

基于稀土元素指纹分析识别葡萄酒原产地

赵芳¹, 林立², 孙翔宇¹, 毛文华³, 孙玉强⁴, 战吉成¹

(1. 中国农业大学食品科学与营养工程学院, 北京 100083) (2. 陕西师范大学食品工程与营养科学学院, 陕西西安 710062) (3. 中国农业机械化科学研究院, 北京 100083) (4. 国家果酒及果蔬饮品质量监督检验中心, 吉林通化 134001)

摘要: 本研究为探讨稀土元素指纹分析对于葡萄酒原产地判别的可行性, 采用电感耦合等离子体质谱法 (inductively coupled plasma-mass spectrometry, ICP-MS) 测定了三个原产地 228 个葡萄酒中的 15 种稀土元素含量, 并对数据进行相关性分析、方差分析和判别分析。葡萄酒原产地和稀土元素含量间显著相关 ($p < 0.01$), 不同原产地间存在显著性差异 ($p < 0.001$), 沙城葡萄酒中稀土元素含量最低, 通化最高, 贺兰山东麓居中。Fisher 线性判别分析 (fisher linear discriminant analysis, FLD) 模型对沙城、贺兰山东麓、通化三产地的交叉验证判别率分别为 92.98%、98.25%、100.00%, 外部验证判别率分别为 84.21%、89.47%、100.00%; 偏最小二乘判别分析 (partial least squares-discriminant analysis, PLS-DA) 模型判别能力略差, 三个产地的交叉验证判别率分别为 82.46%、98.25%、91.23%, 外部验证判别率仅为 73.68%、84.21%、100.00%。说明稀土元素含量结合多元统计分析可以应用于葡萄酒原产地的鉴别。

关键词: 葡萄酒; 稀土元素; ICP-MS; 原产地; 判别分析

文章编号: 1673-9078(2015)2-261-267

DOI: 10.13982/j.mfst.1673-9078.2015.2.041

Identification of Wine of Origins Using Rare Earth Element Fingerprinting

ZHAO Fang¹, LIN Li², SUN Xiang-yu¹, MAO Wen-hua³, SUN Yu-qiang⁴, ZHAN Ji-cheng¹

(1. College of Food Science and Nutritional Engineering, China Agricultural University, Beijing 100083, China)

(2. College of Food Engineering and Nutritional Science, Shanxi Normal University, Xi'an 710062, China)

(3. Chinese Academy of Agricultural Mechanization Sciences, Beijing 100083, China)

(4. National Ratafia and Fruits and Vegetables Drinks Quality Surveillance Test Center, Tonghua 134001, China)

Abstract: In this study, in order to explore the feasibility of using rare earth element fingerprinting to identify wine of origins, inductively coupled plasma-mass spectrometry (ICP-MS) was employed to measure the contents of 15 rare earth elements in 228 wine samples from three origins. The data were analyzed by correlation analysis, analysis of variance (ANOVA), and discriminant analysis. The wine of origin was significantly correlated with the rare earth element contents ($p < 0.01$), while the rare earth element contents were significantly different for the three origins ($p < 0.001$). The highest rare earth element content was found in the wines from Tonghua, followed by wines from the eastern foot of the Helan Mountains, and the lowest was found in the wines from Shacheng. The cross-validated accuracy rates of the wines from Shacheng, the eastern foot of the Helan Mountains, and Tonghua, using Fisher linear discriminant analysis (FLD), were 92.98%, 98.25%, and 100.00%, respectively, while the accuracy rates by external validation were 84.21%, 89.47%, and 100.00%, respectively. The discrimination capacity of partial least squares discriminant analysis (PLS-DA) was slightly lower; the cross-validated accuracy rates of the three wine origins were 82.46%, 98.25%, and 91.23%, respectively, while the accuracy rates by external validation were 73.68%, 84.21%, and 100.00%, respectively. Therefore, it is possible to identify wine of origins by rare earth element fingerprinting and multivariate statistical analysis.

Key words: wine; rare earth elements; inductively coupled plasma-mass spectrometry; place of origin; discriminant analysis

收稿日期: 2014-06-25

基金项目: 国家公益性行业科研专项 (2012104019-3); 国家十二五科技支撑计划 (2012BAD31B07)

作者简介: 赵芳 (1985-), 女, 在读博士, 研究方向为葡萄酒地理标志保护

通讯作者: 战吉成 (1972-), 男, 博士 (后), 副教授, 研究方向为葡萄酒

生理生化与分子生物学研究

葡萄酒的品质主要取决于酿酒葡萄的质量和与之相适应的加工工艺。而酿酒葡萄的质量直接依赖于产区的品种、土壤、气候、肥料、农药和相应的栽培方式, 且在某一特定的区域内, 往往会形成一些特定的生产工艺, 从而生产出具有特定风格的葡萄酒。因此, 葡萄酒是一种明显带有地理标志 (geographical

indications, 简称 GI) 特征的产品, 地理标志葡萄酒不仅限定酿酒葡萄生长的地域特征, 更包括相应的栽培制度和酿造工艺等。

我国作为新兴的葡萄酒国家, 市场潜力巨大, 为使我国的葡萄酒产业实现持续、快速、健康发展, 就必须建立并实施葡萄酒原产地保护体系。地理标志注册, 有利于保护自然资源和人文资源, 也有利于保护生产者、经营者和消费者的权益。目前, 通过我国地理标志认证的葡萄酒有沙城葡萄酒、贺兰山东麓葡萄酒、通化山葡萄酒等^[1-3]。

为有效地保护地理标志葡萄酒, 急需一种能精确快速鉴别葡萄酒原产地的方法。目前, 已有多种检测方法应用于葡萄酒的产地识别^[4-8]。从矿质元素的检测分析来看, Pérez-Trujill 等^[9]借助 ICP-MS 测定葡萄酒中 Ti、Cu、Zn 等 23 种矿质元素, 建立了西班牙加那利群岛 4 个地区的产地判别模型, 正确率达到了 96.7%。Boschetti 等^[10]基于 K、Mn、Rb 和 Sr 4 种元素, 采用连续光源火焰原子吸收光谱法 (high resolution continuum source flame AAS, HR-CS-FAAS), 对巴西不同地区的葡萄酒进行了有效区分。Coetzee 等^[11]通过 ICP-MS 对葡萄酒中 B 的同位素进行了测定, 统计分析结果表明, ¹¹B/¹⁰B 比值可以用来区分南非、意大利和法国的葡萄酒。因此, 利用不同地区葡萄酒矿质元素的指纹差异, 通过数据挖掘, 建立葡萄酒的原产地判别模型是可行的。

稀土元素是周期表中 IIIB 族钇和镧系元素的总称, 化学性质极为相似, 在葡萄酒中的分布和土壤中的分布差异较小, 即可以通过对葡萄酒中的稀土元素含量进行统计分析, 以此来作为产地溯源的手段。已有研究基于稀土元素的指纹分析, 对普洱古树茶和台地茶进行了产地鉴别和身份识别^[12]。本研究的目的在于通过测定葡萄酒中稀土元素的含量, 借助于化学计量学方法, 建立沙城、贺兰山东麓和通化的产地判别模型, 识别葡萄酒的原产地。

1 材料与方法

1.1 材料

根据国标^[1-3]中的规定, 通过酒厂取样方式采集葡萄酒样品共计 228 个, 分别来自沙城、贺兰山东麓和通化三个产地不同的酒庄, 采取的样品中, 每个产区尽可能包括不同年份和不同品种的酒样, 以此减小葡萄品种和年份对产地识别正确率的影响 (如表 1)。

1.2 仪器与试剂

Agilent 7500a 电感耦合等离子体质谱仪, 美国, Agilent 公司; Milli-Q 超纯水系统, 美国, Millipore 公司; 标准贮备液: 10 μg/mL 稀土混合标准贮备液, 含 16 种稀土元素, Sc、Y、La、Ce、Pr、Nd、Sm、Eu、Gd、Tb、Dy、Ho、Er、Tm、Yb、Lu, 美国, Agilent 公司。

表 1 不同产地葡萄酒样品的品种、年份以及数量

| 原产地 | 品种 | 年份 | 数量 |
|-----|-----------------------|-----------|----|
| SC | 赤霞珠、美乐、黑比诺、蛇龙珠、马瑟兰、西拉 | 2008~2012 | 76 |
| HLS | 赤霞珠、美乐、蛇龙珠、西拉、品利珠 | 2008~2012 | 76 |
| TH | 赤霞珠、双优、双红、公酿一号 | 2008~2012 | 76 |

注: SC-沙城; HLS-贺兰山; TH-通化。

内标溶液: 1000 μg/mL Rh 混合标准溶液购于中国计量科学研究院, 用 1% HNO₃ 稀释为 1 μg/mL, 由内标管在线引入质谱仪; 质谱调谐液: Li、Co、Y、Ce、Tl, 浓度均为 10 ng/mL, 美国, Agilent 公司; 浓 HNO₃, 65%, 优级纯, 德国, Merck 公司; 乙醇, 优级纯, 德国, Merck 公司; 超纯水, 18.2 MΩ。

实验中所使用的玻璃仪器及聚四氟乙烯瓶子等均用 20% HNO₃ 浸泡过夜, 再用超纯水冲洗干净, 烘干后使用。

1.3 ICP-MS 工作条件

由质谱调谐液调试仪器, 经全自动调谐优化给出工作参数, 满足 ICP-MS 使用的灵敏度、氧化物、双电荷等各项指标。调谐后的具体参数如下: 射频功率: 1340 W; 冷却气流速: 15 L/min; 载气流速: 1.22 L/min; 辅助气流速: 1.1 L/min; 雾化室温度: 2±0.1 °C; 采样锥和截取锥: Ni; 采样深度: 9 mm; 氧化物: <0.5%; 双电荷: <3%。

1.4 样品测定

空白溶液: 1% HNO₃ 溶液, 其中乙醇含量为 6%, 可以有效消除乙醇对样品测定带来的基体干扰。

标准曲线绘制: 用空白溶液将稀土元素混合标准贮备液逐级稀释为 5、10、20、50、100、200、500、1000 ng/L 的混合标准溶液。在 ICP-MS 的工作条件下, 采集空白溶液和标准溶液系列, 根据葡萄酒中稀土元素的含量范围, 选择合适的标准溶液浓度区间, 由仪器自动绘制标准曲线。

葡萄酒样品用 1% HNO_3 溶液按照 1:1 的比例稀释后, 直接进行 ICP-MS 测定, 采用全定量分析, 结果由标准曲线自动计算所得。

1.5 数据处理

本研究对不同原产地葡萄酒中 15 种稀土元素的含量进行了最小显著性差异 (least significant difference, LSD) 多重比较 ($p < 0.05$) 和单因素方差分析 (ANOVA), 判定葡萄酒中稀土元素含量在不同产地间的差异; 同时对产地与稀土元素含量之间、稀土元素两两之间的含量做了相关性分析, 结果以 Pearson 相关系数表示, 判定产地与稀土元素含量的关系, 以及稀土元素含量彼此间的分布关系。以上方法采用 SPSS 20.0 (美国, IBM 公司) 软件进行分析。

根据稀土元素含量, 分别对每个产地的葡萄酒样品做组内马氏距离分析, 基于马氏距离, 将样品集按照 3:1 的比例分为训练集 ($n=57$) 和测试集 ($n=19$), 最小值和最大值归于训练集; 对训练集中稀土元素含量做 boxcox 变换, 使其尽量满足正态分布, 基于数据变换参数, 对测试集中元素作同样变换, 变换后的数据用于建立原产地判别模型, 以提高模型的稳健性。以上分析采用 Matlab 7.6 (美国, MathWorks 公司) 软件进行处理。

本研究采用 FLD 中的逐步判别法和 PLS-DA 两种模式识别方法对葡萄酒原产地进行识别, 两种方法均用“留一法”交叉验证和测试集外部验证检验模型效果。

FLD 是模式识别分类中一种常用的特征提取方法^[13], 将高维的模式样本投影到低维的最佳鉴别矢量空间, 投影后使样本在新的子空间即典型变量有最大的组间距离和最小的组内距离。本研究中, 有 15 个稀土元素变量, 远小于训练集建模样本数 171, 因此, 不需要先对样本集数据进行主成分分析 (principal component analysis, PCA), 可直接作为输入变量。基于变量入选函数的 F 检验值筛选自变量, 采用 Wilks' Lambda 方法建立 FLD 模型。该分析采用 SPSS 20.0 软件进行。

偏最小二乘回归 (partial least squares, PLS) 法中, 主成分不仅对原始变量有最大限度的综合信息, 还尽可能地解释因变量的相关信息。PLS-DA 法来源于 PLS, 需对每一个样品赋予一个分类变量, 建立分类变量和自变量间的回归模型, 根据分类变量的预测值判定所属类别^[14]。本研究中, 对样品赋予产地分类

变量 Y, 分别设定 SC、HLS、TH 的相应参考值为 1、0 和 -1, 以 0.5 和 -0.5 作为不同产地间的分界线。对训练集样品, 建立 PLS-DA 校正模型, 以预测残差平方和 (predicted residual sum of squares, PRESS) 确定模型的主成分数, 之后以模型相关系数 (correlation coefficient, r)、交互验证均方根误差 (root mean square error of cross validation, RMSEC) 和验证均方根误差 (root mean square error of prediction, RMSEP) 评价模型拟合效果和预测精度。该分析采用 Unscrambler 10.3 (挪威, CAMO 公司) 软件进行。

2 结果与讨论

2.1 差异性分析及相关性分析

通过对沙城、贺兰山东麓和通化葡萄酒中 15 种稀土元素含量的 LSD 多重比较分析 (表 2), 结果如下: 首先, 不同产地葡萄酒中每一种稀土元素含量均存在显著性差异 ($p < 0.001$), 稀土元素在沙城葡萄酒中的含量最低, 在通化葡萄酒中的含量最高, 在贺兰山东麓葡萄酒中的含量介于两者之间。由 F 值可以看出, 葡萄酒产地间差异最大的稀土元素是 Y、Ce、Pr, 这可能使其在产地识别中贡献率更大。其次, 有部分稀土元素含量标准偏差较大, 尤其在贺兰山东麓酒样中, 如 La、Ce、Pr、Nd, 说明同一产地不同酒庄间存在一定的差异。同时, 在三个产地的葡萄酒中, 镧系元素中含量最高的是 La、Ce、Nd, 而钇族元素中 Y 的含量也很高, 在沙城酒样中含量最高, 在贺兰山和通化酒样中含量仅次于 Ce。

基于不同产地葡萄酒中稀土元素含量的相关性分析 (表 2), 可以看到, 葡萄酒的产地和稀土元素含量之间显著相关 ($p < 0.01$), Pearson 相关系数 η 介于 0.43 和 0.69 之间, η 越大, 相应的 F 值越大, 在葡萄酒产地识别中的作用可能就越大。同时, 如表 3 所示, 无论哪一个产地的葡萄酒, 稀土元素含量彼此之间呈显著相关 ($p < 0.01$), Pearson 相关系数介于 0.64 和 0.99 之间, 这种结果可能是由稀土元素相似的化学性质所引起。

方法检测限按照国际纯粹与应用化学联合会 (International Union of Pure and Applied Chemistry, IUPAC) 的规定计算。取 11 次平行测定空白溶液的结果, 按空白溶液响应信号值标准偏差的 3 倍计算各元素的检出限。葡萄酒中各元素的含量均高于方法检出限, 说明葡萄酒经稀释后, 用 ICP-MS 检测其稀土元

素含量的方法是可行的。

表 2 不同产地葡萄酒中稀土元素含量 (ng/L) 的差异性分析、与产地的相关性分析以及方法检出限

Table 2 Analysis of variance (F), correlation (η), and limits of detection (LOD) of the rare earth element contents (ng/L) in wines from different origins

| 稀土元素 | 原产地 | | | F | η | LOD |
|------|----------------------------|----------------------------|-----------------------------|-----------------------|--------------------|-------|
| | SC | HLS | TH | | | |
| | n=76 | n=76 | n=76 | | | |
| Y | 163.23±88.76 ^c | 515.23±400.19 ^b | 1067.31±540.93 ^a | 102.79 ^{***} | 0.69 ^{**} | 1.16 |
| La | 80.21±78.18 ^c | 459.2±497.87 ^b | 899.77±521.53 ^a | 72.93 ^{***} | 0.63 ^{**} | 4.69 |
| Ce | 146.32±116.26 ^c | 651.13±672.55 ^b | 1469.68±832.35 ^a | 87.77 ^{***} | 0.66 ^{**} | 1.66 |
| Pr | 18.76±16.37 ^c | 78.69±83.71 ^b | 210.39±124.63 ^a | 96.06 ^{***} | 0.66 ^{**} | 0.65 |
| Nd | 97.54±72.98 ^c | 436.38±458.02 ^b | 874.68±525.29 ^a | 70.49 ^{***} | 0.62 ^{**} | 2.71 |
| Sm | 53.06±18.64 ^c | 104.97±98.37 ^b | 188.06±118.99 ^a | 43.73 ^{***} | 0.52 ^{**} | 6.92 |
| Eu | 25.36±7.24 ^c | 32.78±17.45 ^b | 48.83±23.61 ^a | 35.86 ^{***} | 0.48 ^{**} | 2.81 |
| Gd | 45.37±18.96 ^c | 123.26±109.84 ^b | 190.11±123.58 ^a | 43.20 ^{***} | 0.53 ^{**} | 11.48 |
| Tb | 5.76±2.82 ^c | 14.44±11.68 ^b | 25.36±13.95 ^a | 64.87 ^{***} | 0.60 ^{**} | 0.60 |
| Dy | 35.41±16.71 ^c | 88.49±75.13 ^b | 140.62±75.74 ^a | 54.11 ^{***} | 0.57 ^{**} | 2.43 |
| Ho | 9.94±3.77 ^c | 20.72±14.78 ^b | 38.14±18.12 ^a | 82.25 ^{***} | 0.64 ^{**} | 0.71 |
| Er | 36.04±16.23 ^c | 61.01±42.04 ^b | 110.53±53.15 ^a | 67.51 ^{***} | 0.60 ^{**} | 3.35 |
| Tm | 7.50±3.00 ^c | 11.61±6.28 ^b | 17.49±9.03 ^a | 44.28 ^{***} | 0.53 ^{**} | 0.51 |
| Yb | 56.99±22.23 ^c | 73.24±39.56 ^b | 116.45±59.94 ^a | 38.10 ^{***} | 0.49 ^{**} | 4.69 |
| Lu | 12.76±4.37 ^c | 15.83±6.98 ^b | 22.78±12.25 ^a | 27.56 ^{***} | 0.43 ^{**} | 0.93 |

注: 1.表中各数据用平均值±标准偏差表示; 2.不同小写字母表示不同产地间该元素含量有显著差异 (p<0.05); 3.***表示显著性差异 (p<0.001), **表示产地与该元素含量显著相关 (p<0.01)。

表 3 葡萄酒中稀土元素含量彼此之间的相关性分析

Table 3 Correlation analysis among the rare earth element contents in wines.

| 稀土元素 | 稀土元素 | | | | | | | | | | | | | | |
|------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| | Y | La | Ce | Pr | Nd | Sm | Eu | Gd | Tb | Dy | Ho | Er | Tm | Yb | Lu |
| Y | 1 | 0.90 ^{**} | 0.92 ^{**} | 0.93 ^{**} | 0.93 ^{**} | 0.92 ^{**} | 0.84 ^{**} | 0.93 ^{**} | 0.96 ^{**} | 0.97 ^{**} | 0.98 ^{**} | 0.97 ^{**} | 0.89 ^{**} | 0.88 ^{**} | 0.81 ^{**} |
| La | 0.90 ^{**} | 1 | 0.98 ^{**} | 0.97 ^{**} | 0.98 ^{**} | 0.93 ^{**} | 0.84 ^{**} | 0.91 ^{**} | 0.90 ^{**} | 0.90 ^{**} | 0.87 ^{**} | 0.84 ^{**} | 0.74 ^{**} | 0.71 ^{**} | 0.64 ^{**} |
| Ce | 0.92 ^{**} | 0.98 ^{**} | 1 | 0.98 ^{**} | 0.98 ^{**} | 0.93 ^{**} | 0.84 ^{**} | 0.90 ^{**} | 0.91 ^{**} | 0.90 ^{**} | 0.89 ^{**} | 0.86 ^{**} | 0.75 ^{**} | 0.73 ^{**} | 0.66 ^{**} |
| Pr | 0.93 ^{**} | 0.97 ^{**} | 0.98 ^{**} | 1 | 0.99 ^{**} | 0.95 ^{**} | 0.87 ^{**} | 0.92 ^{**} | 0.93 ^{**} | 0.91 ^{**} | 0.90 ^{**} | 0.88 ^{**} | 0.77 ^{**} | 0.75 ^{**} | 0.67 ^{**} |
| Nd | 0.93 ^{**} | 0.98 ^{**} | 0.98 ^{**} | 0.99 ^{**} | 1 | 0.97 ^{**} | 0.89 ^{**} | 0.95 ^{**} | 0.94 ^{**} | 0.94 ^{**} | 0.91 ^{**} | 0.88 ^{**} | 0.77 ^{**} | 0.74 ^{**} | 0.67 ^{**} |
| Sm | 0.92 ^{**} | 0.93 ^{**} | 0.93 ^{**} | 0.95 ^{**} | 0.97 ^{**} | 1 | 0.92 ^{**} | 0.98 ^{**} | 0.96 ^{**} | 0.95 ^{**} | 0.92 ^{**} | 0.90 ^{**} | 0.82 ^{**} | 0.77 ^{**} | 0.71 ^{**} |
| Eu | 0.84 ^{**} | 0.84 ^{**} | 0.84 ^{**} | 0.87 ^{**} | 0.89 ^{**} | 0.92 ^{**} | 1 | 0.90 ^{**} | 0.89 ^{**} | 0.87 ^{**} | 0.86 ^{**} | 0.85 ^{**} | 0.79 ^{**} | 0.76 ^{**} | 0.71 ^{**} |
| Gd | 0.93 ^{**} | 0.91 ^{**} | 0.90 ^{**} | 0.92 ^{**} | 0.95 ^{**} | 0.98 ^{**} | 0.90 ^{**} | 1 | 0.97 ^{**} | 0.98 ^{**} | 0.94 ^{**} | 0.91 ^{**} | 0.85 ^{**} | 0.79 ^{**} | 0.74 ^{**} |
| Tb | 0.96 ^{**} | 0.90 ^{**} | 0.91 ^{**} | 0.93 ^{**} | 0.94 ^{**} | 0.96 ^{**} | 0.89 ^{**} | 0.97 ^{**} | 1 | 0.98 ^{**} | 0.97 ^{**} | 0.95 ^{**} | 0.88 ^{**} | 0.84 ^{**} | 0.78 ^{**} |
| Dy | 0.97 ^{**} | 0.90 ^{**} | 0.90 ^{**} | 0.91 ^{**} | 0.94 ^{**} | 0.95 ^{**} | 0.87 ^{**} | 0.98 ^{**} | 0.98 ^{**} | 1 | 0.97 ^{**} | 0.95 ^{**} | 0.88 ^{**} | 0.85 ^{**} | 0.79 ^{**} |
| Ho | 0.98 ^{**} | 0.87 ^{**} | 0.89 ^{**} | 0.90 ^{**} | 0.91 ^{**} | 0.92 ^{**} | 0.86 ^{**} | 0.94 ^{**} | 0.97 ^{**} | 0.97 ^{**} | 1 | 0.98 ^{**} | 0.93 ^{**} | 0.91 ^{**} | 0.86 ^{**} |
| Er | 0.97 ^{**} | 0.84 ^{**} | 0.86 ^{**} | 0.88 ^{**} | 0.88 ^{**} | 0.90 ^{**} | 0.85 ^{**} | 0.91 ^{**} | 0.95 ^{**} | 0.95 ^{**} | 0.98 ^{**} | 1 | 0.95 ^{**} | 0.94 ^{**} | 0.89 ^{**} |
| Tm | 0.89 ^{**} | 0.74 ^{**} | 0.75 ^{**} | 0.77 ^{**} | 0.77 ^{**} | 0.82 ^{**} | 0.79 ^{**} | 0.85 ^{**} | 0.88 ^{**} | 0.88 ^{**} | 0.93 ^{**} | 0.95 ^{**} | 1 | 0.95 ^{**} | 0.93 ^{**} |
| Yb | 0.88 ^{**} | 0.71 ^{**} | 0.73 ^{**} | 0.75 ^{**} | 0.74 ^{**} | 0.77 ^{**} | 0.76 ^{**} | 0.79 ^{**} | 0.84 ^{**} | 0.85 ^{**} | 0.91 ^{**} | 0.94 ^{**} | 0.95 ^{**} | 1 | 0.96 ^{**} |
| Lu | 0.81 ^{**} | 0.64 ^{**} | 0.66 ^{**} | 0.67 ^{**} | 0.67 ^{**} | 0.71 ^{**} | 0.71 ^{**} | 0.74 ^{**} | 0.78 ^{**} | 0.79 ^{**} | 0.86 ^{**} | 0.89 ^{**} | 0.93 ^{**} | 0.96 ^{**} | 1 |

注: **表示稀土元素含量之间显著相关 (p<0.01)。

2.2 FLD 法对葡萄酒原产地的识别

基于 15 种稀土元素, 建立三个产地酒样的 Fisher 判别模型。根据 Wilks' Lambda 统计量, 当被加入的变量的 F 值 ≥ 3.84 时, 该变量进入函数, 当被移出的变量的 F 值 ≤ 2.71 时, 该变量被移出函数。最终筛选出对葡萄酒原产地识别有贡献的 10 个变量 (Y、La、Ce、Pr、Nd、Sm、Gd、Dy、Ho、Yb), 剔除无效的 5 个变量 (Eu、Tb、Er、Tm、Lu)。模型建立过程中, 提取了两个典型判别函数 (discriminant function, DF) DF1 和 DF2, 方差贡献率分别为 79.4% 和 20.6%。

酒样在 DF1 和 DF2 上的散点分布如图 1a 所示, 三个产地样品分布于不同的空间, 贺兰山东麓酒样在 DF2 上和其他两个产地地区明显, 而沙城和通化产地在 DF1 上可以完全区分开来。此外, 通化酒样的聚合度较好, 没有样品与其他产地交叉; 而沙城和贺兰山东麓酒样分布较为离散, 有少量样品与其他产地交叉, 说明这两个产地酒样内部之间存在差异, 可能因为稀土元素含量原本就有较大的组内差异, 也可能因为 boxcox 数据变换增加了这两个产地组内样品间的差异。

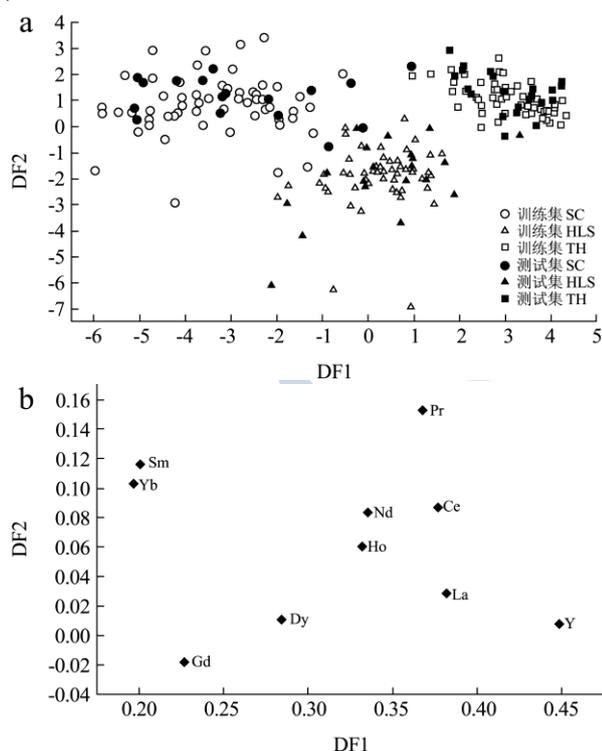


图 1 (a) 酒样前两个典型判别函数的得分散点图; (b) 稀土元素前两个典型判别函数的载荷系数图

Fig.1 (a) Score plot of the first two canonical discriminant functions (DF1 and DF2) of wine samples from different origins (b) Loading plot of the first two canonical discriminant functions (DF1 and DF2) of rare earth elements in wine samples from different origins

由稀土元素在 DF1 和 DF2 上的载荷分布 (图 1b) 所示, DF1 较好地解释了变量 Y、La、Ce 的信息, 而 DF2 仅对变量 Pr 的信息解释较多, 其他元素与 DF1 和 DF2 的关系则较弱。即基于 Y、La、Ce、Pr 四种稀土元素, 可以大致区分沙城、贺兰山东麓、通化三个原产地的酒样, 这与元素含量差异性分析和相关性分析结果较为一致。

经过内部交叉验证, 判别分析结果如表 4 所示, 样本总体判别正确率为 96.49%。沙城样本中有 3 个被错判为贺兰山葡萄酒, 1 个被错判为通化山葡萄酒, 产地判别正确率为 92.98%; 贺兰山样本中有 1 个被错判为通化山葡萄酒, 产地判别正确率为 98.25%; 通化样本中无错判样品, 产地判别正确率为 100%。将 Fisher 判别模型应用于测试集未知样品的产地识别 (图 1 和表 4), 整体判别率为 91.23%。沙城有 3 个样本错判, 判别率为 84.21%; 贺兰山东麓有 2 个样品错判, 判别率为 89.47%; 通化判别率为 100%。总的来说, 模型对沙城酒样识别率最低, 对通化识别率最高, 对贺兰山东麓的识别率居中, 均满足判别分析的要求。

表 4 葡萄酒原产地 FLD 分析结果

Table 4 FLD classification results of wines from different origins

| | | 判属产地 | | | 合计 | |
|----------|----|------|-------|-------|--------|-------|
| | | SC | HLS | TH | | |
| 交叉 验证 | 计数 | SC | 53 | 3 | 1 | 57 |
| | | HLS | 0 | 56 | 1 | 57 |
| | | TH | 0 | 0 | 57 | 57 |
| 判别率/% | | | 92.98 | 98.25 | 100.00 | 96.49 |
| 外部 验证 | 计数 | SC | 16 | 2 | 1 | 19 |
| | | HLS | 0 | 17 | 2 | 19 |
| | | TH | 0 | 0 | 19 | 19 |
| 判别率/% | | | 84.21 | 89.47 | 100.00 | 91.23 |

2.3 PLS-DA 法对葡萄酒原产地的判别分析

利用训练集中 171 个酒样的稀土元素含量建立原产地 PLS-DA 校正模型, 并对 57 个测试集样本进行预判。基于交叉验证 PRESS 值确定模型提取了 7 个主成分, 对自变量的累积方差贡献率为 99.17%, 对因变量信息的综合解释为 87.47%, 模型包含了大多数的自变量和因变量信息。基于样本前三个主成分的得分作图 (图 2), 三个产地有明显的聚类现象, 大部分样品被正确区分, 贺兰山东麓酒样分布范围较广, 与其他两个产地均有交叉现象。

PLS-DA 模型判别分析结果如图 3 和表 5 所示, 交叉验证 r 为 0.92, RMSEC 为 0.30, 整体判别率为

90.64%，模型拟合性较好；外部测试集验证 r 为 0.88，RMSEP 为 0.35，判别率为 87.93%，模型拟合性稍差。当模型应用于沙城酒样判别时，交叉验证判别率为 82.46%，外部验证判别率仅为 73.68%，样本预测值离散程度大，判别效果较差；将模型应用于贺兰山东麓和通化酒样原产地判别时，内部验证和外部验证判别率均大于 80%，判别效果较为理想。

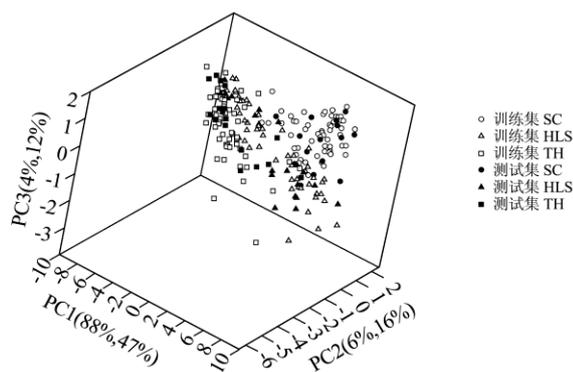


图 2 酒样前三个主成分的得分散点图

Fig.2 Score plot of the first three principle components (PC1, PC2, and PC3) of wine samples from different regions

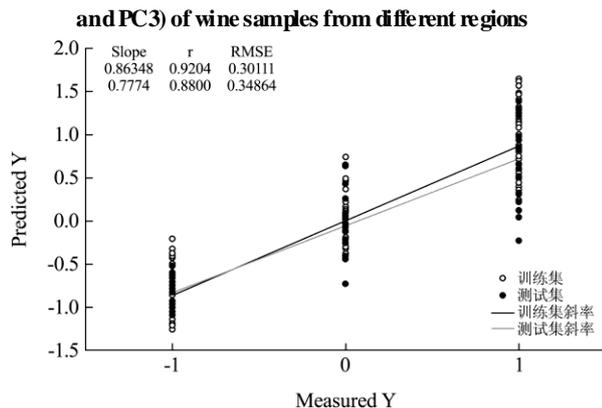


图 3 PLS-DA 校正模型样本分类变量的预测值与参考值

Fig.3 Predicted vs. measured values of categorical variables for calibration samples using PLS-DA

表 5 葡萄酒原产地 PLS-DA 分析结果

Table 5 PLS-DA classification results of wines from different origins

| | | 判属产地 | | | 合计 |
|----------|-------|-------|--------|--------|-------|
| | | SC | HLS | TH | |
| 交叉 验证 | 计数 | SC 47 | HLS 10 | TH 0 | 57 |
| | | HLS 1 | 56 | 0 | 57 |
| | | TH 0 | 5 | 52 | 57 |
| | 判别率/% | 82.46 | 98.25 | 91.23 | 90.64 |
| 外部 验证 | 计数 | SC 14 | HLS 5 | TH 0 | 19 |
| | | HLS 2 | 16 | 1 | 19 |
| | | TH 0 | 0 | 19 | 19 |
| | 判别率/% | 73.68 | 84.21 | 100.00 | 87.93 |

3 结论

3.1 沙城、贺兰山东麓、通化葡萄酒中稀土元素含量存在较大差异，经过FLD，三个原产地酒样明显区分开来，Y、La、Ce、Pr四种稀土元素对典型判别函数的提取贡献更大。模型交叉验证判别率为96.49%，外部验证判别率为91.23%，对每个原产地的判别率均大于80%，满足判别分析的要求。同时，本研究所用样本量较大，具有一定的代表性，分析过程中除了样本内部交叉验证外，还使用未知样本进行了外部验证，模型预测能力和稳健性均较好，可以应用于实践中葡萄酒三个原产地的识别。PLS-DA模型的判别能力稍低，对贺兰山东麓和通化酒样的判别率满足判别分析要求，而对沙城外部验证的判别率仅为73.68%，正确率较低，产地易错判为贺兰山东麓。

3.2 利用稀土元素的指纹分析进行葡萄酒的产地溯源，虽不易受到施肥、农药、气候等因素的影响，但可能会受到酿造工艺的干扰，尤其皂土的使用，会增加或者降低葡萄酒中稀土元素的含量，从而影响判别结果。此外，研究中样本量虽较大，但年份范围较局限，可以通过增加其他年份的葡萄酒样本，扩大模型的应用范围。

参考文献

- [1] GB/T 19265-2008, 地理标志产品 沙城葡萄酒[S]
GB/T 19265-2008, Product of geographical indication-shacheng wine [S]
- [2] GB/T 19504-2008, 地理标志产品贺兰山东麓葡萄酒[S]
GB/T 19504-2008, Product of geographical indication-wine in helan mountain east region [S]
- [3] GB/T 20820-2007, 地理标志产品通化山葡萄酒[S]
GB/T 20820-2007, Product of geographical indication-tonghua amur-wine [S]
- [4] Geana EI, Marinescu A, Iordache AM, et al. Differentiation of romanian wines on geographical origin and wine variety by elemental composition and phenolic components [J/OL]. Food Analytical Methods. 2014 April. <http://link.springer.com/article/10.1007/s12161-014-9846-2>
- [5] Serrano-Lourido D, Saurina J, Hernandez-Cassou S, et al. Classification and characterisation of spanish red wines according to their appellation of origin based on chromatographic profiles and chemometric data analysis [J]. Food Chemistry, 2012, 135(3): 1425-1431

- [6] Cynkar W, Damberg R, Smith P, et al. Classification of tempranillo wines according to geographic origin: combination of mass spectrometry based electronic nose and chemometrics [J]. *Anal. Chim. Acta.*, 2010, 660(1-2): 227-231
- [7] Soufleros EH, Bouloumpasi E, Tsarchopoulos C, et al. Primary amino acid profiles of greek white wines and their use in classification according to variety, origin and vintage [J]. *Food Chemistry*, 2003, 80(2): 261-273
- [8] Bejjani J, Balaban M, Rizk T. A sharper characterization of the geographical origin of Lebanese wines by a new interpretation of the hydrogen isotope ratios of ethanol [J]. *Food Chemistry*, 2014, 165: 134-139
- [9] Pérez-Trujill JP, Barbaste M, Medina B. Chemometric study of bottled wines with denomination of origin from the canary islands (spain) based on ultra-trace elemental content determined by ICP-MS [J]. *Analytical Letters*, 2003, 36(3): 679-697
- [10] Boschetti W, Rampazzo RT, Dessuy MB, et al. Detection of the origin of Brazilian wines based on the determination of only four elements using high-resolution continuum source flame AAS [J]. *Talanta*, 2013, 11: 147-155
- [11] Coetzee PP, Vanhaecke F. Classifying wine according to geographical origin via quadrupole-based ICP-mass spectrometry measurements of boron isotope ratios [J]. *Anal. Bioanal. Chem.*, 2005, 383(6): 977-984
- [12] 林昕,黎其万,和丽忠,等.基于稀土元素指纹分析判别普洱古树茶和台地茶的研究[J].*现代食品科技*,2013,29(12): 2921-2926
LIN Xi, LI Qi-wan, HE Li-zhong, et al. Application of heavy rare earth element fingerprints in discrimination of Pu 'er old plant tea and tableland tea [J]. *Modern Food Science and Technology*, 2013, 29(12): 2921-2926
- [13] Roggo Y, Duponchel L, Huvenne JP. Comparison of supervised pattern recognition methods with McNemar's statistical test application to qualitative analysis of sugar beet by near-infrared spectroscopy [J]. *Analytica Chimica Acta*, 2003, 477: 187-200
- [14] Capron X, Smeyers-Verbeke J, Massart D. Multivariate determination of the geographical origin of wines from four different countries [J]. *Food Chemistry*, 2007, 101(4): 1585-1597