. 4. 4. Machin learning-based biomarker to distinguish different origins of red seabream validation.** Random forest classification was used to evaluate biomarkers distinguishing red seabream origins with 500 trees (A). The model showed a low out-of-bag (OOB) error of 5%, indicating high predictive accuracy. Fifteen key features were identified based on mean decrease accuracy rankings (B). KPM, Korean res seabream (n = 10); JPM, Japanese red seabream (n = 10).

error gradually decreased, indicating improved model stability and generalization performance (Fig. 4. 4A) compared with the pre-RF model (Fig. 4. 2A). Additionally, the post-RF model achieved a low out-of-bag (OOB) error rate of 0.0%, with perfect classification accuracy for KPM and JPM samples (class error = 0%). To identify the most relevant RF analysis was applied to explore variable combinations and rank metabolite features based on their predictive importance with 15 compounds, including carnosine, ornithine, acetoacetamide, citrulline, *N*, *N*-Dimethylglycine, glutathione (GSH), propionylcarnitine, stachydrine, dimethylamino-ethanol, creatine, nicotinic acid, hypotaurine,  $\gamma$ -Glu-ile, and cysteine (Fig. 4. 4B). The MDA range increased from 0.004 - 0.024 to 0.002 - 0.067 following feature selection. The Benjamini–Hochberg false discovery rate (FDR) method is widely applied in multi-omics research to control for false positives that may arise from multiple hypothesis testing. By adjusting *p*-values, it allows the selection of statistically reliable features from high-dimensional datasets (Benjamini et al., 2024; Markussen et al., 2022; Xue et al., 2023). In parallel, fold change (FC) analysis identifies features with biologically meaningful

differences between experimental and control groups, while features with minimal or no change are considered noise. Therefore, dimensionality reduction strategies that incorporate FDR and FC thresholds effectively minimize overfitting caused by noisy variables, thereby improving model generalizability in machine learning-based classification (Cumbo et al., 2025; Rabiei et al., 2023). In this study, dimensionality reduction enabled the random forest model to become more focused on informative variables, thereby improving interpretability and model transparency, as illustrated by the variable importance rankings (Fig. 4.2A and Fig. 4.4A).

Furthermore, to assess the impact of feature selection on model robustness and feature stability, the top 15 selected features were compared using Cohen's  $d$  and bootstrap analysis (Fig. 4.5). These metrics evaluated the effect size and the stability of group-level differences. In the 15 selected features of each models, 12 variables were common while three were unique to each model (Fig. 4. 5A). When comparing the effect sizes of selected features before and after dimensionality reduction, there was a trend toward an increase in effect size with a  $d$ -value of  $0.32 \pm 0.11$  before reduction and  $0.36 \pm 0.10$  after reduction, but no significant change (Fig. 4. 5B.). However, the



**Fig. 4. 5. Comparison of top 15 features selected by random forest before and after feature selection.** Venn diagram showed overlapping 12 and unique three features (A). Average of Cohen's  $d$  value of top 15 features in each model (B). Comparative analysis with bootstrap-based  $p$ -values and Cohen's  $d$  of unique variables from each model. Pre-FS, before feature selection; Post-FS, after feature selection.

effect size of the selected features increased significantly after dimensionality reduction, with a mean Cohen's  $d$  of  $0.32 \pm 0.05$  before reduction and  $0.42 \pm 0.07$  after reduction. Cohen's  $d$  is defined as the difference in means between two groups divided by the pooled standard deviation and is commonly used to quantify the discriminative power between groups (Slinger et al., 2023). It is also applied to evaluate clinically or biologically meaningful differences based on standardized thresholds (Loth et al., 2021). Cohen's  $d$ -suggested effect sizes are as follows:  $d (.01)$  = very small,  $d (0.2)$  = small,  $d (0.5)$  = medium,  $d (0.8)$  = large,  $d (1.2)$  = very large, and  $d (2.0)$  = huge (Sawilowsky, 2009). Although the effect size after feature selection was still below the medium threshold, it demonstrated a meaningful improvement compared to the entire model. This supports that feature selection contributed to eliminating non-informative variables and enhancing the quality of the selected features in the classification model. In addition to effect size, the statistical stability of the selected features was evaluated using bootstrap resampling (Fig. 4. 5C). To address the limited statistical reliability due to the small sample size, a bootstrap method was employed with 5,000 resamples to estimate the distribution of the model parameters (Hesterberg,

2011). All features selected after dimensionality reduction exhibited highly stable selection frequencies, with bootstrap-based  $p$ -values less than 0.001. In contrast, two features from the full model (prior to feature selection) showed lower selection stability ( $p = 0.01$  and  $0.02$ ). These findings suggest that dimensionality reduction improved the discriminative power of individual features and enhanced the statistical reliability of the selected biomarkers. By removing noise and reducing model complexity, the feature selection process effectively improved the interpretability and robustness of the classification model.

#### **4.3.4. Identification of candidate biomarkers to distinguish Korean and Japanese cultured red seabream**

Based on the refined random forest (RF) model, statistical indices, including Cohen's  $d$ , bootstrap  $p$ -value, and fold change, were compared and analyzed to evaluate the reliability and discriminative performance of the selected features. Among the top-ranked variables, three features,  $N$ ,  $N$ -dimethylglycine, creatine, and hypotaurine, were identified as robust

biomarker candidates, meeting the criteria of effect size (Cohen's  $d$ )  $\geq 0.4$  and a bootstrap  $p$ -value  $< 0.001$  (Table 4.1). The three selected features were all specific to the Japanese farmed red seabream group, exhibiting  $\log_2$  fold changes greater than 1 and FDR-adjusted  $p$ -values below 0.005. These variables showed Cohen's  $d$  values approaching 5, indicating a large effect size and strong group separation (Sawilowsky, 2009). Although the small sample size posed limitations in biomarker reliability, bootstrap resampling ( $n = 5,000$ ) confirmed statistically significant differences between groups for all three features ( $p < 0.001$ ), supporting their robustness (Hesterberg, 2011). In addition to statistical validation, biological relevance was examined to support the validity of the selected features as candidate biomarkers (Broadhurst & Kell, 2006; Goodacre et al., 2004).

*N, N*-Dimethylglycine (DMG) is a methylated derivative of glycine involved in glycine, serine, and threonine metabolism (Zhang et al., 2022). Previous studies have reported that levels of DMG, along with dimethylamine and phosphocholine, are altered in fish fed plant-based alternative protein diets, suggesting a dietary sensitivity of these metabo-

**Table 4. 1. Evaluation of statistical metrics for biomarker selection to distinguish the origins of Korean and Japanese red seabreams**

Compounds	$\text{Log}_2(\text{Fc})$ <sup>1</sup>	$p$ - adjusted ( $\times 10^{-2}$ ) <sup>2</sup>	Cohen's $d$ <sup>4</sup>	Bootstrap- $p$ <sup>5</sup>
<i>N, N</i> -Dimethylglycine	2.35	0.003	0.4253	<0.001
Cysteine	1.42	0.007	0.4194	<0.001
Hypotaurine	1.87	0.575	0.5403	<0.001

<sup>1</sup> The fold change (FC) with  $\text{log}_2$  was computed using average detection values

<sup>2</sup> The  $p$ -adjusted value is a multiple testing-corrected  $p$ -value, typically obtained via the Benjamini-Hochberg FDR method.

<sup>3</sup> Mean decrease accuracy (MDA) was obtained by random forest

<sup>4</sup> Bootstrap-based  $p$ -value was employed with 5,000 resamples.

lites (Roques et al., 2020). Moreover, exposure to environmental stressors such as the pesticide diazinon has been shown to reduce DMG levels in fish, while DMG supplementation may help mitigate some of the associated physiological effects (Hajirezaee et al., 2019).

Creatine plays a crucial role in cellular energy metabolism through the phosphocreatine system, buffering ATP levels during periods of high energy demand (Wyss & Kaddurah-Daouk, 2000). It reflects muscle metabolic activity and is sensitive to changes in muscle development, nutritional status, and physical stress (Villasante et al., 2023). A diet high in plant proteins can lead to low creatine levels, as a lack of arginine, glycine, and methionine can result in differences in endogenous synthesis due to deficiencies in these precursors (Wuertz & Reiser, 2023).

Hypotaurine is a sulfur-containing amino acid derivative formed as an intermediate in the cysteine-sulfinic acid pathway (Aruoma et al., 1988). It functions as a potent antioxidant and radical scavenger, particularly under stress conditions in fish (Martínez-Páramo et al., 2013). In rainbow trout, hypotaurine levels were shown to vary significantly across tissues, with notably high accumulation in the liver following dietary cystine

supplementation (Yokoyama & Nakazoe, 1998). These findings suggest that hypotaurine is highly responsive to dietary amino acid composition and may serve as a sensitive biomarker reflecting the nutritional and physiological status of fish.

There are notable differences in feeding practices between Korean and Japanese aquaculture systems for red sea bream. While raw fish-based feed is commonly used in Korea, Japanese farms predominantly use formulated compound feeds. These differences likely reflect distinct nutritional strategies and farm management approaches. The selection of the identified features may reflect the metabolic responses to these dietary regimes, thereby supporting their biological relevance. This biological link reinforces the potential utility of these features as biomarkers for distinguishing fish according to their origin in aquaculture.

#### **4.4. Conclusion**

This study demonstrated a comprehensive strategy for biomarker discovery using metabolomic datasets derived from Korean and Japanese farmed red seabreams, analyzed via CE-TOF/MS. While traditional

multivariate methods, such as OPLS-DA, are susceptible to overfitting in high-dimensional, small-sample data, the application of random forest (RF) offers improved classification performance with a reduced risk of overfitting. Initial RF modeling, using all 233 metabolic features, achieved high accuracy (OOB error: 5%), but including irrelevant variables limited the model's robustness. Dimensionality reduction using volcano plot-based feature selection (FDR-adjusted  $p$ -value  $< 0.05$ ,  $|\log_2FC| > 2$ ) enhanced model stability, reducing the OOB error to 0% and increasing variable importance scores. Statistical validation confirmed that the effect size (Cohen's  $d$ ) of selected features increased from 0.32 to 0.42, with all selected variables showing significant bootstrap  $p$ -values ( $< 0.001$ ).  $N,N$ -dimethylglycine, creatine, and hypotaurine were identified as promising biomarkers, supported by both statistical metrics and biological relevance. While Chapter 3 identified statistically significant metabolites using univariate and multivariate analysis, the feature selection and machine learning-based approach in Chapter 4 refined biomarker discovery by focusing on model robustness and biological interpretability. The emergence of creatine and hypotaurine, initially not ranked highest,

underscores the added value of integrative approaches in identifying robust candidate biomarkers using a discrimination model.

This Chapter proposes an interpretable and reproducible framework for metabolomics-based biomarker discovery in food authenticity. Nonetheless, further validation using larger, seasonally balanced sample sets and targeted quantification is essential to confirm the practical utility of the proposed biomarkers.



## Summary

Red seabream (*Pagrus major*) is a major aquaculture species and is a widely consumed seafood product in Korea, Japan, and China. Approximately 99% of the imported red seabream in Korea originates from Japanese aquaculture, significantly influencing the domestic seafood market. However, public concern over marine contamination caused by radioactive wastewater discharge from Japanese nuclear power plants has increased consumer avoidance of imported seafood. Consequently, incidents of mislabeling or omission of origin labeling, particularly for red seabream, have been reported in domestic markets. Despite governmental efforts to enforce origin labeling regulations, there remains a lack of scientific methods to distinguish the geographical origin of seafood, unlike agricultural products. This study aimed to develop a robust method for discriminating the geographical origin of red seabream using chemical composition analysis and metabolomics.

In the primary study, Fatty acid, amino acid, and mineral compositions of Korean and Japanese red seabream were analyzed. 29 fatty acids, 17

amino acids, and four minerals were quantified. Hierarchical clustering and orthogonal partial least squares discriminant analysis (OPLS-DA) revealed significant group separation based on fatty acid profiles. Among these, linoleic acid was identified as a key marker with high sensitivity and specificity through receiver operating characteristic (ROC) analysis, with a threshold of 6.45% (relative to total fatty acids) enabling origin discrimination.

Additionally, non-targeted metabolomics using CE-TOF/MS was conducted to evaluate the physiological differences between Korean and Japanese red seabream. 233 metabolites were identified and analyzed using principal component analysis (PCA) and hierarchical clustering analysis (HCA), which revealed distinct metabolic profiles based on geographical origin. Among these, 11 metabolites exhibited statistically significant differences, and KEGG pathway analysis indicated strong associations with histidine metabolism and arginine biosynthesis. These findings support the feasibility of origin discrimination based on metabolomic profiles.

Finally, to propose a biomarker discovery strategy, metabolomic data

from Korean and Japanese cultured red seabream were analyzed using a random forest (RF) classification approach. To identify potential biomarkers, the metabolite dataset was preprocessed using sum normalization, log transformation, and Pareto scaling. To address the risk of overfitting associated with high-dimensional and small-sample datasets, feature selection was performed using univariate filtering criteria based on FDR-adjusted  $p$ -values ( $< 0.05$ ) and  $|\log_2$  fold change| ( $> 2$ ). The RF model trained on the reduced dataset demonstrated improved model stability and predictive performance, with the out-of-bag (OOB) error rate decreasing from 5% to 0% and an increase in mean decrease accuracy. Statistical validation confirmed enhanced effect sizes (Cohen's  $d$ : 0.32 to 0.42) and strong significance via bootstrap resampling ( $p < 0.001$ ). Among the selected features,  $N$ ,  $N$ -dimethylglycine, creatine, and hypotaurine exhibited consistent discriminatory power, supported by both statistical metrics and biological relevance. These findings highlight the utility of univariate-based dimensionality reduction as a robust strategy for biomarker discovery in small-scale metabolomics studies.

In this study, compositional analysis and metabolomics-based

approaches were employed to identify metabolic differences between Korean and Japanese red seabream, aiming to develop biomarkers for geographical origin discrimination. The integrated analysis demonstrated the potential of metabolite-based biomarkers for distinguishing the origin of red seabream. However, the sample set used in this study was limited in terms of geographical and seasonal scope, which may restrict the generalizability of the proposed method. Since the metabolic profile of red seabream can be influenced by various aquaculture and environmental factors, such as feed, water temperature, and farming systems, continuous and long-term monitoring is necessary to validate the biomarkers under diverse conditions. Further investigations are needed to examine seasonal variations and the impact of different aquaculture systems on metabolic profiles. Given the high dimensionality of metabolomic data, larger sample sizes are crucial for confidently proposing robust metabolome-based biomarkers. Therefore, the reliability and generalizability of the biomarkers identified in this study remain to be validated.

In conclusion, this study proposes a novel strategy for authenticating seafood origins by integrating chemometrics and untargeted

metabolomics, supported by multivariate statistical analysis. This integrative approach offers a scientific foundation for establishing a reliable seafood traceability system.



## **Acknowledgment**

I would like to express my deepest gratitude to my advisor, Professor Ji-young Yang, for their invaluable guidance, unwavering support, and insightful feedback throughout my doctoral research. Their expertise and encouragement were instrumental in shaping this work.

I also extend my sincere thanks to the members of my dissertation committee, Professors Young-Mog Kim, Gun-Do Kim, Hyeoung Lee, and Kil-Bo Shim, for their constructive suggestions and critical insights, which greatly enriched this study. Special appreciation is given to my colleagues and collaborators at the Food Fermentation Lab, Department of Food Science and Technology, Pukyong National University, whose support and camaraderie made the research journey more meaningful.

I am grateful to the funding agencies, including the Ministry of Oceans and Fisheries, under the project titled Development of Discrimination Method and On-site Kit for Geographical Origin (grant number: 20200425). for supporting this research project.

Lastly, I would like to thank my family and friends for their endless

patience, encouragement, and love. Their belief in me provided the strength needed to complete this dissertation.



## References

- Abdullah M, Kornegay JN, Honcoop A, Parry TL, Balog-Alvarez CJ, O'Neal SK, Bain JR, Muehlbauer MJ, Newgard CB, Patterson C, Willis MS. (2017). Non-targeted metabolomics analysis of golden retriever muscular dystrophy-affected muscles reveals alterations in arginine and proline metabolism, and elevations in glutamic and oleic acid in vivo. *Metabolites*, 7, 38.
- Abe H. (1983). Distribution of free L-histidine and related dipeptides in the muscle of fresh-water fishes. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, 76(1), 35-39.
- Abe H, Brill RW, Hochachka PW. (1986). Metabolism of L-histidine, carnosine, and anserine in skipjack tuna. *Physiological zoology*, 59(4), 439-450.
- Acharjee A, Larkman J, Xu Y, Cardoso VR, Gkoutos GV. (2020). A random forest based biomarker discovery and power analysis framework for diagnostics research. *BMC medical genomics*, 13, 1-14.

Acquaviva A, Romero LM, Castells CB. (2016). Analysis of citrulline and metabolic related amino acids in plasma by derivatization and RPLC. Application of the extrapolative internal standard calibration method. *Microchemical Journal*, 129, 29-35.

Adam O, Beringer C, Kless T, Lemmen C, Adam A, Wiseman M, Adam P, Klimmek R, Forth W. (2003). Anti-inflammatory effects of a low arachidonic acid diet and fish oil in patients with rheumatoid arthritis. *Rheumatology international*, 23, 27-36.

Ahmad A, Imran M, Ahsan H. (2023). Biomarkers as biomedical bioindicators: Approaches and techniques for the detection, analysis, and validation of novel biomarkers of diseases. *Pharmaceutics*, 15(6), 1630.

Ahmed I, Khan M. (2006). Dietary branched-chain amino acid valine, isoleucine and leucine requirements of fingerling Indian major carp, *Cirrhinus mrigala* (Hamilton). *British Journal of Nutrition*, 96(3), 450-460.

Ali AAH. (2023). Overview of the vital roles of macro minerals in the human body. *Journal of Trace Elements and Minerals*, 4, 100076.

Aristoy MC, Toldrá F. (2004). Histidine dipeptides HPLC-based test for the detection of mammalian origin proteins in feeds for ruminants. *Meat Science*, 67(2), 211-217.

Aruoma OI, Halliwell BARRY, Hoey BM, Butler J. (1988). The antioxidant action of taurine, hypotaurine and their metabolic precursors. *Biochemical Journal*, 256(1), 251-255.

Azeredo R, Machado M, Fontinha F, Fernández-Boo S, Conceição LE, Dias J, Costas B. (2020). Dietary arginine and citrulline supplementation modulates the immune condition and inflammatory response of European seabass. *Fish & Shellfish Immunology*, 106, 451-463.

Bai K, Xu W, Zhang J, Kou T, Niu Y, Wan X, Zhang L, Wang C, Wang T. (2016). Assessment of free radical scavenging activity of dimethylglycine sodium salt and its role in providing protection against

lipopolysaccharide-induced oxidative stress in mice. *PLoS One*, 11(5), e0155393.

Bayer E, Grom E, Kaltenecker B, Uhlmann R. (1976). Separation of amino acids by high performance liquid chromatography. *Analytical Chemistry*, 48(8), 1106-1109.

BBC news. Fukushima: Japan approves releasing wastewater into ocean, 13 April 2021. Available from: <https://www.bbc.com/news/world-asia-56728068>.

Benjamin KJ, Chen Q, Eagles NJ, Huuki-Myers LA, Collado-Torres L, Stolz JM, Perteua G, Shin J, Paquola AM, Hyde T, Kleinman J, Jaffe A, Han S, Weinberger DR. (2024). Analysis of gene expression in the postmortem brain of neurotypical Black Americans reveals contributions of genetic ancestry. *Nature neuroscience*, 27(6), 1064-1074.

Bianchi M, Hallström E, Parker RW, Mifflin K, Tyedmers P, Ziegler F. (2022). Assessing seafood nutritional diversity together with climate

impacts informs more comprehensive dietary advice. *Communications Earth & Environment*, 3(1), 188.

Bidlingmeyer BA, Cohen SA, Tarvin TL. (1984). Rapid analysis of amino acids using pre-column derivatization. *Journal of Chromatography B: Biomedical Sciences and Applications*, 336(1), 93-104.

Bifarin OO. (2023). Interpretable machine learning with tree-based shapley additive explanations: Application to metabolomics datasets for binary classification. *Plos one*, 18(5), e0284315.

Bligh EG, Dyer WJ. (1959). A rapid method of total lipid extraction and purification. *Canadian journal of biochemistry and physiology*, 37(8), 911-917.

Boizard, F., Brunchault, V., Moulos, P., Breuil, B., Klein, J., Lounis, N., Caubet C, Tellier S, Bascands JL, Decramer S, Schanstra JP, Buffin-Meyer B. (2016). A capillary electrophoresis coupled to mass spectrometry pipeline for long term comparable assessment of the urinary metabolome. *Scientific reports*, 6(1), 34453.

- Boldyrev AA, Aldini G, Derave W. (2013). Physiology and pathophysiology of carnosine. *Physiological reviews*.
- Boldyrev AA. (2007). *Carnosine and Oxidative Stress in Cells and Tissue*, 4th ed.; Nova Science Publishers Inc.: New York, NY, USA.
- Bouzari A, Holstege D, Barrett DM. (2015). Mineral, fiber, and total phenolic retention in eight fruits and vegetables: a comparison of refrigerated and frozen storage. *Journal of agricultural and food chemistry*, 63(3), 951-956.
- Breiman L. (2001). Random forests. *Machine learning*, 45, 5-32.
- Brosnan ME, Brosnan JT. (2020). Histidine metabolism and function. *Journal of Nutrition*, 150, 2570S-2575S.
- Broadhurst DI, Kell DB. (2006). Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics*, 2, 171-196.
- Brosset P, Cooke SJ, Schull Q, Trenkel VM, Soudant P, Lebigre C. (2021).

Physiological biomarkers and fisheries management. *Reviews in Fish Biology and Fisheries*, 31(4), 797-819.

Buentello JA, Gatlin DM. (2000). The dietary arginine requirement of channel catfish (*Ictalurus punctatus*) is influenced by endogenous synthesis of arginine from glutamic acid. *Aquaculture*, 188(3-4), 311-321.

Calder PC, Yaqoob P. (2009). Omega-3 polyunsaturated fatty acids and human health outcomes. *Biofactors*, 35(3), 266-272.

Califf RM. (2018). Biomarker definitions and their applications. *Experimental biology and medicine*, 243(3), 213-221.

Camin F, Bontempo L, Heinrich K, Horacek M, Kelly SD, Schlicht C, Thomas F, Monahan FJ, Hoogewerff J, Rossmann A. (2007). Multi-element (H, C, N, S) stable isotope characteristics of lamb meat from different European regions. *Analytical and bioanalytical chemistry*, 389, 309-320.

- Carvalho MS, Mendonça MA, Pinho DM, Resck IS, Suarez PA. (2012).  
Chromatographic analyses of fatty acid methyl esters by HPLC-UV  
and GC-FID. *Journal of the Brazilian Chemical Society*, 23, 763-769.
- Chaouch A, Bouhlel I, Chraief I, Hammami M, El Hani A, Romdhane MS,  
El Cafsi M. (2003). Seasonal variation of polyunsaturated fatty acids  
(n-3) composition in *Diplodus annularis* from the Gulf of Tunis:  
nutritional benefits. *Journal-Societe Chimique De Tunisie*, 5(1), 55–64.
- Chiu HH, Kuo CH. (2020). Gas chromatography-mass spectrometry-  
based analytical strategies for fatty acid analysis in biological samples.  
*Journal of food and drug analysis*, 28(1), 60-73.
- Choi JH, Rhim CH, Choi YJ, Park KD, Oh SK. (1985). Comparative study  
on amino acid profiles of wild and cultured carp, and Israeli carp.  
*Korean Journal of Fisheries Aquatic Sciences*, 18(6), 545-549.
- Choi YS, Park KJ, Song JH, Yoon SP, Chung SO, An KH. (2012).  
Contents of inorganic elements in shellfish and geochemical  
characteristics of surface sediments on the west coast of Korea. *Korean*

Journal of Malacology, 28(3), 225–232.

Chongtham N, Bisht MS, Santosh O, Bajwa HK, Indira A. (2021). Mineral elements in bamboo shoots and potential role in food fortification. Journal of Food Composition and Analysis, 95, 103662.

Clark TC, Tinsley J, Sigholt T, Macqueen DJ, Martin SA. M. (2020). Arginine, ornithine and citrulline supplementation in rainbow trout: Free amino acid dynamics and gene expression responses to bacterial infection. Fish & shellfish immunology, 98, 374–390.

Colin N, Porte C, Fernandes D, Barata C, Padrós F, Carrassón M, Monray M, Cano-Rocabayera O, Sostoa A, Piña B, Maceda-Veiga A. (2016). Ecological relevance of biomarkers in monitoring studies of macro-invertebrates and fish in Mediterranean rivers. Science of the Total Environment, 540, 307-323.

Coristine LE, Robillard CM, Kerr JT, O'Connor CM, Lapointe D, Cooke SJ. (2014). A conceptual framework for the emerging discipline of conservation physiology. Conservation Physiology, 2(1), cou033.

- Cumbo F, Truglia S, Weitschek E, Blankenberg D. (2025). Feature selection with vector-symbolic architectures: a case study on microbial profiles of shotgun metagenomic samples of colorectal cancer. *Briefings in Bioinformatics*, 26(2), bbaf177.
- Cummings J, Raynaud F, Jones L, Sugar R, Dive C. (2010). Fit-for-purpose biomarker method validation for application in clinical trials of anticancer drugs. *British journal of cancer*, 103(9), 1313-1317.
- Danish M, Nizami M. (2019). Complete fatty acid analysis data of flaxseed oil using GC-FID method. *Data in brief*, 23, 103845.
- D'Ascenzo N, Antonecchia E, Angiolillo A, Bender V, Camerlenghi M, Xie Q, Di Costanzo A. (2022). Metabolomics of blood reveals age-dependent pathways in Parkinson's Disease. *Cell & Bioscience*, 12(1), 102.
- Dedon PC, DeMott MS, Elmquist CE, Prestwich EG, McFaline JL, Pang B. (2007). Challenges in developing DNA and RNA biomarkers of inflammation. *Biomarkers in medicine*, 1(2), 293-312.

Degenhardt F, Seifert S, Szymczak S. (2019). Evaluation of variable selection methods for random forests and omics data sets. *Briefings in bioinformatics*, 20(2), 492-503.

Drouin N, Ramautar R. (2021). Capillary electrophoresis-mass spectrometry for metabolomics: possibilities and perspectives. *Separation techniques applied to omics sciences: from principles to relevant applications*, 159-178.

Ecker J, Scherer M, Schmitz G, Liebisch G. (2012). A rapid GC-MS method for quantification of positional and geometric isomers of fatty acid methyl esters. *Journal of Chromatography B*, 897, 98-104.

Ekman DR., Skelton DM, Davis JM, Villeneuve DL, Cavallin JE, Schroeder A, Collette TW. (2015). Metabolite profiling of fish skin mucus: A novel approach for minimally-invasive environmental exposure monitoring and surveillance. *Environmental Science & Technology*, 49(5), 3091-3100.

Eriksson L, Trygg J, Wold S. (2008). CV-ANOVA for significance testing

of PLS and OPLS® models. *Journal of Chemometrics: A Journal of the Chemometrics Society*, 22(11-12), 594-600.

Farag MA, Abib B, Qin Z, Ze X, Ali SE. (2023). Dietary macrominerals: Updated review of their role and orchestration in human nutrition throughout the life cycle with sex differences. *Current Research in Food Science*, 6, 100450.

Farrell AP, Gallaughier PE, Routledge R. (2001). Rapid recovery of exhausted adult coho salmon after commercial capture by troll fishing. *Canadian Journal of Fisheries and Aquatic Sciences*, 58(12), 2319-2324.

Ferreira N, Henriques B, Viana T, Carvalho L, Tavares D, Pinto J, Jacinto J, Colónia J, Pereira E. (2023). Validation of a methodology to quantify macro, micro, and potentially toxic elements in food matrices. *Food Chemistry*, 404, 134669.

Fiehn O. (2016). Metabolomics by gas chromatography–mass spectrometry: Combined targeted and untargeted profiling. *Current*

protocols in molecular biology, 114(1), 30-4.

Florkowski CM. (2008). Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests. *The Clinical Biochemist Reviews*, 29 (Suppl 1), S83.

Food and Agriculture Organization of the United Nations (FAO). (2021). Food fraud – Intention, detection and management. Food Safety Technical Toolkit for Asia and the Pacific No. 5. Food and Agriculture Organization of the United Nations. <https://www.fao.org/publications>.

Frostegård A, Bååth E. (1996). The use of phospholipid fatty acid analysis to estimate bacterial and fungal biomass in soil. *Biology and Fertility of soils*, 22, 59-65.

Gan L, Jiang WD, Wu P, Liu Y, Jiang J, Li SH, Tang L, Kuang SY, Feung L, Zhou XQ. (2014). Flesh quality loss in response to dietary isoleucine deficiency and excess in fish: a link to impaired Nrf2-dependent antioxidant defense in muscle. *PLoS One* 2014;9: e115129.

Garcia A, Barbas C. (2010). Gas chromatography-mass spectrometry (GC-MS)-based metabolomics. In *Metabolic profiling: Methods and protocols* (pp. 191-204). Totowa, NJ: Humana Press.

German JB, Hammock BD, Watkins SM. (2005). Metabolomics: Building on a century of biochemistry to guide human health. *Metabolomics*, 1, 3–9.

Gil-Solsona R, Raro M, Sales C, Lacalle L, Díaz R, Ibáñez M, Beltran J, Sancho JV, Hernández FJ. (2016). Metabolomic approach for extra virgin olive oil origin discrimination making use of ultra-high performance liquid chromatography–quadrupole time-of-flight mass spectrometry. *Food Control*, 70, 350–359.

Gonzalvez A, Armenta S, De La Guardia M. (2009). Trace-element composition and stable-isotope ratio for discrimination of foods with Protected Designation of Origin. *TrAC Trends in Analytical Chemistry*, 28(11), 1295-1311.

Greene HM, Wickler SJ, Bray RE, Burrill MJ, London C. The effect of

N,N-dimethylglycine on athletic performance at altitude in horses and mules. *Pferdeheilkunde*. 1996;12: 499-501.

Grissa D, Pétéra M, Brandolini M, Napoli A, Comte B, Pujos-Guillot E. (2016). Feature selection methods for early predictive biomarker discovery using untargeted metabolomic data. *Frontiers in molecular biosciences*, 3, 30.

Guideline of Institutional animal care and use committee. (2020). Animal and Plant Quarantine Agency and Ministry of Food and Drug Safety.

Hajirezaee S, Rafieepour A, Shafiei S. (2019). A NMR-based metabonomic study on the ameliorating effects of Ginkgo biloba extract in rainbow trout, *Oncorhynchus mykiss* exposed to organophosphate pesticide, diazinon. *Aquaculture*, 513, 734450.

Hardy RW, Lee CS. (2010). Aquaculture feed and seafood quality. *Bulletin of Fisheries Research and Development Agency*, 31, 43–50.

Haury AC, Gestraud P, Vert JP. (2011). The influence of feature selection methods on accuracy, stability and interpretability of molecular signatures. PloS one, 6(12), e28210.

Hellberg R, and Morrissey MT. (2011). Advances in DNA-based techniques for the detection of seafood species substitution on the commercial market. Journal of the Association for Laboratory Automation, 16(4), 308-321.

Heo S, Lee DY, Choi HK, Lee J, Kim JH, Cho SM, Lee HJ, Auh JH. (2011). Metabolite fingerprinting of bokbunja (*Rubus coreanus* Miquel) by UPLC-qTOF-MS. Food Sci Biotechnol. 2011;20: 567-570.

Hesterberg T. (2011). Bootstrap. Wiley Interdisciplinary Reviews: Computational Statistics, 3(6), 497-526.

Hiemori-Kondo M, Shinya D, Ueta R. (2022). Development of a quantitative method for analyzing three imidazole dipeptides using high-performance liquid chromatography and its application for meat and fish. Journal of Food Composition and Analysis, 106, 104323.

Hixson SM. (2014). Fish nutrition and current issues in aquaculture: The balance in providing safe and nutritious seafood, in an environmentally sustainable manner. *Journal of Aquaculture Research and Development*, 5(3).

Höcker O, Flottmann D, Schmidt TC, Neusüß C. (2021). Non-targeted LC-MS and CE-MS for biomarker discovery in bioreactors: Influence of separation, mass spectrometry, and data processing tools. *Science of The Total Environment*, 798, 149012.

Hulbert AJ, Kelly MA, Abbott SK. (2014). Polyunsaturated fats, membrane lipids and animal longevity. *Journal of Comparative Physiology B*, 184, 149-166.

Hwang SY, Bae JH, Lim SY. (2015). Heavy metal contents and antioxidant activity and cytotoxic effect of red sea bream (*Pagrus major*): Comparative studies in domestic and imported red sea bream (*Pagrus major*). *Journal of Life Science*, 25(4), 450–455.

Jang MS, Kang YJ, Kim KW, Kim KD, Lee HY, Heo SB. (2009). Quality

characteristics of cultured olive flounder *Paralichthys olivaceus* fed with extruded pellets; I. Comparison of fatty acid and amino acid contents. *Korean Journal of Food Science and Technology*, 41(1), 42-49.

Jia TD, Kelleher SD, Hultin HO, Petillo D, Maney R, Krzynowek J. (1996). Comparison of quality loss and changes in the glutathione antioxidant system in stored mackerel and bluefish muscle. *Journal of Agricultural and Food Chemistry*, 44(5), 1195-1201.

Jin SK, Kim IS, Kim SJ, Jeong KJ, Lee JR. (2006). Fatty acid, amino acid composition and sensory traits of pork from pigs fed artificial culture medium of wild ginseng. *Food Science of Animal Resources*, 26(3), 349-355.

Johnson CH, Gonzalez FJ. (2012). Challenges and opportunities of metabolomics. *Journal of cellular physiology*, 227(8), 2975-2981.

Johnson CH, Patterson AD, Idle JR, Gonzalez FJ. (2012). Xenobiotic metabolomics: major impact on the metabolome. *Annual review of*

pharmacology and toxicology, 52(1), 37-56.

Johnstone CP, Lill A, Reina RD. (2017). Use of erythrocyte indicators of health and condition in vertebrate ecophysiology: a review and appraisal. *Biological Reviews*, 92(1), 150-168.

Jørgensen JT. (2021). Predictive biomarkers and clinical evidence. *Basic & Clinical Pharmacology & Toxicology*, 128(5), 642-648.

Kasamatsu S, Komae S, Matsukura K, Kakihana Y, Uchida K, Ihara H. (2021). 2-oxo-imidazole-containing dipeptides play a key role in the antioxidant capacity of imidazole-containing dipeptides. *Antioxidants* 10, 1434.

Katafuchi A, Kamegawa M, Goto S, Kuwahara D, Osawa Y, Shimamoto S, Ishihara S, Ohtsuka A, Ijiri D. (2024). Effects of cyclic high ambient temperature on muscle imidazole dipeptide content in broiler chickens. *Journal of Poultry Science*. 61, 2024004.

Kaushik SJ, Seiliez I. (2010). Protein and amino acid nutrition and

metabolism in fish: Current knowledge and future needs. *Aquaculture Research*, 41(3), 322-332.

Kawarazuka N. (2010). The contribution of fish intake, aquaculture, and small-scale fisheries to improving nutrition: A literature review. The World Fish Center Working Paper No. 2106, 44 p.

Khalili Tilami S, Sampels S. (2018). Nutritional value of fish: Lipids, proteins, vitamins, and minerals. *Reviews in Fisheries Science & Aquaculture*, 26(2), 243-253.

Kim HY, Shin JW, Park HO, Choi SH, Jang YM, Lee SO. (2000). Comparison of taste compounds of red sea bream, rockfish, and flounders differing in the localities and growing conditions. *Korean Journal of Food Science and Technology*, 32(3), 550-563.

Kim KW, Kim KD, Kim SK, Son MH, Jang MS, Kang YJ, Bai SC, Lee KJ. (2010). Quality characteristics of Olive flounder muscle fed with extruded pellet and raw fish-based moist pellet. *Korean Journal of Fisheries and Aquatic Sciences*, 43(5), 451–456.

- Kim RA, Chong YK. (2023). The influence of perceived risk of consuming seafood due to Fukushima nuclear disaster wastewater discharge on seafood consumption attitude and intention. *Korean Journal of Food Service Industry*, 19(6), 41-54.
- Kim S, Kim A, Ma S, Lee W, Lee S, Yoon D, Kim D, Kim S. (2019). Glutathione injection alleviates the fluctuation of metabolic response under thermal stress in olive flounder, *Paralichthys olivaceus*. *Metabolites*, 10(1), 3.
- Kim SH, Yang SO, Kim KH, Kim YS, Liu KH, Yoon YR, Lee DH, Lee CH, Hwang GS, Chung MW, Choi KH, Choi HK. (2009). Research trends, applications, and domestic research promotion strategies of metabolomics. *Korean Journal of Biotechnology and Bioengineering*, 24(2), 113–121.
- Kim YH, Kang YJ. (1991). Food habits of sand eel, *Ammodytes personatus*. *Korean Journal of Fisheries and Aquatic Sciences*, 24(2), 89-98.

Klockmann S, Reiner E, Bachmann R, Hackl T, Fischer M. (2016). Food fingerprinting: Metabolomic approaches for geographical origin discrimination of hazelnuts (*Corylus avellana*) by UPLC-QTOF-MS. *Journal of Agricultural and Food Chemistry*, 64(48), 9253–9262.

Konosu S, Watanabe K. (1976). Comparison of nitrogenous extractives of cultured and wild red sea breams. *Bulletin of the Japanese Society of Scientific Fisheries*, 42, 1263-1266.

Korean Trade Statistic Service. Import Trends by Major Countries and Types. Available online: <https://www.bandtrass.or.kr/customs/total.do?command=CUS001View&viewCode=CUS00101>. (accessed on March 13, 2023).

Kratz M, Cullen P, Kannenberg F, Kassner A, Fobker M, Abuja P M, Assmann G, Wahrburg, U. (2002). Effects of dietary fatty acids on the composition and oxidizability of low-density lipoprotein. *European journal of clinical nutrition*, 56(1), 72-81.

Kuchta K, Volk RB, Rauwald HW. (2013). Stachydrine in Leonurus

cardiaca, Leonurus japonicus, Leonotis leonurus: detection and quantification by instrumental HPTLC and  $^1\text{H}$ -qNMR analyses. *Die Pharmazie-An International Journal of Pharmaceutical Sciences*, 68(7), 534-540.

Kwon DH, Lee H, Park C, Hong SH, Hong SH, Kim GY, Cha HJ, Kim S, Kim HS, Hwang HJ, Choi YH. (2019). Glutathione induced immunostimulatory activity by promoting M1-like macrophages polarization via potential ROS scavenging capacity. *Antioxidants*, 8(9), 413.

Laberge T, Walsh PJ, McDonald MD. (2009). Effects of crowding on ornithine-urea cycle enzyme mRNA expression and activity in gulf toadfish (*Opsanus beta*). *Journal of Experimental Biology*, 212(15), 2394-2402.

Lands WE, Libelt B, Morris A, Kramer NC, Prewitt TE, Bowen P, Schmeisser D, Davidson MH, Burns JH. (1992). Maintenance of lower proportions of (n-6) eicosanoid precursors in phospholipids of human plasma in response to added dietary (n-3) fatty acids. *Biochimica et*

Biophysica Acta (BBA)-Molecular Basis of Disease, 1180(2), 147-162.

Lazar IM, Lee ED, Rockwood AL, Lee ML. (1997) Evaluation of an electrospray interface for capillary electrophoresis-time-of-flight mass spectrometry. *Journal of Chromatography A*, 791(1-2), 269-278.

Le Boucher J, Charret C, Coudray-Lucas C, Giboudeau J, Cynober L. (1997). Amino acid determination in biological fluids by automated ion-exchange chromatography: performance of Hitachi L-8500A. *Clinical chemistry*, 43(8), 1421-1428.

LeMoine CM, Walsh PJ. (2013). Ontogeny of ornithine-urea cycle gene expression in zebrafish (*Danio rerio*). *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 304(11), R991-R1000.

Lee B, Mahmud I, Marchica J, Dereziński P, Qi F, Wang F, Joshi P, Valerio F, Rivera I, Patel V, Pavlovich CP, Garrett TJ, Schroth GP, Sun Y, Perera RJ. (2020). Integrated RNA and metabolite profiling of urine liquid biopsies for prostate cancer biomarker discovery. *Scientific*

reports, 10(1), 3716.

Lee JW, Devanarayan V, Barrett YC, Weiner R, Allinson J, Fountain S, Keller S, Weinryb I, Green M, Duan L, Rogers JA, Millham R, O'Brien PJ, Sailstad J, Khan M, Ray C, Wagner JA. (2006). Fit-for-purpose method development and validation for successful biomarker measurement. *Pharmaceutical Research*, 23, 312–328.

Lee SM, Asaduzzaman AKM, Chun, BS. (2012). Characterization of lecithin isolated from anchovy (*Engraulis japonica*) residues deoiled by supercritical carbon dioxide and organic solvent extraction. *Journal of Food Science*, 77(7), C773–C778.

Lee WJ, An BK, In JJ, Han HG, Park JY, Bang IC, Shim KB. (2022). Comparison of the Food Quality of Hybrid Grouper *Epinephelus fuyscoguttatus* × *E. labnceolatus* and Hybrid Lon gtooth Grouper *E. moara* × *E. lanceolatus*. *Korean Journal of Fisheries and Aquatic Sciences*, 55(2), 129–136.

Leggatt RA, Brauner CJ, Schulte PM, Iwama GK. (2007). Effects of

acclimation and incubation temperature on the glutathione antioxidant system in killifish and RTH-149 cells. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 146(3), 317-326.

Lei Z, Huhman DV, Sumner LW. (2011). Mass spectrometry strategies in metabolomics. *Journal of Biological Chemistry*, 286(29), 25435-25442.

Li W. (2012). Volcano plots in analyzing differential expressions with mRNA microarrays. *Journal of bioinformatics and computational biology*, 10(06), 1231003.

Lim DK, Mo C, Lee JH, Long NP, Dong Z, Li J, Lim J, Kwon SW. (2018). The integration of multi-platform MS-based metabolomics and multivariate analysis for the geographical origin discrimination of *Oryza sativa* L. *Journal of Food and Drug Analysis*, 26(2), 769–777.

Linley TD, Gerring ME, Yancey PH, Drazen JC, Weinstock CL, Jamieson AJ. (2016). Fishes of the hadal zone including new species, in situ observations and depth records of Liparidae. *Deep Sea Research*

Part 1: Oceanographic Research Papers. 114:99-110.

Loktionov A. (2020). Biomarkers for detecting colorectal cancer non-invasively: DNA, RNA or proteins?. *World journal of gastrointestinal oncology*, 12(2), 124.

Loth E, Ahmad J, Chatham C, López B, Carter B, Crawley D, Oakley B, Hayward H, Cooke J, Jose' Ca'ceres AS, Bzdok D, Jones E, Charman T, Beckmann C, Bourgeron T, Toro R, Buitelaar J, Murphy D, Dumas, G. (2021). The meaning of significant mean group differences for biomarker discovery. *PLoS computational biology*, 17(11), e1009477.

Lozano Muñoz I, Díaz NF. (2020). Minerals in edible seaweed: Health benefits and food safety issues. *Critical Reviews in Food Science and Nutrition*, 62(6), 1592-1607.

Lund EK. (2013). Health benefits of seafood; Is it just the fatty acids?. *Food Chemistry*, 140(3), 413-420.

Makarov A, Scigelova M. (2010). Coupling liquid chromatography to

Orbitrap mass spectrometry. *Journal of Chromatography A*, 1217(25), 3938–3945.

Marie B, Gallet A. (2022). Fish metabolome from sub-urban lakes of the Paris area (France) and potential influence of noxious metabolites produced by cyanobacteria. *Chemosphere*, 296, 134035.

Markussen LK, Rondini EA, Johansen OS, Madsen JG, Sustarsic EG, Marcher AB, Hansen J, Gerhart-Hines Z, Granneman J, Mandrup S. (2022). Lipolysis regulates major transcriptional programs in brown adipocytes. *Nature Communications*, 13(1), 3956.

Martínez-Páramo S, Diogo P, Dinis MT, Soares F, Sarasquete C, Cabrita E. (2013). Effect of two sulfur-containing amino acids, taurine and hypotaurine in European sea bass (*Dicentrarchus labrax*) sperm cryopreservation. *Cryobiology*, 66(3), 333-338.

Martins C, Brandão T, Almeida A, Rocha SM. (2020). Enlarging knowledge on lager beer volatile metabolites using multidimensional gas chromatography. *Foods*, 9(9), 1276.

Maunder MN, Piner KR. (2015). Contemporary fisheries stock assessment: many issues still remain. *ICES Journal of Marine Science*, 72(1), 7-18.

Mehraj H, Nishimura Y, Shimasaki K. (2017). Analysis of essential macro-micro mineral content of twelve hosta taxa. *Annals of Agricultural Sciences*, 62(1), 71-74.

Min B H, Myeong JI, Kang HS. (2017). Effects of thermal and salinity stress on expression of FK506BP in the red seabream (*Pagrus major*). *Journal of Marine Life Science*, 2, 34–38.

Ministry of Food and Drug Safety (MFDS). Imported Food Safety Management. Available online: <https://radsafe.mfds.go.kr/CFQCC01F01>. Accessed on August 12, 2023.

Ministry of Food and Drug Safety (MFDS). (2023). Food code. Retrieved from <https://www.mfds.go.kr/eng>. Accessed on August 7, 2023

Ministry of Oceans and Fisheries (MOF). Statistics System of Korean Ministry of Oceans and Fisheries. Available online:

<https://www.mof.go.kr/statPortal/main/portalMain.do>. Accessed on 15 August 2021.

Ministry of Oceans and Fisheries (MOF). (2022). Special inspection of 24,391 fishery products on the occasion of Chuseok holiday. Retrieved from <https://www.mof.go.kr/article/view>. Accessed on March 13, 2023.

Mir-Marques A, Cervera ML, de la Guardia M. (2016). Mineral analysis of human diets by spectrometry methods. *TrAC Trends in Analytical Chemistry*, 82, 457-467.

Misra BB. (2020). Data normalization strategies in metabolomics: Current challenges, approaches, and tools. *European Journal of Mass Spectrometry*, 26(3), 165-174.

Moco S, Vervoort J, Bino RJ, De Vos RC, Bino R. (2007). Metabolomics technologies and metabolite identification. *TrAC Trends in Analytical Chemistry*, 26(9), 855-866.

Mohanty B, Mahanty A, Ganguly S, Sankar TV, Chakraborty K,

Rangasamy A, Sharma AP. (2014). Amino acid compositions of 27 food fishes and their importance in clinical nutrition. *Journal of Amino Acids*, 2014(1), 269797.

Montero L, Schmitz OJ, Meckelmann SW. (2020). Chemical characterization of eight herbal liqueurs by means of liquid chromatography coupled with ion mobility quadrupole time-of-flight mass spectrometry. *Journal of Chromatography A*, 1631, 461560.

Morris Jr SM. (2002). Regulation of enzymes of the urea cycle and arginine metabolism. *Annual review of nutrition*, 22(1), 87-105.

Morris SM. (2016). Arginine metabolism revisited. *The Journal of nutrition*, 146(12), 2579S-2586S.

Morrison WR., Smith LM. (1964). Preparation of fatty acid methyl esters and dimethylacetals from lipids with boron fluoride–methanol. *Journal of lipid research*, 5(4), 600-608.

Muller EI, Souza JP, Muller CC, Muller AL, Mello PA, Bizzi CA (2016).

Microwave-assisted wet digestion with H<sub>2</sub>O<sub>2</sub> at high temperature and pressure using single reaction chamber for elemental determination in milk powder by ICP-OES and ICP-MS. *Talanta*, 156, 232-238.

Muller PY, Dieterle F. (2009). Tissue-specific, non-invasive toxicity biomarkers: translation from preclinical safety assessment to clinical safety monitoring. *Expert Opinion on Drug Metabolism & Toxicology*, 5(9), 1023-1038.

Munjal Y, Tonk RK., Sharma R. (2022). Analytical techniques used in metabolomics: A review. *Syst. Rev. Pharm*, 13, 550-556.

Myong JG. (2009). Red seabream. *The Sea: Journal of the Marine Federation*, 32, 64-67.

Nakayasu ES, Gritsenko M, Pichowski PD, Gao Y, Orton DJ, Schepmoes AA, Fillmore TL, Frohnert BI, Rewers M, Krischer JP, Ansong C, Suchy-Dickey AM, Evans-Molina C, Qian WJ, Webb-Robertson BM, Metz TO. (2021). Tutorial: best practices and considerations for mass-spectrometry-based protein biomarker discovery and validation.

Nature Protocols, 16(8), 3737-3760.

National Agriculture Products Quality Management Service (NAPQ). (2021). Development of a 5-minute rapid determination tool for pork origin. Retrieved from <https://www.naqs.go.kr/bbs/boardDetail.do>. Accessed March 13, 2023.

National Agriculture Products Quality Management Service (NAPQ). (2022) Developed a method for determining the origin of chicken and chrysanthemums with the Korea Agriculture Organization. Retrieved from <https://www.naqs.go.kr/bbs/boardDetail.do>. Accessed March 13, 2023.

National Fishery Products Quality Management Service (NFQS). (2023). Country of Origin Labeling System. Retrieved from <https://www.nfqs.go.kr/hpmsg/orig/actionOriginIndicationForm.do?menuId=M0000208>. Accessed March 13, 2023.

National Institute of Fishery Science (NIFS). (2014). General status of compounded feed, General Information. Retrieved from

[https://www.nifs.go.kr/mixfeed/web/sub/02\\_01.jsp](https://www.nifs.go.kr/mixfeed/web/sub/02_01.jsp). Accessed March 13, 2023.

National Institute of Fishery Science (NIFS). (2018). Developed technology to identify the origin of eel using DNA analysis. Retrieved from <https://www.nifs.go.kr>. Accessed June 7, 2024.

National Institute of Fishery Science (NIFS) (2014), General status of compounded feed, General Information. Available online: <http://www.nifs.go.kr/fishfeed/view/ep/Commonjsp> (accessed on 13 October 2021).

National Institute of Fishery Science (NIFS). (2018). Red seabream farming, a symbol of longevity. Retrieved from <https://www.nifs.go.kr>. Accessed March 13, 2023

National Institute of Fishery Science (NIFS). (2020). Bream aquaculture technical manual. Ministry of Oceans and Fisheries.

Nicodemus KK, Malley JD. (2009). Predictor correlation impacts machine

learning algorithms: implications for genomic studies. *Bioinformatics*, 25(15), 1884-1890.

Nie B, Fang S, Jiang M, Wang L, Ni M, Zheng J, Yang Z, Li F. (2021). Anthropogenic tritium: Inventory, discharge, environmental behavior and health effects. *Renewable and Sustainable Energy Reviews*, 135, 110188.

Nielsen SS, Ward RE, Carpenter CE. (2010). Traditional methods for mineral analysis. *Food analysis*, 201-215.

Ninomiya K. (2002). Umami: a universal taste. *Food review international*, 18(1), 23-38.

Ogata H, Murai T. (1994). White muscle of masu salmon, *Oncorhynchus masou masou*, smolts possesses a strong buffering capacity due to a high level of anserine. *Fish Physiology and Biochemistry*, 13, 285-293.

Okumura S, Kurihara A, Iwamoto A, Takeuchi T. (2005). Improved survival and growth in *Octopus vulgaris* paralarvae by feeding large

type *Artemia* and Pacific sandeel, *Ammodytes personatus*: Improved survival and growth of common octopus paralarvae. *Aquaculture*, 244(1–4), 147–157.

Özogul Y, Özogul FH, Çiçek E, Polat A, Kuley E. (2009). Fat content and fatty acid compositions of 34 marine water fish species from the Mediterranean Sea. *International Journal of Food Sciences and Nutrition*, 60(6), 464–475.

Oruna-Concha MJ, Methven L, Blumenthal H, Young C, Mottram DS. (2007). Differences in glutamic acid and 5'-ribonucleotide contents between flesh and pulp of tomatoes and the relationship with umami taste. *Journal of Agriculture and Food Chemistry*, 55(14), 5776-5780.

Ou FS, Michiels S, Shyr Y, Adjei AA, Oberg AL. (2021). Biomarker discovery and validation: statistical considerations. *Journal of Thoracic Oncology*, 16(4), 537-545.

Park SJ, Kim KY, Yim SB, Park MJ, Kim BS, Yu YJ, Jeong YH. (2006). Fatty acid composition and mineral content of marketed mackerels.

Journal of the East Asian Society of Dietary Life, 16(6), 670–676.

Pecoraro C, Babbucci M, Franch R, Rico C, Papetti C, Chassot E, Bodin N, Cariani A, Bargelloni L, Tinti F. (2018). The population genomics of yellowfin tuna (*Thunnus albacares*) at a global geographic scale challenges current stock delineation. *Scientific Reports*, 8(1), 13890.

Peiretti PG, Medana C, Visentin S, Giancotti V, Zunino V, & Meineri G. (2011). Determination of carnosine, anserine, homo carnosine, pentosidine and thiobarbituric acid reactive substances contents in meat from different animal species. *Food Chemistry*, 126, 1939–1947.

Quehenberger O, Armando AM, Dennis EA. (2011). High sensitivity quantitative lipidomics analysis of fatty acids in biological samples by gas chromatography–mass spectrometry. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids*, 1811(11), 648-656.

Rabiei N, Soltanian AR, Farhadian M, Bahreini F. (2023). the performance evaluation of the random forest algorithm for a gene selection in identifying genes associated with resectable pancreatic cancer in

microarray dataset: A retrospective study. *Cell Journal (Yakhteh)*, 25(5), 347

Rahimi R, Hajirezaee S, Pordanjani HR. (2020). A <sup>1</sup>HNMR-based molecular study of anesthesia in fish. *Aquaculture*, 520, 734995.

Rahimnejad S, Lee KJ. (2013) Dietary valine requirement of juvenile red sea bream *Pagrus major*. *Aquaculture*. 416, 212-218.

Rharrabti Y, Elhani S, Martos-Núñez V, García del Moral LF. (2001). Protein and lysine content, grain yield, and other technological traits in durum wheat under Mediterranean conditions. *Journal of Agricultural and Food Chemistry*, 49(8), 3802-3807.

Rudneva II., Skuratovskaya EN, Kuzminova NS, Kovyreshina TB. (2010). Age composition and antioxidant enzyme activities in blood of Black Sea teleosts. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology*, 151(2), 229-239.

Rutherford SM, Gilani GS. (2009). Amino acid analysis. *Current*

protocol protein science. 58(1), 11-9.

Roques S, Deborde C, Richard N, Skiba-Cassy S, Moing A, Fauconneau B. (2020). Metabolomics and fish nutrition: a review in the context of sustainable feed development. *Reviews in Aquaculture*, 12(1), 261-282.

Ryuji Takasaki (2024). Significance of Biochemical Markers and their Vital Role in Clinical Research and Drug Development. *Clinical & Medical Biochemistry*, 10(1).

Saito-Shida S, Hamasaka T, Nemoto S, Akiyama H. (2018). Multiresidue determination of pesticides in tea by liquid chromatography-high-resolution mass spectrometry: Comparison between Orbitrap and time-of-flight mass analyzers. *Food Chemistry*, 256, 140-148.

Santacruz L, Hurtado DX, Doohan R, Thomas OP, Puyana M, Tello E. (2020). Metabolomic study of soft corals from the Colombian Caribbean: PSYCHE and <sup>1</sup>H-NMR comparative analysis. *Scientific Reports*, 10(1), 5417.

Sawayama E, Nakao H, Kobayashi W, Minami T, Takagi M. (2019). Using microsatellite DNA markers, identifying and quantifying farmed red sea bream escapees from a large aquaculture area in Japan. *Aquatic Living Resources*, 32, 26.

Sawilowsky SS. (2009). New effect size rules of thumb. *Journal of modern applied statistical methods*, 8(2), 26.

Sayed I, Parvez S, Pandey S, Bin-Hafeez B, Haque R, Raisuddin, S. (2003). Oxidative stress biomarkers of exposure to deltamethrin in freshwater fish, *Channa punctatus* Bloch. *Ecotoxicology and environmental safety*, 56(2), 295-301.

Selamat J, Rozani N, Murugesu S. (2021). Application of the metabolomics approach in food authentication. *Molecules*, 26(24), 7565.

Seppänen-Laakso T, Laakso I, Hiltunen R. (2002). Analysis of fatty acids by gas chromatography, and its relevance to research on health and nutrition. *Analytica Chimica Acta*, 465(1-2), 39-62.

Shin GM, Ahn YS, Shin DM, Kim HS, Kim HJ, Yoon MS, Kim JS. (2008).

Comparison of muscle color, taste, and nutrition components between red seabreams cultured by feeding and starving. *Journal of the Korean Society of Food Science and Nutrition*, 37(9), 1142-1147.

Shin J, Yang J, Cha E, Kim H, Lee Y, Kim S, Yang J. (2021). Analyzing

the Metabolomic Profile of Yellowtail (*Seriola quinquerdiata*) by Capillary Electrophoresis–Time of Flight Mass Spectrometry to Determine Geographical Origin. *Metabolites*, 11(11), 793.

Shin YK, Kim YD, Kim WJ. (2018). Survival and physiological responses

of red sea bream *Pagrus major* with decreasing sea water temperature. *Korean Journal of Ichthyology*, 30(3), 131-136.

Shulgina LV, Davletshina TA, Pavlovsky AM, Solodova EA, Pavel KG.

(2019). Composition of lipids and fatty acids in muscle tissue of chub mackerel *Scomber japonicus*. *Izvestiya Tinro*, 196(1), 193–203.

Singla D, Sangha MK, Singh M, Pathak M, Bala M. (2023). Variation of

mineral composition in different fruit parts of bitter gourd (*Momordica*

*charantia* L.). Biological Trace Element Research, 201(10), 4961-4971.

Slinger G, Stevelink R, van Diessen E, Braun KP, Otte WM. (2023). The importance of discriminative power rather than significance when evaluating potential clinical biomarkers in epilepsy research. *Epileptic Disorders*, 25(3), 285-296.

Song IS, Jin DH, Choi SJ, Lee SG. (2004). Polymorphism analysis of the ND-4 gene for the origin determination of olive flounder, *Paralichthys olivaceus*. *Korean J Life Sci.* 16: 627-635.

Song JH. (2003). A study on development process of fish aqua culture on Japan—case by seabream aquaculture. *Journal of Fisheries Business Administration*, 34, 75–90.

Staub A, Schappler J, Rudaz S, Veuthey JL. (2009). CE-TOF/MS: fundamental concepts, instrumental considerations and applications. *Electrophoresis*, 30(10), 1610-1623.

Stoffel W, Chu F, Ahrens EH. (1959). Analysis of long-chain fatty acids

- by gas-liquid chromatography. *Analytical Chemistry*, 31(2), 307-308.
- Strimbu K, Tavel JA. (2010). What are biomarkers?. *Current Opinion in HIV and AIDS*, 5(6), 463-466.
- Suyama M, Hirano T, Okada N, Shibuya T. (1977). Quality wild and cultured Ayu-I on the proximate composition, free amino acids and related compounds. *Bulletin of the Japanese Society Scientific Fishery*, 43(5), 535-540.
- Takeuchi T, Shiina Y, Watanabe T. (1991). Suitable protein and lipid levels in diet for fingerlings of red sea bream *Pagrus major*. *Nippon Suisan Gakkishi*, 57(2), 293-299.
- Tauchi H, Imamura H, Inoue, Komatsu K, Tachibana A. (2011). Assessment of biological effect of tritiated water by using hypersensitive system. *Fusion Science and Technology*, 60(3), 1173-1178.
- Taşbozan O, Gökçe MA. (2017). Fatty acids in fish. *Fatty acids*, 1, 143–

159.

- Terjesen BF, Chadwick TD, Verreth JA, Rønnestad I, Wright PA. (2001). Pathways for urea production during early life of an air-breathing teleost, the African catfish *Clarias gariepinus* Burchell. *Journal of Experimental Biology*, 204(12), 2155-2165.
- Thirumdas R, Janve M, Siliveru K, Kothakota A. (2019). Determination of food quality using atomic emission spectroscopy. In *Evaluation technologies for food quality* (pp. 175-192). Woodhead Publishing.
- Týčová A, Ledvina V, Klepárník K. (2017). Recent advances in CE-MS coupling: Instrumentation, methodology, and applications. *Electrophoresis*, 38(1), 115-134.
- Tomonaga S, Kaji Y, Tachibana T, Denbow DM, Furuse M. (2005). Oral administration of  $\beta$ -alanine modifies carnosine concentrations in the muscles and brains of chickens. *Journal of Animal Science*, 76(3), 249–254.

Triba MN, Le Moyec L, Amathieu R, Goossens C, Bouchemal N, Nahon P, Rutledge DN, Savarin, P. (2015). PLS/OPLS models in metabolomics: the impact of permutation of dataset rows on the K-fold cross-validation quality parameters. *Molecular BioSystems*, 11(1), 13-19.

Vaidyanathan S, Dunn WB, Harrigan GG, Kell DB. (2004). Metabolomics by numbers: acquiring and understanding global metabolite data. *TRENDS in Biotechnology*, 22(5), 245-252.

Van den Berg RA, Hoefsloot HC, Westerhuis JA., Smilde AK, Van der Werf MJ. (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC genomics*, 7, 1-15.

Villasante A, Ramírez C, Figueroa Villalobos E, Pereira WA, Powell MS, Gatlin III DM, Dantngnan P, Romero J. (2023). Creatine in sustainable fish aquaculture. *Reviews in Fisheries Science & Aquaculture*, 31(3), 420-451.

Wang CY, Li YR, Pan C, Chen J, Jiang W, Li WN, Zhang XL, Liao Z, Yan XJ. (2021). Quantitative analysis of carnosine, anserine, and homocarnosine in skeletal muscle of aquatic species from east China sea. *Biochemistry and Biophysics Reports*, 25, 100880.

Wang Q, Xu Z, Ai Q. (2021). Arginine metabolism and its functions in growth, nutrient utilization, and immune nutrition of fish. *Animal Nutrition*, 7(3), 716-727.

Wheelock ÅM, Wheelock CE. (2013). Trials and tribulations of 'omics data analysis: assessing quality of SIMCA-based multivariate models using examples from pulmonary medicine. *Molecular BioSystems*, 9(11), 2589-2596.

White JA, Hart RJ, Fry JC. (1986). An evaluation of the Waters Pico-Tag system for the amino-acid analysis of food materials. *Journal of Automatic Chemistry*, 8(4), 170-177.

Whitmire M, Ammerman J, De Lisio P, Killmer J, Kyle D, Mainstone E, Porter L, Zhang T. (2011). LC-MS/MS bioanalysis method

development, validation, and sample analysis: points to consider when conducting nonclinical and clinical studies in accordance with current regulatory guidances. *Journal of Analytical and Bioanalytical Techniques*, 4(2).

Wishart DS. (2008). Metabolomics: applications to food science and nutrition research. *Trends in food science & technology*, 19(9), 482-493.

Wright PA, Felskie A, Anderson PM. (1995). Induction of ornithine-urea cycle enzymes and nitrogen metabolism and excretion in rainbow trout (*Oncorhynchus mykiss*) during early life stages. *Journal of Experimental Biology*, 198(1), 127-135.

Wright PA, Land MD. (1998). Urea production and transport in teleost fishes. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 119(1), 47-54.

Wuertz S, Reiser S. (2023). Creatine: A valuable supplement in aquafeeds?. *Reviews in Aquaculture*, 15(1), 292-304.

- Xia J, Broadhurst DI, Wilson M, Wishart DS. (2013). Translational biomarker discovery in clinical metabolomics: An introductory tutorial. *Metabolomics*, 9, 280–299.
- Xiccato G, Trocino A, Tulli F, Tibaldi E. (2004). Prediction of chemical composition and origin identification of European sea bass (*Dicentrarchus labrax* L.) by near infrared reflectance spectroscopy (NIRS). *Food Chemistry*, 86(2), 275-281.
- Yamada T, Kamiya M, Higuchi M. (2020). Gas chromatography–mass spectrometry-based metabolomic analysis of Wagyu and holstein beef. *Metabolites*, 10(3), 95.
- Yamaguchi S, Ninomiya K. What is umami? (1998). *Food Review International*, 14(2-3), 123-138.
- Yamaguchi S, Ninomiya K. (2000). Umami and food palatability. *Journal of nutrition*, 130(4), 921S-926S.
- Yancey PH, Gerringer ME, Drazen JC, Rowden AA, Jamieson A. (2014).

Marine fish may be biochemically constrained from inhabiting the deepest ocean depths. *Proceedings of the National Academy Sciences*, 111(12), 4461-4465.

Yang J, Shin J, Cha E, Kim H, Lee Y, Kim S, Yang, J. (2022). Analysis of metabolites of red seabream (*Pagrus major*) from different geographical origins by capillary electrophoresis time-of-flight mass spectrometry. *Plos One*, 17(7), e0270280.

Yang J, Shin J, Kim H, Sim Y, Cha E, & Yang J. (2023). Analysis of Metabolite Differences between South Korean and Chinese Yellow goosefish (*Lophius litulon*) using Capillary Electrophoresis Time-of-Flight Mass Spectrometry. *Journal of Chromatography B*, 123863.

Yoda H, Takahashi M, Sengoku T. (2015). Developments in the Synthesis of 3-Acyltetramic Acid Natural Products. *Studies in Natural Products Chemistry*, 46, 99-131.

Yokoyama M, Nakazoe JI. (1998). Effect of oral administration of L-cystine on hypotaurine level in rainbow trout. *Fisheries science*, 64(1),

144-147.

Zehra S, Khan MA. (2014). Dietary valine requirement of fingerling *Catla catla*. *Journal of Applied Aquaculture*, 26(3), 232-251.

Zengin H. (2021). The effects of feeding and starvation on antioxidant defence, fatty acid composition and lipid peroxidation in reared *Oncorhynchus mykiss* fry. *Scientific Reports*, 11(1), 16716.

Zenitani H, Onishi Y, Kobayashi S, Fujiwara T. (2009). Spawning season, spawning grounds, and egg production of red sea bream in Hiuchi-nada, Seto Inland Sea. *Fisheries Science*, 75, 55-62.

Zhang H, Zhang X, Zhao X, Xu J, Lin C, Jing P, Hu L, Zhao S, Wang X, Li, B. (2019). Discrimination of dried sea cucumber (*Apostichopus japonicus*) products from different geographical origins by sequential windowed acquisition of all theoretical fragment ion mass spectra (SWATH-MS)-based proteomic analysis and chemometrics. *Food Chemistry*, 274, 592–602.

Zhao G, Zhao W, Han L, Ding J, Chang Y. (2020). Metabolomics analysis of sea cucumber (*Apostichopus japonicus*) in different geographical origins using UPLC–Q-TOF/MS. *Food Chemistry*, 333, 127453.

Zhao J, Zhao Y, Hu C, Zhao C, Zhang J, Li L, Zeng J, Peng X, Lu X, Xu G. (2016). Metabolic profiling with gas chromatography–mass spectrometry and capillary electrophoresis–mass spectrometry reveals the carbon–nitrogen status of tobacco leaves across different planting areas. *Journal of Proteome Research*, 15(2), 468-476.

Zhou J, Zhong L. (2022). Applications of liquid chromatography-mass spectrometry based metabolomics in predictive and personalized medicine. *Frontiers in Molecular Biosciences*, 9, 1049016.

Zhou WH, Limbu SM, Luo Y, Li RX, Ren J, Qiao F, Zhang ML, Du ZY. (2023). Dietary acetate promotes growth and nutrients deposition in Nile tilapia (*Oreochromis niloticus*) through increasing acetyl-CoA-triggered energy production. *Aquaculture*, 575, 739750.

Zlatanov S, Laskaridis K. (2007). Seasonal variation in the fatty acid

composition of three Mediterranean fish—sardine (*Sardina pilchardus*), anchovy (*Engraulis encrasicolus*) and picarel (*Spicara smaris*). Food Chemistry, 103(3), 725–728.

Zou S, Wang X, Liu P, Ke C, Xu S. (2019). Arginine metabolism and deprivation in cancer therapy. Biomedicine & Pharmacotherapy, 118, 109210.

Zuk M, Stoehr AM. (2002). Immune defense and host life history. The American Naturalist, 160(S4), S9-S22.

