

## 基于 PCA-SVM 的仿刺参产地溯源在线系统研究

吴 鹏<sup>1,2</sup>, 李 颖<sup>1,2\*</sup>, 刘 瑀<sup>2,3</sup>, 付金宇<sup>1,2</sup>, 李亚芳<sup>1,2</sup>, 冉明衢<sup>1,2</sup>, 赵新达<sup>3</sup>

1. 大连海事大学航海学院, 辽宁 大连 116026
2. 大连海事大学环境信息研究所, 辽宁 大连 116026
3. 大连海事大学环境科学与工程学院, 辽宁 大连 116026

**摘 要** 仿刺参是具有极高经济价值的水产资源,是海水养殖产业的重要组成部分,研发出一种灵活、稳定、高效的仿刺参产地溯源方法对于水产养殖产业具有极强的现实意义。仿刺参主要有三种养殖方式,分别是底播增殖、圈养养殖和筏式养殖。不同产地采用不同的养殖方式,仿刺参的营养价值、药用价值和经济价值都存在着明显差异。不同产地初级生产者的构成不同,作为初级生产者的不同藻类与浮游生物体内的脂肪酸特征也各不相同,通过食物链的传递,不同产地的仿刺参具有了不同的脂肪酸特征。气相色谱指纹图谱法是一种快速准确地食品产地溯源技术,碳稳定同位素质谱法不仅可以鉴别产地还可以区分出食品的营养价值。采集9个最具代表性产地的仿刺参样品,先利用 Folch 法对样品进行总脂提取,再通过气相色谱仪测定出各种脂肪酸的种类及其相对含量;最后使用稳定同位素质谱仪测定出每种脂肪酸各自的碳稳定同位素组成数据。使用单因素方差分析法对脂肪酸相对含量和脂肪酸碳稳定同位素组成数据进行显著性检验,各筛选出17种脂肪酸数据作为两个模型的输入。主成分分析(PCA)法可以降低数据的维度,聚合不同种脂肪酸数据的溯源特征,提高产地溯源的精度。支持向量机(SVM)是一种以结构化风险最小为目标的分类识别算法,具有优秀的泛化能力。研究结果表明,不同产地仿刺参的脂肪酸相对含量和脂肪酸碳稳定同位素组成数据存在明显差异。通过主成分变换后,脂肪酸数据的聚类特征更加明显,运用随机交叉验证法确定前6个主成分作为两个支持向量机分类器的输入。采用基于遗传交叉因子改进的粒子群优化算法(GPSO),以粒子不同K值各100次交叉验证的平均准确率作为其适应度,寻找支持向量机分类器模型的最优参数组合。最终计算得到脂肪酸相对含量产地溯源模型的最优参数组合为 $\sigma=6.247\ 599$ 和 $C=14.313\ 042$ ,平均准确率为79.49%;脂肪酸碳稳定同位素组成产地溯源模型的最优参数组合为 $\sigma=7.626\ 194$ 和 $C=2.193\ 410$ ,平均准确率为98.33%。对比交叉验证的结果,脂肪酸碳稳定同位素组成产地溯源模型具有更高的准确率和更强的泛化性能。在两个模型的识别结果不一致时,采用脂肪酸碳稳定同位素组成模型的识别结果。将实验室检测与互联网技术进行整合,构建了仿刺参产地溯源在线系统。实现了“互联网+产地溯源”的一体化溯源模式,为进一步开展食品产地溯源研究提供了科学依据和技术支撑。

**关键词** 仿刺参;产地溯源;脂肪酸;主成分分析;支持向量机;互联网+

**中图分类号:** O657.7 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2019)05-1604-06

## 引 言

海参是所有棘皮动物门海参纲(Holothuroidea)生物物种的统称,目前已知的1400多种海参全部栖息在海洋之中,其中的60多种可被人食用<sup>[1-2]</sup>。仿刺参(*Apostichopus japonicus*)是营养价值与经济价值最高的一类海参,是我国重要的

水产养殖资源<sup>[3]</sup>。2016年中国仿刺参的年产量达到204359t,养殖区主要分布在辽宁,山东,福建等东部沿海地区。由于不同产地气候不同,仿刺参的生长周期存在明显差异,福建等南方地区养殖不到一年时间即可长为成参,但辽宁、山东等北方地区则需要养殖3~5年才可以出栏上市。为了最大程度制止不法商贩以次充好、牟取暴利,有效地维护消费者的利益,保证水产养殖产业的平稳高速发展。精度高、检

收稿日期:2018-04-08,修订日期:2018-09-17

基金项目:国家科技支撑计划项目(2015BAD17B05),国家海洋公益类专项(201305002)资助

作者简介:吴 鹏,1994年生,大连海事大学航海学院硕士研究生 e-mail:18840866641@163.com

\* 通讯联系人 e-mail:yldmu@126.com

验周期短、灵活高效地仿刺参产地溯源在线系统是迫切需要的。

脂肪酸具有可被生物体直接吸收、可沿食物链传递和基本结构在代谢过程中不被破坏的特性<sup>[4]</sup>。通过长期不断的进食积累,不同产地仿刺参体内的脂肪酸相对含量和脂肪酸碳稳定同位素组成呈现明显差异,因此脂肪酸特征可以作为仿刺参产地溯源的有效依据<sup>[5-7]</sup>。

近些年来,国内外研究人员分析了不同产地仿刺参在无机元素、稳定同位素<sup>[8]</sup>、近红外光谱<sup>[9]</sup>和多环芳烃<sup>[10]</sup>等方面的特性,并应用聚类分析、判别分析和主成分分析等进行了初步的产地溯源研究。但以往的研究存在着采样产地少、未对主要产地进一步细分、没有建立具体产地溯源模型、模型精度不高和模型参数未优化等问题。

本研究提出了根据脂肪酸相对含量和脂肪酸碳稳定同位素组成特征追溯仿刺参产地的新方法,建立了主成分分析(PCA)与支持向量机(SVM)的联合模型,并在此基础上构建了仿刺参产地溯源在线系统,实现了“互联网+产地溯源”的一体化溯源模式,对实现食品产地溯源的现代化、规模化与标准化具有重要的参考价值及借鉴意义。

## 1 实验部分

### 1.1 样品

有效仿刺参样品共 109 个:其中脂肪酸相对含量样品 54 个,脂肪酸碳稳定同位素组成样品 55 个。样品包括 9 个产地,除普兰店皮口样品为 6 个与 7 个外,其他产地样品均为 6 个,其中长海县、獐子岛、烟台担子岛、烟台莱州和烟台牟平的样品为底播增殖;普兰店皮口、瓦房店和威海乳山的样品为圈养养殖;福建霞浦的样品为筏式养殖。

仿刺参采集时间为 2015 年 11 月,除福建霞浦样品为当年的成参外,其他样品均为参龄 2 年的成参,体长 15~19 cm,体重 100~130 g。捕捞后立即用水将样品洗净,去除内脏、沙石及石灰环,体壁用超纯水洗净,冷冻干燥 48 h,磨制成粉末状,过 80 目网筛并干燥。参照 Folch<sup>[11]</sup>法对仿刺参进行总脂提取,然后用 1 mL 1% 硫酸-甲醇溶液水浴甲酯化 20 mg 总脂,冷却后振荡,待静置后取上清液保存。

### 1.2 仪器

Flash EA 1112 型元素分析仪,Delta V Advantage 型稳定同位素质谱仪,Trace GC Ultra 型气相色谱仪和 ISQ 型气质联用仪,均为美国赛默飞世尔科技公司生产。

### 1.3 数据测定

按照 1.1 方法提取到的仿刺参甲酯化脂肪酸溶液 1  $\mu\text{L}$ ,进行气相色谱分离和质谱分析。色谱分离条件为:无分流方式进样,进样口温度 280  $^{\circ}\text{C}$ ;初始加热至 70  $^{\circ}\text{C}$ 并保持 1 min,以 8  $^{\circ}\text{C}\cdot\text{min}^{-1}$ 的速度加热至 170  $^{\circ}\text{C}$ 并保持 6 min,再以 4  $^{\circ}\text{C}\cdot\text{min}^{-1}$ 的速度加热至 280  $^{\circ}\text{C}$ 并保持 10 min,最后以 1.2 mL $\cdot\text{min}^{-1}$ 的恒定流速充入纯度 $\geq 99.999\%$ 的 He 作为载气。质谱分析条件为:传输线温度 250  $^{\circ}\text{C}$ ;离子源温度 230  $^{\circ}\text{C}$ ;通过能量为 70 eV 的 EI 电子进行电离。

测量脂肪酸相对含量时,直接将气相色谱分离后的甲酯

化脂肪酸,通过气质联用仪进行 GC-MS 实验并与标准质谱图进行对比分析,确定脂肪酸的种类及其相对含量;测量脂肪酸碳稳定同位素组成时,取 1/10 气相色谱分离后的甲酯化脂肪酸,通过气质联用仪进行 GC-MS 实验并与标准质谱图进行对比分析,确定脂肪酸的种类,同时将剩余的 9/10 甲酯化脂肪酸,通过稳定同位素质谱仪测定相应脂肪酸的碳稳定同位素组成。

### 1.4 方法

仿刺参产地溯源研究的整体思路与具体方法如图 1 所示。

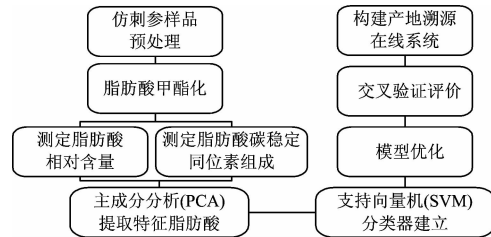


图 1 整体思路流程图

Fig. 1 Flow chart of principle idea

#### 1.4.1 利用主成分分析(PCA)提取特征

主成分分析(PCA)是统计学中最常用的一种降维方法。仿刺参体内含有多种脂肪酸,借助主成分分析法提取出最具溯源特性的脂肪酸种类,并利用其降维特性,提高样品的采样密度聚合溯源特性,消除噪声与实验误差的干扰,提高分类器精度。

#### 1.4.2 支持向量机(SVM)分类器

支持向量机(SVM)是机器学习领域中最常用且效果最好的一种分类识别算法。相对于神经网络等分类算法,支持向量机的优化目标是结构化风险最小,而不是经验风险最小,使其具有了优秀的泛化能力和鲁棒特性,避免了过拟合、神经网络结构选择和局部最优等问题。

支持向量机的分类性能取决于核函数的选择,采用交叉验证的方法,测试线性核函数、多项式核函数与高斯径向基核函数的分类性能。由于经过主成分分析后的数据具有明显的聚类特性,因此各核函数分类器的准确率相差不大,但高斯径向基核函数经过模型校正与参数优化后,具有更小的海明距离与泛化误差。因此,选择高斯径向基作为支持向量机分类器的核函数。

#### 1.4.3 交叉验证优化及评价

交叉验证也称为循环估计,是统计学中最常用的一种预测模型性能的评价方式。通过交叉验证评价可以挖掘出有限数据中尽可能多的有效信息,并根据验证结果不断对模型进行优化,一定程度上可以减小与预防模型过拟合的发生,提高模型的精度、稳定性与泛化性能。

但对于不同的数据划分模型会产生截然不同的验证结果。因此,采用蒙特卡洛随机采样的方法,对产地溯源模型进行不同  $K$  值各 100 次随机划分的交叉验证评价,并计算出不同  $K$  值模型准确率的均值和标准差<sup>[12]</sup>。

## 2 结果与讨论

### 2.1 仿刺参脂肪酸测定结果

54 个脂肪酸相对含量样品, 测定出 28 种脂肪酸; 55 个脂肪酸碳稳定同位素组成样品, 测定出 20 种脂肪酸。对各种脂肪酸数据进行显著性检验, 采用单因素方差分析法, 进行置信水平为 95% 的单总体图基 (Tukey) 检验, 各筛选出 17 种显著性差异  $< 0.001$  的脂肪酸, 并根据实验结果, 分别建立各产地仿刺参脂肪酸数据库。图 2 为脂肪酸样品的气相色谱图。

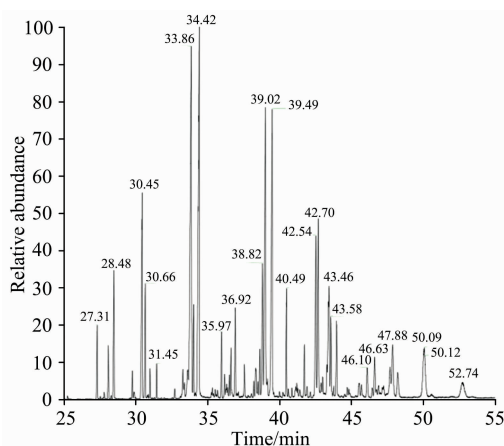


图 2 脂肪酸样品气相色谱图

Fig. 2 Chromatogram of fatty acids samples

### 2.2 主成分分析 (PCA) 提取结果

图 3 和图 4 分别是脂肪酸相对含量和脂肪酸碳稳定同位素组成前 3 个主成分数据的散点图, 从图中可以清晰地看出提取出的不同产地仿刺参脂肪酸特征呈现出明显的聚类特征。

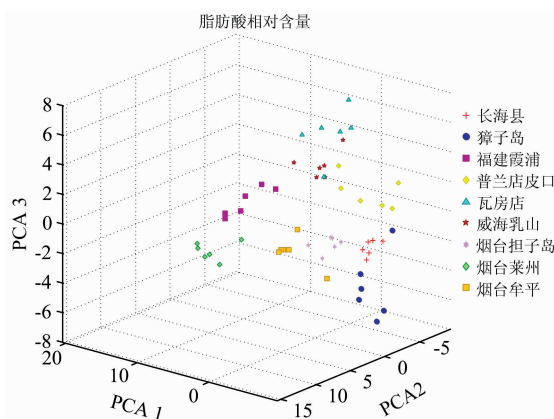


图 3 脂肪酸相对含量主成分分析结果

Fig. 3 PCA results of fatty acids relative content

初选贡献率大于 1 的主成分作为产地溯源模型的输入变量, 脂肪酸相对含量数据的前 6 个主成分满足条件; 脂肪酸碳稳定同位素组成数据的前 8 个主成分满足条件。针对不同的主成分个数在 (0, 1 000] 区间范围内, 进行随机参数运算,

计算得到参数的最优区间为 (0, 100]。

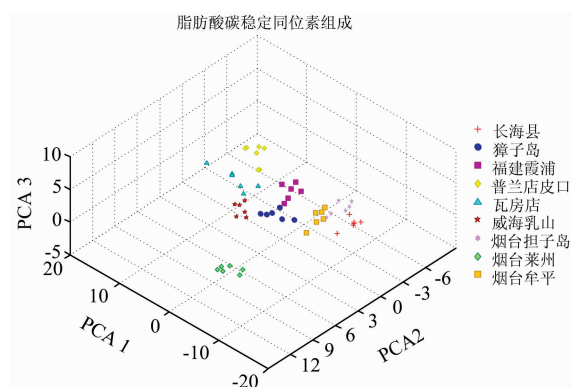


图 4 脂肪酸碳稳定同位素组成主成分分析结果

Fig. 4 PCA results of fatty acids carbon stable isotope compositions

表 1 和表 2 分别为脂肪酸相对含量和脂肪酸碳稳定同位素组成在保证训练集上每类至少都包含一个数据的前提下, 依次对前  $N$  个主成分进行最优区间上 100 次随机参数的模型运算, 计算得到的不同  $K$  值各 100 次随机采样交叉验证的平均准确率。因此, 选取前 6 个主成分作为脂肪酸相对含量和脂肪酸碳稳定同位素组成模型的输入变量, 前 6 个主成分的累计贡献率分别为 97.815% 和 92.344%。

表 1 脂肪酸相对含量模型平均准确率

Table 1 Average accuracy of fatty acid relative content model

主成分个数	$K=3/\%$	$K=6/\%$	$K=9/\%$	$K=18/\%$	$K=54/\%$
$N=2$	43.624	47.080	48.222	49.603	51.400
$N=3$	65.308	70.907	72.852	75.367	76.690
$N=4$	63.891	67.611	68.465	69.200	70.020
$N=5$	68.276	72.517	73.518	74.640	74.930
$N=6$	67.880	72.953	74.497	76.450	77.630

表 2 脂肪酸碳稳定同位素组成模型平均准确率

Table 2 Average accuracy of fatty acids carbon stable isotope compositions model

主成分个数	$K=3/\%$	$K=6/\%$	$K=9/\%$	$K=18/\%$	$K=55/\%$
$N=2$	62.987	68.391	70.127	72.810	74.790
$N=3$	82.066	85.731	86.080	86.233	86.450
$N=4$	91.168	95.069	95.685	96.657	96.810
$N=5$	90.960	93.662	94.212	94.250	94.440
$N=6$	92.912	96.124	96.833	97.863	98.220
$N=7$	81.440	84.572	85.032	85.777	86.110
$N=8$	83.199	97.514	88.625	89.930	90.240

### 2.3 支持向量机 (SVM) 分类结果

利用 Accord.NET 框架下的机器学习程序集设计并优化支持向量机分类器, 并运用基于遗传交叉因子改进的粒子群优化算法 (GPSO) 建立最优产地溯源模型<sup>[13]</sup>。在最优参数区间内, 随机设置高斯径向基核函数参数  $\sigma$  和复杂度参数  $C$  的

初始值,进行初始种群规模为 50,遗传进化代数为 100,自我认知因子  $c_1$  为 1.49 618,社会认知因子  $c_2$  为 1.49 618,权重因子  $w$  为 0.752 9 的分类器参数寻优。

在每一次遗传进化中,先计算出粒子不同  $K$  值各 100 次交叉验证的平均准确率并作为其适应度,然后更新种群最优和个体最优,最后根据适应度大小对粒子进行降序排序,较优秀的前一半粒子直接进入下一次进化,较差的后一半粒子依次随机与前一半中的粒子进行交叉遗传,这样既可以提高模型收敛的速度也避免了陷入局部最优的问题。

最终计算得到脂肪酸相对含量模型的最优参数组合为  $\sigma = 6.247\ 599$  和  $C = 14.313\ 042$ ,平均准确率为 79.49%;脂肪酸碳稳定同位素组成模型的最优参数组合为  $\sigma = 7.626\ 194$  和  $C = 2.193\ 410$ ,平均准确率为 98.33%。

每一类随机筛选出 1/3 数据作为预测样本,剩下 2/3 数据作为训练样本,使用建立的最优产地溯源模型进行预测,识别结果分别如图 5 和图 6 所示。图中纵坐标为 0 时,代表分类器未给出最终分类结果。从图中可以清楚的看出产地溯源模型具有良好的准确率和极低的错误率,其中脂肪酸碳稳定同位素组成模型具有更高的产地溯源能力。

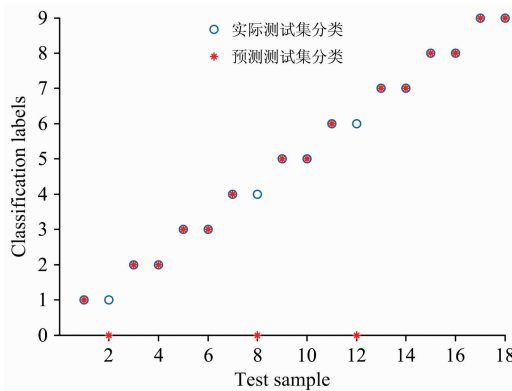


图 5 脂肪酸相对含量模型识别结果

Fig. 5 Recognition results of fatty acid relative content model

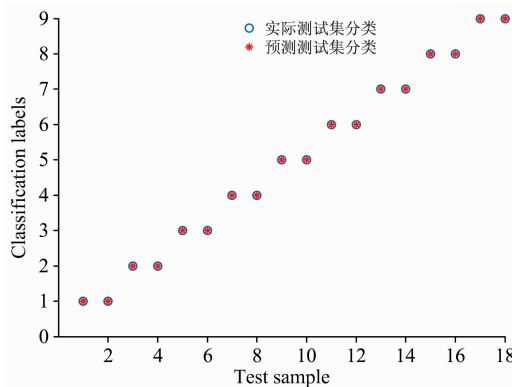


图 6 脂肪酸碳稳定同位素组成模型识别结果

Fig. 6 Recognition results of fatty acids carbon stable isotope compositions model

2.4 交叉验证结果

表 3 和表 4 分别是脂肪酸相对含量模型和脂肪酸碳稳定

同位素组成模型不同  $K$  值 100 次随机交叉验证的平均准确率与标准差。对比两个模型不同  $K$  值时的准确率,随着  $K$  值的减小脂肪酸碳稳定同位素组成模型的准确率没有明显的下降,更加稳定并且具有更好的泛化特性。因此,选择脂肪酸碳稳定同位素组成模型作为仿刺参产地溯源的首选模型。在两个模型的识别结果不一致时,采用脂肪酸碳稳定同位素组成模型的识别结果。

表 3 脂肪酸相对含量模型交叉验证结果

Table 3 Results of cross validation of fatty acid relative content model

K 取值	平均准确率/%	标准差
3	70.44	0.098 407
6	74.89	0.132 049
9	76.50	0.178 099
18	80.00	0.210 819
54	86.00	0.346 987

表 4 脂肪酸碳稳定同位素组成模型交叉验证结果

Table 4 Results of cross validation of fatty acids carbon stable isotope compositions model

K 取值	平均准确率/%	标准差
3	94.44	0.072 436
6	97.78	0.058 794
9	99.00	0.039 581
18	100.00	0
55	100.00	0

3 在线系统构建

为了使消费者与市场监管部门能有效地追溯仿刺参产地信息,基于微软的 .net framework 4.5.2 框架,利用 HTML, CSS 与 JavaScript 等语言搭建系统的前台,并利用 C# 与 SQL server 等语言架构网站的后台及数据库。采用了开放式、模块化、集成化的设计思想,构建了体系架构完善、使用灵活高效地仿刺参产地溯源在线系统。在线系统主要由前



图 7 仿刺参产地溯源在线系统界面

Fig. 7 Web page of apostichopus japonicus origin identification online system

台用户化界面模块、脂肪酸数据输入模块、后台数据分析模块与产地信息显示模块等构成。仿刺参产地溯源在线系统的网站界面如图 7 所示。

## 4 结 论

讨论了脂肪酸数据进行仿刺参产地溯源的可行性,通过主成分分析法降低脂肪酸相对含量和脂肪酸碳稳定同位素组成数据的维度,聚合产地溯源特性,并利用以高斯径向基为

核函数的支持向量机构建产地溯源模型。根据交叉验证的结果,运用基于遗传交叉因子改进的粒子群优化算法对模型参数不断优化,最终构建了平均准确率分别为 79.49% 和 98.33% 的仿刺参产地溯源模型。

实验证明:脂肪酸碳稳定同位素组成模型的准确率高、泛化误差小、稳定性高,是一种实用高效的仿刺参产地溯源方法。借助互联网技术整合传统的实验室检测方式,构建了仿刺参产地溯源在线系统,充分利用“互联网+”的机遇,推动了我国食品产地溯源的现代化、规模化与标准化。

## References

- [ 1 ] CHEN Zi-hao, WANG Mao-jian, ZHANG Jian, et al(陈子豪, 王茂剑, 张 健, 等). Food Science(食品科学), 2017, 38(12): 196.
- [ 2 ] ZHAO Xin-tao(赵欣涛). Agricultural Development & Equipments(农业开发与装备), 2016, 12: 38.
- [ 3 ] YANG Ning, GUO Zhong-shuai, WANG Zheng-li(杨 宁, 郭中帅, 王正丽). Fisheries Science(水产科学), 2016, 35(5): 498.
- [ 4 ] Alfaro A C. Estuarine Coastal and Shelf Science, 2008, 79(4): 718.
- [ 5 ] Zhang X F, Liu Y, Li Y, et al. Food Chemistry, 2017, 218: 269.
- [ 6 ] Liu Y, Zhang X F, LI Y, et al. Journal of the Science of Food and Agriculture, 2017, 97(14): 4912.
- [ 7 ] Petursdottir H, Gislason A, Falk-Petersen S, et al. Deep Sea Research Part II: Topical Studies in Oceanography, 2008, 55(1): 83.
- [ 8 ] LIU Yu, ZHANG Xu-feng, LI Ying, et al(刘 瑀, 张旭峰, 李 颖, 等). Marine Environmental Science(海洋环境科学), 2017, 36(1): 37.
- [ 9 ] TAO Lin, WU Zhong-chen, ZHANG Peng-yan, et al(陶 琳, 武中臣, 张鹏彦, 等). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2011, 27(5): 364.
- [10] XU Lin-bo(徐林波). Ocean University of China(中国海洋大学), 2012.
- [11] Folch J, Lees M, Sloane-Stanley G H. Journal of Biological Chemistry, 1957, 226(1): 497.
- [12] LU Peng-fei, FAN Ya, ZHOU Lin-hua, et al(卢鹏飞, 范 雅, 周林华, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2017, 37(12): 3828.
- [13] ZHANG Jin, DING Sheng, LI Bo(张 进, 丁 胜, 李 波). Journal of Computer Applications(计算机应用), 2016, 36(5): 1330.

# Study on Apostichopus Japonicus Origin Identification Online System Based on PCA-SVM

WU Peng<sup>1,2</sup>, LI Ying<sup>1,2\*</sup>, LIU Yu<sup>2,3</sup>, FU Jin-yu<sup>1,2</sup>, LI Ya-fang<sup>1,2</sup>, RAN Ming-qu<sup>1,2</sup>, ZHAO Xin-da<sup>3</sup>

1. Navigation College, Dalian Maritime University, Dalian 116026, China

2. Environmental Information Institute, Dalian Maritime University, Dalian 116026, China

3. College of Environmental Science and Engineering, Dalian Maritime University, Dalian 116026, China

**Abstract** *Apostichopus japonicus*, an important part of the mariculture, is a fishery resources with extremely high economic value. Therefore, it is of great practical significance for the mariculture to study a flexible, stable and efficient method to identify the origin information of *apostichopus japonicas*. There are three main aquaculture methods for the *apostichopus japonicas*, including bottom sowing culture, captive culture and raft culture. Different aquaculture methods are used in *apostichopus japonicus* of different producing areas, and there also exist great differences in the nutritional value, medicinal value and economical value from different producing areas. The compositions of primary producers vary from one provenance to another, and the characteristics of fatty acids in different algae and plankton as primary producers are also different. Through the transmission of the food chain, *apostichopus japonicus* from different producing areas have different fatty acid characteristics. Gas chromatographic fingerprint is a fast and accurate traceability technology for food origin. Carbon stable isotope ratio mass spectrometry can not only identify origin but also distinguish the nutritional value of food. Samples of *apostichopus japonicus* were collected from nine representative producing areas, and total lipid data were extracted by using the Folch method. Then determined the data of fatty acid kinds and relative content through gas chromatography. Finally, stable isotope ratio mass spectrometer was used to determine the data of fatty acids carbon stable isotope compositions. One-factor analysis of variance (ANOVA) was used to test the

significance of the data of fatty acid relative content and fatty acids carbon stable isotope compositions, and then selected 17 kinds of fatty acid data as inputs for the two models. The principal component analysis(PCA) method can reduce the dimensions of the data, and aggregate the origin characteristics of different fatty acids to improve the accuracy of origin identification model. Support vector machine (SVM) is a classification algorithm that aims to minimize structural risk and has good ability of generalization. The results indicated that there were significant differences in the fatty acids relative content and fatty acids carbon stable isotope compositions data of *Apostichopus japonicus* from different producing areas. After the principal component analysis, the clustering characteristics of the fatty acids data were more obvious. With the cross-validation method, the first six principal components were determined as inputs of the two support vector machine classifiers. Applied the particle swarm optimization improved based on genetic crossover factor(GPSO), and the average accuracy of 100 cross-validation results of different  $K$  values of the particle was used as the fitness to find optimal parameter combinations of the SVM classifier model. Finally, the optimal parameter combinations of fatty acids relative content model were  $\sigma=6.247\ 599$  and  $C=14.313\ 042$ , and average accuracy was 79.49%; the optimal parameter combinations of fatty acids carbon stable isotope compositions model were  $\sigma=7.626\ 194$  and  $C=2.193\ 410$ , and average accuracy was 98.33%. Compared with the results of cross-validation, the fatty acids carbon stable isotope compositions origin identification model has higher accuracy and stronger ability of generalization. When origin identification results of two models were inconsistent, the results of fatty acids carbon stable isotope compositions model were used. The laboratory detection was integrated with Internet technology to build an *Apostichopus japonicus* origin identification online system. The Integrated model of "Internet+origin identification" has achieved to provide a scientific basis and technical support for the further studies on origin identification of food.

**Keywords** *Apostichopus japonicus*; Origin identification; Fatty acids; Principal component analysis; Support vector machine; Internet+

(Received Apr. 8, 2018; accepted Sep. 17, 2018)

\* Corresponding author

---

## 敬告读者——《光谱学与光谱分析》已全文上网

从 2008 年第 7 期开始在《光谱学与光谱分析》网站([www.gpxygpx.com](http://www.gpxygpx.com))“在线期刊”栏内发布《光谱学与光谱分析》期刊全文,读者可方便地免费下载摘要和 PDF 全文,欢迎浏览、检索本刊当期的全部内容;并陆续刊出自 2004 年以后出版的各期摘要和 PDF 全文内容。2009 年起《光谱学与光谱分析》每期出版日期改为每月 1 日。

《光谱学与光谱分析》期刊社