基于机器学习和元素指纹的牛肉产地溯源

向晓蒙[1,2]，赵超敏[2,3*]，陈念念[2,3]，张润何[2,3]，Diego Cristos[4]，王粮子[2,3]，曾静[2,3]，
吴池莹[2,3]，蔡文瑄[2,3]，伊雄海[2,3*]，邓晓军[1]，刘冰[1*]
（1.上海体育大学兴奋剂检测上海研究院，上海 200438）（2.上海海关动植物与食品检验检疫技术中心，上海 201210）
（3.上海出入境食品与饲料安全专业服务平台，上海 201210）
（4.食品技术研究所-农业产业研究中心，可持续食物体系科技研究所，布宜诺斯艾利斯赫林汉姆 1686）

摘要：为了研究中国不同产地牛肉在元素上的差异并构建可靠的产地溯源方法，本研究采用电感耦合等离子体质谱（ICP-MS）与电感耦合等离子体发射光谱（ICP-OES）技术对来自山东、云南、宁夏、新疆、江苏、甘肃和四川七个产地的牛肉样品中 52 种元素含量进行了系统性定量分析。利用元素含量信息构建原始数据集，并结合多元统计与机器学习方法，包括多重方差分析（MANOVA）、随机森林（RF）、线性判别分析（LDA）以及余弦相似度（Cosine Similarity）算法，联合建立牛肉产地判别模型。通过特征重要性评估与方差分析，筛选出 27 个具有显著地区差异的关键元素（Sc、Co、K、Se、Rb、V、Ba、As、Sn、Pb、Na、Cu、Mg、Ag、Al、P、Ni、Tl、Fe、Ti、Zn、Mn、Cd、Ca、Cs、Sb、Mo），并据此构建模型核心特征空间。通过引入 12 例外部验证样品进行产地鉴别，结果显示判别准确率达到 91.67%，验证了所建模型的可行性与稳健性。研究表明，不同产地牛肉在元素组成上存在显著差异，元素指纹信息可作为产地溯源的重要依据，该成果为提升牛肉产地真实性认证和保障肉类产品质量安全提供了理论基础。

关键词：牛肉；产地溯源；多元素；机器学习
文章编号：1673-9078(2026)10-332-340 DOI: 10.13982/j.mfst.1673-9078.2026.10.1167

Beef Origin Traceability Based on Machine Learning and Elemental Fingerprints

XIANG Xiaomeng[1,2], ZHAO Chaomin[2,3*], CHEN Niannian[2,3], ZHANG Runhe[2,3], DIEGO Cristos[4], WANG Liangzi[2,3], ZENG Jing[2,3], WU Chiying[2,3], CAI Wenxuan[2,3], YI Xionghai[2,3*], DENG Xiaojun[1], LIU Bing[1*]
(1.Shanghai Institute of Doping Analyses, Shanghai University of Sport, Shanghai 200438, China)
(2.Technical Center for Animal, Plant and Food Inspection and Quarantine, Shanghai Customs, Shanghai 201210, China)
(3.Shanghai Entry and Exit Food and Feed Safety Professional Service Platform, Shanghai 201210, China)
(4.Food Technology Institute-Agroindustry Research Center, Institute of Science and Technology of Sustainable Food Systems, Hurlingham 1686, Buenos Aires, Argentina)

Abstract: To investigate the elemental differences in beef from different regions of China and construct a reliable origin traceability method, the contents of 52 elements in beef samples from seven production regions (Shandong, Yunnan, Ningxia, Xinjiang, Jiangsu, Gansu, and Sichuan) were systematically quantified using inductively coupled plasma mass spectrometry (ICP-MS) and inductively coupled plasma optical emission spectrometry (ICP-OES). An original dataset was constructed based on the elemental contents, and a discriminant model for beef origin was developed by applying multivariate statistical and machine learning methods, including multivariate analysis of variance (MANOVA), random forest (RF), linear discriminant analysis (LDA),

收稿日期：2025-08-15；修回日期：2025-09-24；接受日期：2025-09-25
基金项目：国家重点研发计划项目（2023YFE0106900）
作者简介：向晓蒙（1998-），男，硕士研究生，研究方向：食品掺假溯源；E-mai：psnamd9600hk@gmail.com
通讯作者：赵超敏（1979-），女，博士，高级工程师，研究方向：食品掺假溯源；E-mail：chaomin829@163.com；共同通讯作者：伊雄海（1979-），男，博士，正高级工程师，研究方向：食品安全检测，E-mail：xhyi1118@sina.com；刘冰（1989-），女，博士，副教授，研究方向：兴奋剂检测，E-mail：liubing2019@sus.edu.cn

and cosine similarity algorithms. Through feature importance evaluation and variance analysis, 27 key elements (Sc, Co, K, Se, Rb, V, Ba, As, Sn, Pb, Na, Cu, Mg, Ag, Al, P, Ni, Tl, Fe, Ti, Zn, Mn, Cd, Ca, Cs, Sb, and Mo) with significant regional differences were identified and used to construct the core feature space of the model. Subsequently, 12 external validation samples were introduced for origin identification, and the results showed a discrimination accuracy of 91.67%, confirming the feasibility and robustness of the established model. The findings demonstrated that significant differences exist in the elemental composition of beef from different regions, and elemental fingerprints can serve as reliable indicators for origin traceability. This achievement provides a theoretical basis for improving beef origin authentication and ensuring the quality and safety of meat products.

Key words: beef; origin traceability; multi-element; machine learning

引文格式：

向晓蒙,赵超敏,陈念念,等.基于机器学习和元素指纹的牛肉产地溯源[J].现代食品科技,2026,42(10):332-340.

XIANG Xiaomeng, ZHAO Chaomin, CHEN Niannian, et al. Beef origin traceability based on machine learning and elemental fingerprints [J]. Modern Food Science and Technology, 2026, 42(10): 332-340.

牛肉是重要的高价值畜产品，因其蛋白质含量高，风味独特、富含微量元素而广受消费者青睐[1,2]。随着人民收入与消费水平的提高，人们对牛肉的安全与品质也提出更高的要求，牛肉品质因其产地来源不同也具有较大差异性，产地成为影响牛肉品质的主导因素，也是牛肉价格差异大的主要原因[3-5]。近年来市场上常常出现牛肉以次充好的现象，食品安全问题层出不穷，消费者因高价购买到低价的产地假冒牛肉而遭受经济损失[6]。因此亟需一种新的技术手段来准确分析牛肉的产地信息，为牛肉的真实性与安全提供保障。中国幅员辽阔，肉牛品种多样，产地分布全国各地，主要产区有新疆维吾尔自治区、山东省、四川省等地，因其自然条件优秀，养殖体系完整等因素成为我国牛肉产品输出大省[7]。然而，不同牛肉产地的地质背景、水源、饲料及肉牛养殖方式的差异，导致牛肉中元素指纹特征呈显著差异并具有地理特征。元素指纹被认为是食品溯源的重要技术之一[8-10]，在植物产品中应用主要有水果、咖啡、茶叶等[11-13]，动物产品中应用主要有鱼、猪、羊等[14-17]。在牛肉溯源领域，Fernandes等[18]通过 9 种元素构建的牛肉溯源 MLP 模型完成了对巴西、澳大利亚、阿根廷、乌拉圭和巴拉圭 5 个国家牛肉的产地区分。Zhang 等[19]使用 ICP-MS 测定牛骨 50 种元素含量，通过 HCA 与 LDA 等方法判别牛骨的产地，正确率达到 90% 以上。随着计算机技术不断发展，人工智能出现在人们的视野中，而实现人工智能的核心驱动力，即机器学习也在农产品溯源中的应用变得普遍起来[20,21]。Zong 等[22]使用近红外光谱，结合反向传播神经网络构建藏牦牛产地判别模型，准确率达 95%。Qi 等[23]利用不同地区猪肉元素指纹，结合前馈神经网络算法构建产地判别模型，整体正确率达为 95.7%，足以区分 7 个不同地区猪肉的地理产地。An 等[24]使用 ICP-MS 与 ICP-OES 测定了不同鸡肉样品中的 60 种元素构建 OPLS-DA（正交偏最小二乘判别）模型，获得 28 个高 VIP 值元素作为影响鸡肉溯源的关键影响因子。Qie 等[25]通过对羊肉中的元素指纹、稳定同位素与脂肪酸进行测定，构建 PCA、HCA 与 PCA-LDA 产地判别模型，鉴别率 100%。

本研究通过电感耦合等离子体质谱与电感耦合等离子体发射光谱测定山东、云南、宁夏、新疆、江苏、甘肃、四川 7 个省份牛肉样品中 Li、Na、Mg 等 52 种元素含量，创新采用 Python 构建了随机森林（Random Forest, RF）、线性判别分析（Linear Discriminant Analysis, LDA）与余弦相似度（Cosine Similarity）组合产地预测模型，本研究实现了省份之间牛肉产地的分类预测，拓展了机器学习在牛肉产地溯源中的应用，为牛肉产地真实性与安全性提供了技术保障。

1 材料与方法

1.1 实验样品

牛肉样品 107 个均购自各省市，其中山东样品 10 个、云南样品 14 个、宁夏样品 21 个、新疆样品 21 个、江苏样品 20 个、甘肃样品 14 个、四川样品 7 个，样品存于 −20 ℃冰箱中备用。

1.2 仪器与试剂

5100 电感耦合等离子体发射光谱仪（ICP-OES），美国 Agilent 公司；7900 电感耦合等离子体质谱仪（ICP-MS），美国 Agilent 公司；微波消解仪，上海屹尧公司；控温电加热器，上海屹尧公司；分析天平，瑞士 Mettler Toledo 公司；硝酸，上海安谱公司；52 种元素标准品（1 000 mg·L^{-1}），上海安谱公司。

1.3 标准曲线

52 种元素标准曲线配制参考国家标准 GB 5009.268-2016《食品安全国家标准 食品中多元素的测定》[26]，参考附录 ICP-MS 与 ICP-OES 方法的标准溶液系列质量浓度绘制标准曲线。

1.4 样品前处理

参考 GB 5009.268-2016 标准。冷冻牛肉样品自然解冻，取适量牛肉样品经绞肉机均匀切割为肉泥，准确称取 1.0 g（精确到 0.001 g）牛肉样品于消解管中，加入 5 mL 硝酸（≥65 wt.%），加盖放置 1 h 或过夜，旋紧管盖，将其放入已预热的微波消解仪中进行消解。消解程序条件：初始温度 120 ℃，保持 5 min；升至 150 ℃保持 10 min；然后升至 190 ℃，保持 20 min。消解后，取出溶液并冷却至室温，缓慢打开管盖排气，用少量水冲洗内盖，将消解管放在电热器上于 100 ℃加热 30 min，用水定容至 50 mL，使用滤棉漏斗过滤消解液中的杂质，滤液经 ICP-MS 与 ICP-OES 测定元素含量。

1.5 仪器工作条件

电感耦合等离子体发射光谱仪的基本参数：重复项：3 次；泵速：12 r·min^{-1}；提升延迟：20 s；冲洗时间：15 s；读取时间：3 s；射频功率：1.2 kW；稳定时间：10 s；观察模式：SVDV（径向和轴向双向观察）；观察高度 8 mm；等离子体气流量：12.0 L·min^{-1}；雾化气流量：0.70 L·min^{-1}；辅助气流量：1.0 L·min^{-1}。所有气体如载气和等离子体气体等均为纯度 99.999% 的氩气（Ar）。

电感耦合等离子体质谱仪的基本参数：采集方式：跳峰；每峰测定点数：3 个点；每次重复读取时间：3 次；采样深度：9 mm；雾化室温度：2 ℃；样品提升速率：0.5 r·s^{-1}；载气流量：0.82 L·min^{-1}；氦气流量：4 mL·min^{-1}；稀释气体流量：0.2 L· min^{-1}；射频功率 1 600 W；等离子体气流量：16 L· min^{-1}。

1.6 数据分析

Microsoft Office Excel 对数据进行统计与初步分析。使用 Python 3.12.6 构建分析程序，Python 的 statsmodels 0.14.5 包构建多重方差检测程序，matplotlib 3.9.2 包和 sklearn 1.5.2 包实现分类模型的训练与验证，seaborn 0.13.2 包进行图形绘制。

2 结果与讨论

2.1 多重方差分析

因气候、地形、水土与饲养方式等原因，不同产地的牛肉产品中元素含量存在明显差异，表 1 为不同省份牛肉样品中 52 种元素（Li、Na、Mg、Al、P、K、Ca、Sc、Ti、V、Cr、Mn、Fe、Co、Ni、Cu、Zn、As、Se、Rb、Sr、Y、Mo、Ru、Rh、Ag、Cd、Sn、Sb、Te、Cs、Ba、La、Ce、Pr、Nd、Sm、Eu、Gd、Tb、Dy、Ho、Er、Tm、Yb、Lu、Ir、Pt、Au、Hg、Tl、Pb）平均含量定量检测结果与标准差，其中 Na、K、P、Mg、Ca、Fe、Zn 的单位为 mg/100 g，其余元素为 μg·kg^{-1}。所有元素测定值经多重方差分析（MANOVA），结果 p 值于表 1 中尾列显示。通常 p 值小于 0.05 说明该元素属性与分类显著相关，分析结果显示共有 18 种元素的 p 值与产地分类呈现显著相关的关系，这些元素在表 1 中以星号标记，分别为 Li、Na、Mg、P、K、Ca、Sc、V、Fe、Co、Zn、As、Rb、Rh、Sn、Ba、Tl、Pb。这 18 种元素具有作为牛肉产地溯源主要关键参数的潜力。

各产地牛肉元素含量不同明显反映了土壤和地质条件的差异性。分析显示，多种宏量元素和痕量元素在不同产地间存在显著差异。如图 1a 所示，新疆和云南牛肉中镁（Mg）含量较高（均值分别为 21.35 和 21.26 mg/100 g），这可能与两地土壤和水源中丰富的镁元素有关，镁元素通过牧草和饮水在牛体内富集；图 1b 中，山东牛肉中的锂（Li）含量（152.47 μg·kg^{-1}）远超其他地区，这或许暗示了其沿海地域独特的土壤或水源背景，例如某些特定类型的矿物质或地下卤水影响。水源中的矿化成分直接影响牛体内矿物质沉积；图 1c 表明，其甘肃牛肉中钠（Na）的含量（94.89 mg/100 g）表现最为突出，这很可能是因为当地部分牧区的饮用水矿化度高，含盐量较大；云南牛肉中钾（K）（353.96 mg/100 g）和钙（Ca）（14.28 mg/100 g）含量均显著高于其他地区（图 1d 与图 1f），这清晰地指向云南省复杂的地质背景，特别是广泛分布的富钾页岩和石灰岩喀斯特地貌，高含量的钾和钙通过风化进入土壤和水源，并最终传递到牛体内；此外，图 1f 反映了江苏牛肉中铷（Rb）含量最高（442.18 μg·kg^{-1}），铷的地球化学行为与钾相似，推测其高含量可能与江苏平原地区特定的土壤类型或农业活动有关。

表 1 不同省份牛肉中矿物元素含量测定结果

Table 1 Mineral element contents in beef across provinces

元素	山东	云南	宁夏	新疆	江苏	甘肃	四川	*P*-value
Li*	152.47±262.12	19.04±28.60	7.22±11.07	29.09±34.62	17.11±24.12	15.68±23.43	24.85±35.86	0.007
Na*	43.26±14.42	53.09±12.56	64.19±14.79	63.56±16.59	61.72±24.36	94.89±79.86	38.60±10.20	0.015
Mg*	16.67±4.12	21.26±3.23	20.14±2.29	21.35±1.98	17.75±4.71	17.25±3.81	16.27±4.09	0.001
Al	671.88±792.53	984.45±1173.09	703.20±1055.26	435.16±761.24	517.56±340.26	844.24±1256.83	806.12±1201.82	0.211
P*	142.34±34.39	171.75±24.47	164.87±20.51	174.48±16.26	143.58±38.09	144.63±35.19	140.40±36.09	0.007
K*	273.20±68.73	353.96±63.23	339.40±44.30	329.61±26.87	290.94±82.14	279.52±71.39	248.62±75.23	0.001
Ca*	8.54±3.01	14.28±5.38	12.84±4.25	12.01±3.68	13.74±6.91	12.72±3.75	7.77±2.66	0.036
Sc*	9.20±3.66	76.88±61.01	116.26±111.19	429.89±438.28	88.27±70.93	121.63±109.81	99.97±236.92	<0.000
Ti	325.84±158.59	449.71±236.12	425.48±189.83	545.08±254.38	431.86±245.07	401.64±206.97	273.60±259.71	0.182
V*	1.34±1.20	1.48±1.28	1.21±1.10	187.10±177.78	1.47±1.22	0.99±1.36	33.42±83.65	<0.000
Cr	23.22±8.28	18.84±10.76	19.62±12.24	31.72±32.44	17.29±13.58	24.19±13.33	24.81±11.32	0.345
Mn	151.55±44.44	169.55±63.96	165.39±49.35	181.64±43.97	153.46±30.53	144.08±61.90	173.00±44.50	0.435
Fe*	1.93±0.59	2.02±0.94	2.08±0.67	2.13±0.91	1.62±0.54	1.68±0.46	3.12±2.35	0.046
Co*	1.47±0.85	3.61±2.55	3.17±1.66	601.59±588.87	2.72±1.53	1.82±1.23	118.83±306.07	<0.000
Ni	2.16±2.75	12.92±13.07	16.25±26.72	15.28±13.91	19.09±24.35	10.51±8.26	9.46±14.10	0.435
Cu	458.13±252.32	647.77±87.74	654.29±138.39	623.25±112.51	611.23±261.92	607.74±223.76	671.31±207.88	0.299
Zn*	3.50±1.72	3.92±0.95	4.71±1.16	5.25±1.13	4.34±1.55	4.12±1.09	4.89±0.89	0.019
As*	10.78±4.16	12.19±6.76	12.84±5.95	12.75±7.60	6.47±3.54	11.10±8.70	7.46±3.16	0.046
Se	72.16±21.04	105.23±55.66	141.00±95.19	102.43±30.46	119.37±74.83	89.87±59.41	92.98±36.65	0.182
Rb*	398.86±244.92	387.00±166.61	353.33±125.95	233.06±125.59	442.18±278.25	234.09±111.52	346.46±236.51	0.022
Sr	281.93±100.54	489.80±223.83	430.80±168.22	501.88±270.77	436.77±183.92	429.93±135.69	254.65±132.66	0.054
Y	2.71±2.65	1.63±2.25	2.19±3.39	1.83±5.41	1.73±2.80	2.96±2.90	1.82±1.47	0.933
Mo	0.00±0.00	0.73±1.57	2.40±3.55	4.24±4.40	3.63±3.51	3.83±6.40	0.94±2.49	0.054
Ru	0.00±0.00	0.00±0.00	0.00±0.00	0.24±0.80	0.00±0.00	0.00±0.00	0.00±0.00	0.412
Rh*	0.13±0.13	0.01±0.04	0.05±0.12	0.00±0.01	0.02±0.05	0.04±0.08	0.01±0.02	0.019
Ag	71.64±81.87	88.48±73.65	73.18±70.35	29.41±41.67	64.97±38.69	60.46±42.57	73.96±37.63	0.182
Cd	0.47±0.37	1.38±1.26	0.96±0.37	0.97±0.39	1.09±0.42	0.93±0.47	0.83±0.57	0.062
Sn*	1.65±0.62	3.62±2.17	2.33±1.17	2.89±1.13	2.52±1.62	1.80±0.90	3.68±2.74	0.019
Sb	0.90±0.69	2.23±1.24	1.90±0.93	1.61±1.36	2.19±1.02	2.24±3.23	6.94±15.57	0.194
Te	0.21±0.66	0.00±0.00	0.00±0.00	0.03±0.13	0.00±0.00	0.00±0.00	0.50±1.32	0.14
Cs	10.34±10.66	26.32±38.21	21.45±16.23	13.84±14.23	25.41±23.79	12.83±11.23	24.82±37.98	0.418
Ba*	92.90±29.16	135.66±52.02	117.84±33.24	157.47±67.98	132.81±48.79	138.43±63.14	78.07±39.88	0.019
La	10.94±12.18	4.78±8.75	7.76±17.77	4.38±11.11	14.20±52.60	5.85±10.66	3.26±3.29	0.931
Ce	5.71±4.80	2.98±4.42	2.49±3.50	6.16±15.70	3.66±6.81	1.86±2.61	3.64±3.96	0.82
Pr	0.55±0.49	0.49±0.73	0.30±0.41	0.20±0.29	0.60±1.33	0.31±0.44	3.54±8.23	0.069
Nd	2.60±1.80	1.61±2.60	1.52±2.18	8.21±19.34	0.76±1.37	3.38±5.46	1.53±1.39	0.28
Sm	0.25±0.29	0.39±0.68	0.32±0.62	0.13±0.23	0.37±1.17	0.29±0.46	0.39±0.31	0.931
Eu	1.02±0.65	0.71±0.68	0.81±0.70	0.74±0.65	1.04±1.16	1.01±0.90	0.69±0.30	0.866
Gd	0.98±0.78	0.58±0.66	0.62±0.80	0.54±0.43	0.70±1.23	0.61±0.87	1.03±0.82	0.83
Tb	0.22±0.23	0.14±0.21	0.21±0.25	0.11±0.19	0.08±0.18	0.17±0.24	0.15±0.13	0.476
Dy	0.67±0.53	0.31±0.46	0.40±0.77	0.52±1.21	0.40±0.66	0.22±0.33	0.51±0.53	0.896

续表1

元素	山东	云南	宁夏	新疆	江苏	甘肃	四川	*P*-value
Ho	0.07±0.08	0.04±0.09	0.04±0.09	0.04±0.07	0.12±0.24	0.08±0.20	0.08±0.16	0.698
Er	0.46±0.44	0.58±0.60	0.49±0.72	0.15±0.22	0.26±0.46	0.20±0.31	0.24±0.23	0.182
Tm	0.03±0.05	0.01±0.02	0.02±0.04	0.00±0.02	0.04±0.10	0.00±0.01	0.00±0.00	0.436
Yb	0.16±0.20	0.14±0.26	0.16±0.35	0.14±0.21	0.14±0.26	0.05±0.11	0.19±0.20	0.931
Lu	0.02±0.05	0.04±0.07	0.02±0.05	0.01±0.03	0.02±0.05	0.02±0.05	0.06±0.07	0.418
Ir	0.30±0.93	0.05±0.13	0.07±0.16	0.00±0.00	0.04±0.15	0.08±0.22	0.00±0.00	0.436
Pt	0.00±0.00	0.11±0.19	0.53±1.57	0.59±1.02	0.32±0.70	0.63±1.07	0.38±1.00	0.708
Au	1.41±3.07	4.43±10.00	1.50±2.48	3.35±5.72	1.46±1.96	0.48±1.40	1.40±2.39	0.418
Hg	0.00±0.00	1.65±3.89	3.01±6.89	0.88±3.37	4.08±6.75	0.93±2.37	0.00±0.00	0.28
Tl*	0.01±0.03	0.06±0.13	0.02±0.06	10.51±13.05	0.28±1.09	0.02±0.05	2.00±4.65	<0.000
Pb*	16.81±7.83	27.93±22.30	24.14±14.26	81.68±95.28	34.19±44.77	26.23±33.28	47.78±50.22	0.019

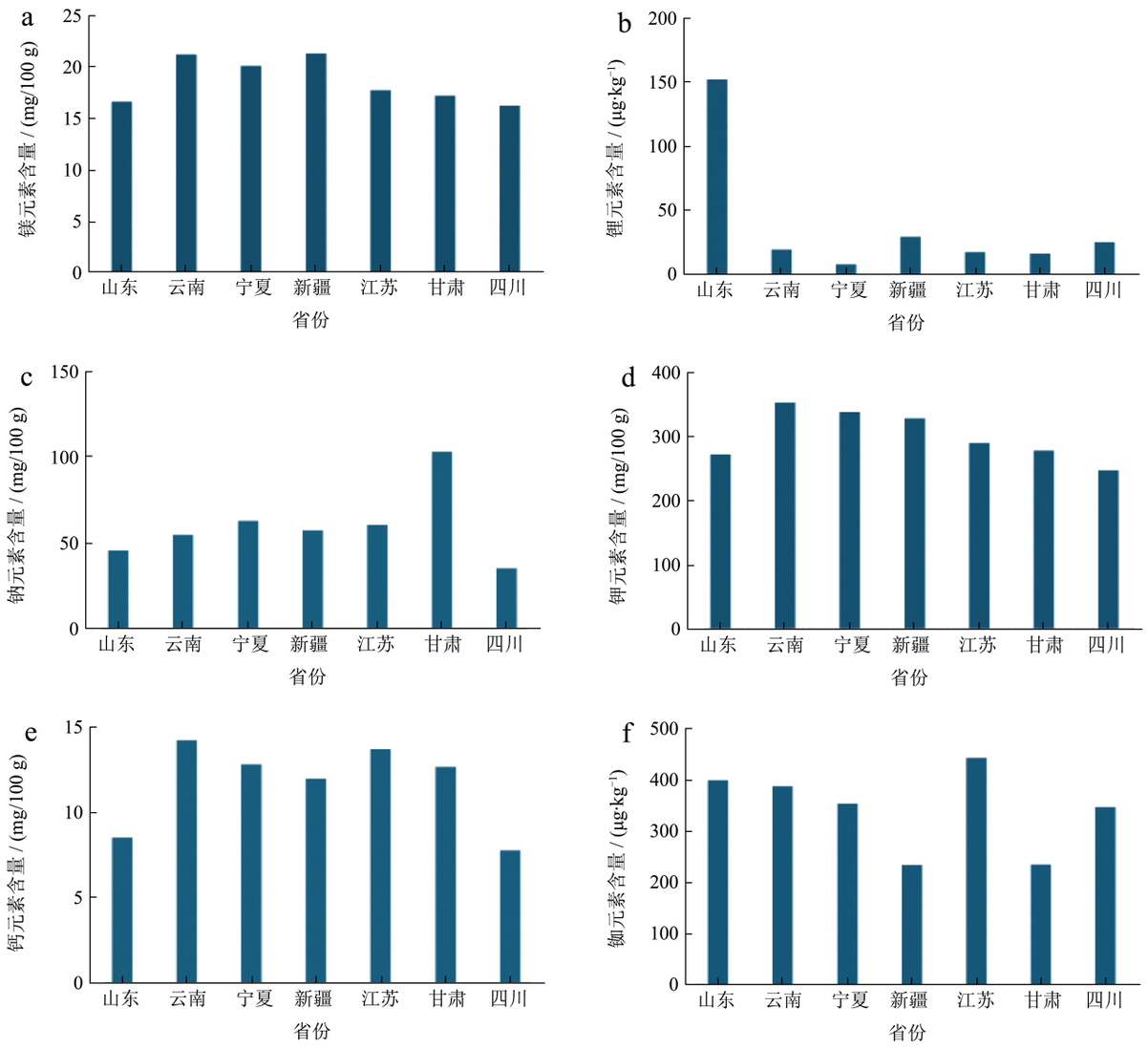


图 1 不同省份牛肉中典型矿物元素含量比较

Fig.1 Comparison of typical mineral element contents in beef from different provinces

2.2 数据模型分析

对于产地溯源研究，元素的地区差异性是构建模型的关键依据。通常优先选取在不同地区之间具有显著统计差异的元素作为建模特征，以最大程度保留产地指导性信号，提高模型判别能力。相反，那些在不同产地间差异较小或无统计显著性的元素，往往反映的是基础营养需求或系统背景水平，在判别模型中缺乏识别能力，若不筛除可能引入噪声变量，降低模型稳定性和泛化性能。因此，在本研究后续的建模过程，将综合采用多因素方差分析与随机森林算法，对所有测定元素进行统计显著性与判别贡献度的联合筛选，提取最具代表性的差异元素特征，并剔除无效变量，以此构建具有地理溯源能力的元素判别模型，为牛肉产地识别提供科学、可靠的技术支持。

数据分析技术路线如图 2 所示，为实现牛肉产地的精准溯源，引入随机森林（RF）算法对元素特征进行筛选与重要性评估。随机森林是一种集成学习方法，基于多棵决策树构建模型，能够有效处理高维数据并对特征的重要性进行量化分析。在随机森林中，特征重要性用于衡量各变量对模型预测结果的相对贡献程度，特征重要性越高代表该特征对分类性能影响越大。某一元素的特征重要性越高，意味着其在区分不同牛肉产地方面的作用越显著，反之如果特征重要性越小，则说明该元素对于产地溯源作用越小，过低特征重要性的元素变量可能会导致模型分类性能下降[27]。因此，首先我们使用随机森林算法的特征重要性对元素进行筛选，以剔除对溯源没有帮助的元素。

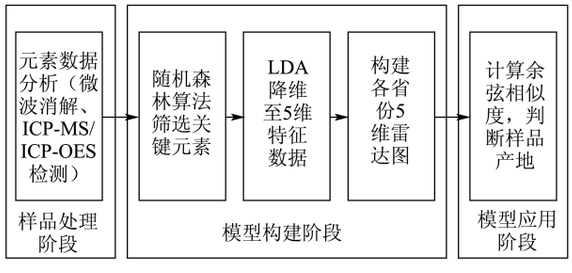


图 2 技术路线图

Fig.2 Technology roadmap

数据分析过程设定特征重要性阈值为 0.02，仅保留重要性高于该阈值的元素变量，结果显示 52 种元素中筛去 25 种，分别为 Cr、Ru、Ir、Pt、Au、Li、Sr、Rh、Te、Hg、Y、La、Ce、Pr、Nd、Sm、Eu、Gd、Tb、Dy、Ho、Er、Tm、Yb、Lu。图 3 结果为特征重要性大于 0.02 的 27 种元素，作为后续主要分析变量，分别为 Sc、Co、K、Se、Rb、V、Ba、As、Sn、Pb、Na、Cu、Mg、Ag、Al、P、Ni、Tl、Fe、Ti、Zn、Mn、Cd、Ca、Cs、Sb、Mo。这些被筛选的元素与表 1 中多重方差检验 p 值小于 0.05 的 18 个元素存在 16 种重合，占比达到 88.89%。表明随机森林所筛选出的元素在统计意义上亦具有显著的地区差异性，进一步验证了所选元素在牛肉产地识别中的判别能力。图 3 显示特征重要性大于 0.02 的元素分布，这些元素是构成牛肉省份分类模型的核心变量，可作为产地显著差异性的关键元素，同时也与农产品溯源相关文献中的关键指标具有高度一致性[28-30]。因此，采用随机森林结合特征重要性评估作为变量筛选手段，在牛肉产地溯源研究中具有较高的科学性与实用价值。

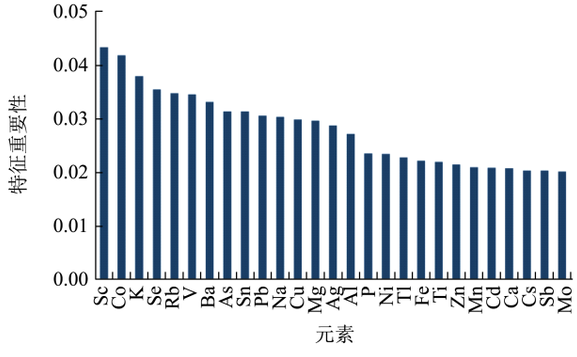


图 3 特征重要性大于 0.02 的元素

Fig.3 Elements with feature importance greater than 0.02

尽管已筛选至 27 个特征变量，但是直接用于构建雷达图仍面临维度过高、不利于可视化分析与后续建模的问题。为此，数据分析过程进一步引入线性判别分析（LDA）方法对元素数据进行降维处理。图 4 为 LDA 特征载荷热图，该图展示了原始属性在构建各线性判别轴时的影响力，其中 X 轴中的数据代表通过 LDA 计算出的线性判别轴，Y 轴中的数据代表各元素，图中数值的符号代表贡献方向，红色（正值）表示该原始属性值的增加，会正向推动样本在该 LDA 轴上的得分，蓝色（负值）表示该原始属性值的增加，会反向推动样本在该 LDA 轴上的得分。值的绝对值大小代表贡献强度，绝对值越大（颜色越深），说明该原始属性对构建这个 LDA 轴的影响力越强，接近 0 的值意味着该属性对这个轴基本没有贡献。从图中可知，第一判别轴（LDA1）主要由 V、Ba、Mg 等元素的强正向载荷和 Ti、Ca、Mn 等元素的强负向载荷所定义，具体而言，具有较高 V、Ba、Mg 等元素含量值和较低 Ti、Ca、Mn 等元素含量值的样本在 LDA1 上会获得较高的分数，LDA2 至 LDA5 同理。为了在降维后的判别空间中直观地比较各类别的特征轮廓，采用 5 个 LDA 轴并将其压缩为一维表示，从而形成每个省份牛肉样本的五维特征向量，基于该五维数据构建各省份牛肉的雷达图如图 5 所示。

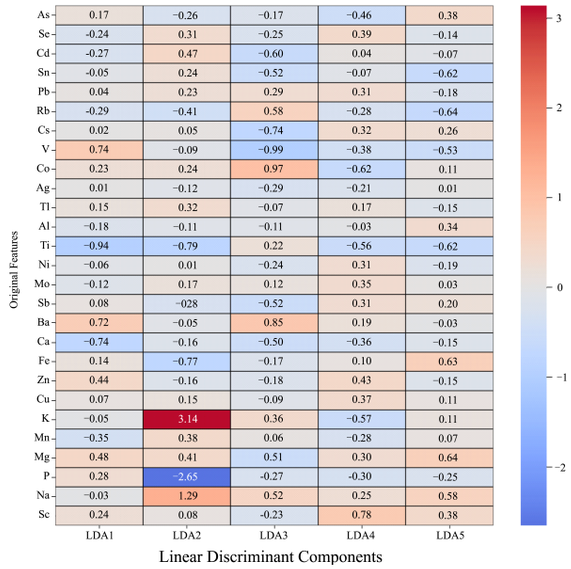


图 4 LDA 特征载荷热图

Fig.4 Elements with feature importance greater than 0.02

图 5 显示 7 个省份牛肉的元素特征差异。坐标轴的刻度表示 LDA 投影得分的范围，坐标轴的元素标签代表该 LDA 轴影响力最强的前 5 个元素，线条上的点距离中心点的远近，表示该类别的所有样本的平均得分在该 LDA 轴上的大小，每个类别多边形的形状和大小，共同构成了该类别的轮廓。图中清晰地显示，不同类别的多边形在形状和方向上存在显著差异。例如，新疆牛肉的元素轮廓在 LDA1 轴上表现出强烈的正向延伸，但在 LDA5 轴上得分偏低。与之形成鲜明对比的是，甘肃牛肉的元素轮廓在 LDA1 轴上则得分偏低而在 LDA5 中得分较高。投影分数的差异代表着不同省份牛肉元素含量的差异，各省份样本根据 LDA 降维结果呈现出不同的轮廓分布，说明该方法在产地溯源的鉴别能力较强，具备较好的图像识别与匹配潜力。为进一步验证模型的实用性与稳健性，通过外部样品进行验证实验，7 个省份的 12 例外部样品作为验证集，旨在评估模型在未知样本上的分类性能。

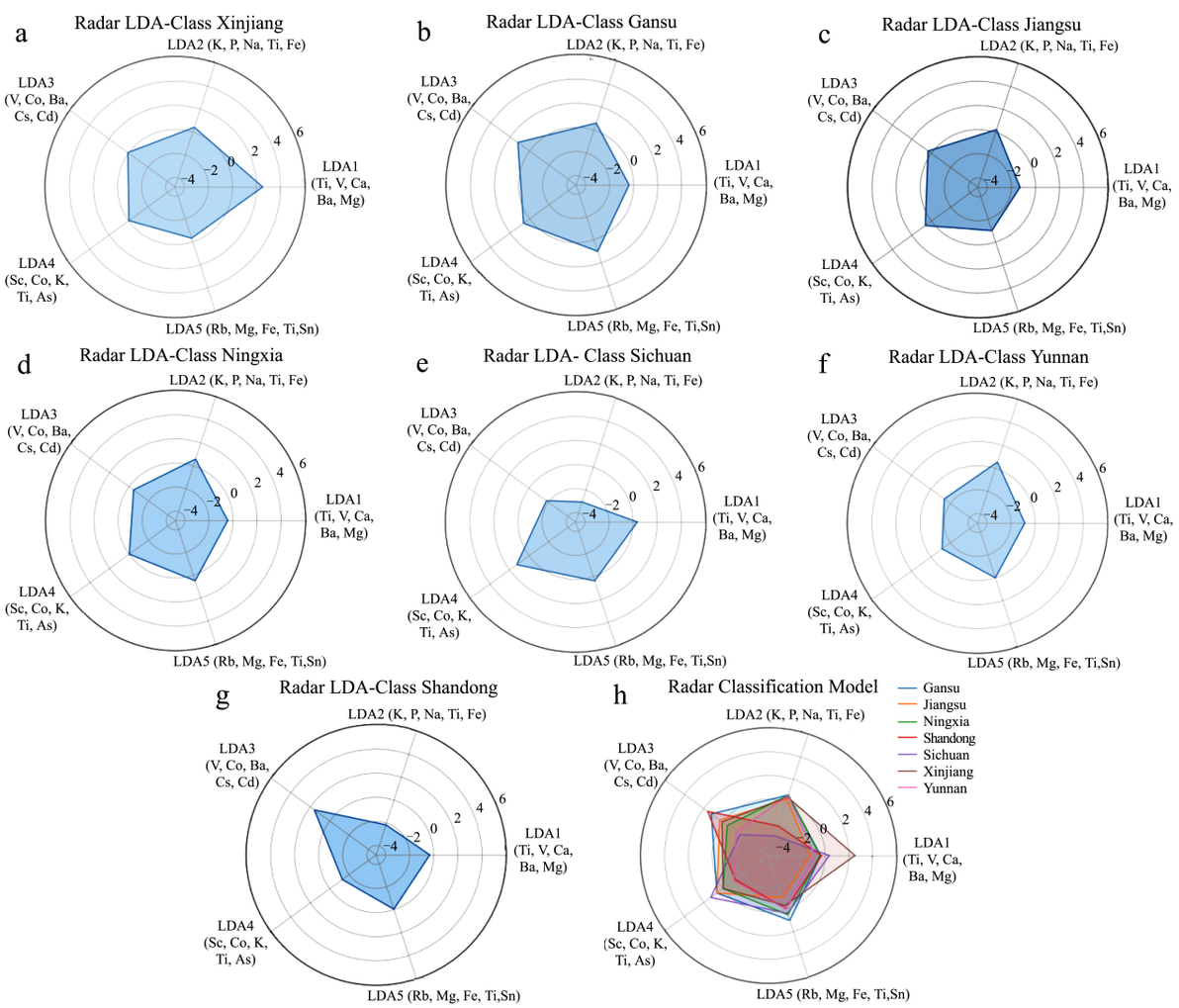


图 5 各省份牛肉元素轮廓雷达图

Fig.5 Radar contour maps of beef elements by province

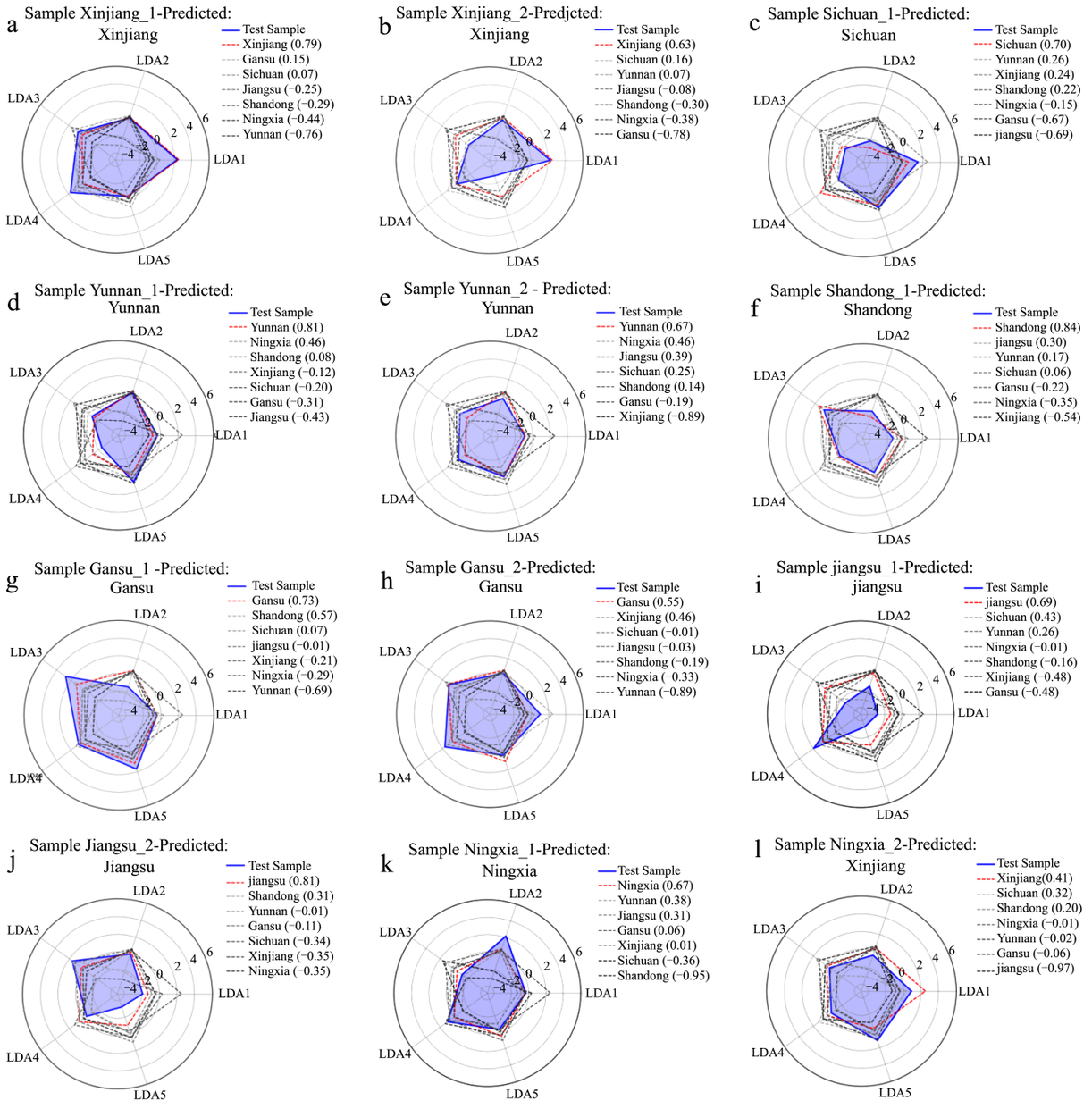


图 6 验证集样品与模型匹配结果

Fig.6 Matching results between validation set samples and the model

余弦相似度是一种衡量两个向量方向相似程度的指标，广泛用于文本、图像、用户偏好等领域的相似性计算[31]。在匹配过程中，本研究借助余弦相似度算法度量外部样本雷达图与各省份模型轮廓图之间的相似性，训练集所划分的省份类别会以不同的灰色虚线标识于雷达图中，将样品得到的元素数据筛除同样的 25 个低特征重要性的元素，进行相同的 LDA 降维之后，将其构建的样品雷达图使用余弦相似度算法与模型总雷达图进行匹配，算法会找到最符合测试样品雷达图轮廓的省份类别，并以红色虚线标出。12 例外部样品分类结果如图 6 所示，通过外部样品验证，结果显示 12 例外部样品中 11 例样品都得到了正确分类，预测错误 1 例，宁夏产地的牛肉样品 Ningxia_2 被错误判断为新疆，原因可能是宁夏和新疆两地同属西北干旱半干旱区，重叠纬度带的草原、绿洲土壤中元素的自然背景值有共性，肉牛饲养方式类似，导致该牛肉样品中元素种类和含量与新疆样品呈现近似，本研究所构建模型对外部样品的产地判别准确率达到了 91.67%，因此该牛肉产地判别模型对于 7 个省份牛肉样品的溯源能力是足够的，同时，为进一步提高牛肉产地判别模型的精准性，后续研究可持续增加训练集样品量。

3 结论

通过山东、云南、宁夏、新疆、江苏、甘肃、四川等7个省份的牛肉样品，利用电感耦合等离子体质谱（ICP-MS）与电感耦合等离子体发射光谱（ICP-OES）技术对样品中52种元素的含量进行了定量分析。基于随机森林算法的特征重要性，筛选出27种在不同产地间具有显著差异的关键元素（Sc、Co、K、Se、Rb、V、Ba、As、Sn、Pb、Na、Cu、Mg、Ag、Al、P、Ni、Tl、Fe、Ti、Zn、Mn、Cd、Ca、Cs、Sb、Mo），这些元素在构建牛肉产地鉴别模型中具有重要的分类判别价值。为了实现对牛肉产地的准确判别，采用随机森林算法筛选有效元素，随后使用线性判别分析（LDA）将高维数据降维为五个维度的数据，并以降维结果构建各省份的元素雷达图特征模板。通过余弦相似度算法进行样品匹配与溯源判别，构建了牛肉产地预测模型，并引入外部样品进行模型验证，最终判别准确率达91.67%，表现出良好的泛化性能和判别能力。

不同省份牛肉中矿物元素的显著差异主要受到以下三类因素的综合影响：一是地质背景与土壤矿物组成决定了各地区元素的基础赋存能力，二是水源中矿化成分直接影响牛肉中矿物质含量，三是饲料与饲养环境对牛肉中宏量元素和痕量元素的沉积造成的影响。上述多重因素叠加作用，形成了各省牛肉样品的元素指纹特征差异，为区域来源判别提供了坚实的科学依据。因此，本研究构建的多元素指纹识别与机器学习溯源模型，不仅实现了对牛肉产地的高效判别，也为食品地理溯源技术的发展提供了重要的理论支撑与方法参考，具有良好的实际应用前景。

参考文献

[1] ARTHINGTON J D, RANCHES J. Trace mineral nutrition of grazing beef cattle [J]. Animals, 2021, 11(10): 2767.

[2] ZHENG H Q, LI H X, KONG L C, et al. Association of dietary patterns with cognitive function in older people: Results from the Chinese longitudinal healthy longevity survey [J]. Medicine Advances, 2023, 1(3): 246-259.

[3] 罗菲.消费者牛肉偏好的原产地效应及影响因素研究——基于选择实验[D].南昌:江西财经大学,2021.

[4] FU R, LI C G, WANG L M, et al. Chinese consumer preference for beef with geographical indications and other attributes [J]. Meat Science, 2024, 212: 109475.

[5] ELPHUS J M, INGRID M M, DAVIES V N, et al. Determination of intrinsic sodium in Grass-Fed bonsmara (*Bos taurus indicus*) meat [J]. Applied Sciences, 2025, 15(12): 6612.

[6] LI X M, ZANG M W, LI D, et al. Meat food fraud risk in Chinese markets 2012-2021 [J]. npj Science of Food, 2023, 7(1): 12.

[7] 赵航,程玛丽.2023年我国肉牛产业发展回顾与2024年展望[J].畜牧产业,2024,4:32-39.

[8] DEHELEAN A, CRISTEA G, PUSCAS R, et al. Assigning the geographical origin of meat and animal rearing system using isotopic and elemental fingerprints [J]. Applied Sciences, 2022, 12(23): 12391.

[9] GREET D, CARLOS V P, R D. Influence of the consumption of meat contaminated with anabolic steroids on doping tests [J]. Analytica Chimica Acta, 1993, 275(1-2): 49-56.

[10] ALJAZ M, CLAUDIO M. Advances in Food Metabolomics [J]. Applied Sciences, 2025, 15(15): 8283.

[11] 张雨婷.基于矿物元素和有机成分指纹分析技术的红心猕猴桃产地溯源研究[D].重庆:重庆三峡学院,2025.

[12] 王曦奥,彭贤瑞,胡丽松,等.基于矿质元素指纹特征的海南咖啡产地溯源研究[J].热带作物学报,2025,46(2):398-406.

[13] 黄锐敏,李星,何肖云,等.基于矿物元素指纹图谱技术的武夷岩茶产地溯源研究[J].福建农业科技,2025,56(6):17-25.

[14] RENATO M, IRINA A D, ISABEL C, et al. Elemental fingerprinting of wild and farmed fish muscle to authenticate and validate production method [J]. Foods, 2022, 11(19): 3081.

[15] WANG Q, LIU H J, ZHAO S S, et al. Discrimination of mutton from different sources (regions, feeding patterns and species) by mineral elements in Inner Mongolia, China [J]. Meat Science, 2021, 174: 108415.

[16] 齐婧,李莹莹,姜锐,等.基于矿物元素指纹分析中国地理标志猪肉的产地溯源[J].现代食品科技,2020,36(3):267-274,55.

[17] ZHAO L Y, ZHANG H R, HUANG F, et al. Authenticating Tibetan pork in China by tracing the species and geographical features based on stable isotopic and multi-elemental fingerprints [J]. Food Control, 2023, 145: 109411.

[18] FERNANDES E A D N, MAZOLA Y T, BACCHI M A, et al. Discriminating beef producing countries by multi-element analysis and machine learning [J]. Advances in Artificial Intelligence and Machine Learning, 2021, 1(1): 1-11.

[19] ZHANG H R, LIU W Y, SHEN Q S, et al. Discrimination of geographical origin and species of China's cattle bones based on multi-element analyses by inductively coupled plasma mass spectrometry [J]. Food Chemistry, 2021, 356: 129619.

[20] LI J L, QIAN J P, CHEN J Y, et al. Recent advances of machine learning in the geographical origin traceability of food and agro-products: A review [J]. Comprehensive Reviews in Food Science and Food Safety, 2025, 24(1): e70082.

[21] BIGLIA A, BARGE P, TORTIA C, et al. Artificial intelligence to boost traceability systems for fraud prevention in the meat industry[J]. Journal of Agricultural Engineering, 2022, 53(4): 1328.

[22] ZONG W L, ZHAO S S, LI Y L, et al. Quality control and origin traceability of Tibetan yak meat using mid-infrared spectroscopy combined with multivariate analysis and machine learning [J]. Journal of Consumer Protection and Food Safety, 2025, 20: 235-244.

[23] QI J, LI Y Y, ZHANG C, et al. Geographic origin discrimination of pork from different Chinese regions using mineral elements analysis assisted by machine learning techniques [J]. Food Chemistry, 2021, 337: 127779.

[24] AN J M, HUR S H, KIM H, et al. Determination of the geographical origin of chicken (breast and drumstick) using ICP-OES and ICP-MS: chemometric analysis [J]. Food Chemistry, 2024, 437: 137836.

[25] QIE M J, ZHANG B, LI Z, et al. Data fusion by ratio modulation of stable isotope, multi-element, and fatty acids to improve geographical traceability of lamb [J]. Food Control, 2021, 120: 107549.

[26] GB 5009.268-2016,食品安全国家标准 食品中多元素的测定[S].

[27] GONZÁLEZ-DOMÍNGUEZ R, SAYAGO A, FERNÁNDEZ-RECAMALES A. An overview on the application of chemometrics tools in food authenticity and traceability [J]. Foods, 2022, 11(23): 3940.

[28] SALIM N A A, DAUD N M, GRIBOFF J, et al. Elemental assessments in paddy soil for geographical traceability of rice from peninsular malaysia [J]. Rice Science, 2023, 30(5): 486-498.

[29] GUO N, WU Q, SHI C, et al. Geographical discrimination of *Cyclocarya paliurus* tea for origin traceability based on multielement analysis by ICP-OES and chemometrics multivariate [J]. Chinese Herbal Medicines, 2023, 15(1): 63-68.

[30] JI X Y. Multielemental analysis using inductively coupled plasma mass spectrometry and optical emission spectroscopy for tracing the geographical origin of food [J]. Journal of Analytical Chemistry, 2025, 80(7): 1140-1151.

[31] XIA P P, ZHANG L, LI F Z. Learning similarity with cosine similarity ensemble [J]. Information Sciences, 2015, 307: 39-52.