

采用精确质量四极杆-飞行时间液质联用仪及分子特征提取、数据挖掘和主成分分析进行食物品种区分

应用简报

食品

作者

Alyson E. Mitchell, Jihyun Lee
食品科学与技术系,
加州大学戴维斯分校
One Shields Avenue
Davis, CA 95616

Jerry Zweigenbaum
安捷伦科技有限公司
2850 Centerville Road
Wilmington, DE 19808

摘要

洋葱是重要的膳食黄酮类化合物槲皮素的主要来源。我们使用 Agilent MassHunter PCDL Manager 软件、公共文献和在线数据库如 Phenol-Explorer 和 ChemSpider, 建立了一个精确质量洋葱黄酮类化合物数据库。为了与黄酮类化合物数据库进行匹配, 采用 Agilent MassHunter 定性分析软件检索 7 种洋葱提取物的四极杆-飞行时间液质联用仪 (Q-TOF LC/MS) 的数据。初步鉴定了 19 种黄酮类化合物。根据初步鉴定的黄酮类化合物, 采用 Agilent Mass Profiler Professional 软件对洋葱的种类和颜色差异进行主成分分析 (PCA)。PCA 的结果可以区分洋葱的品种和颜色。为了便于比较, 对可能存在的未知化合物进行了 PCA 分析, 这些化合物由分子特征提取所确定, 代表了洋葱的品种独特性。虽然该 PCA 分析获得了类似的结果, 但前 10 种初步鉴定的非花青素黄酮类化合物与洋葱的颜色和品种具有较强的相关性。



Agilent Technologies

前言

本研究提供了区分洋葱种类的精确质量四极杆-飞行时间液质联用方法研究的补充报告，发表于 ACS 研讨会系列刊物 *Physical Methods in Food Analysis* 上 [1]。

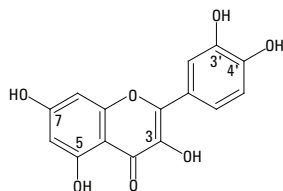
水果和蔬菜是维生素、矿物质、纤维素和大量非必须营养性植物化学成分如多酚类抗氧化剂、黄酮类化合物、类胡萝卜素、生物碱和芥子油苷的主要食物来源。这些化合物多数都有很强的生物活性。物种和品种决定了生物活性植物化学成分的范围。生物活性植物化学成分具有多种化学结构。黄酮类化合物是植物来源多酚类生物活性化合物中含量最多的次生代谢产物，可以进行多种化学修饰，如酰化、丙二酰化、硫酸化、甲基化和糖基化等。其配糖体可以是单、双和三糖取代。食用水果和蔬菜之所以有益健康，公认是因为这类生物活性植物化学成分具有协同作用。

在美国，槲皮素这种重要膳食黄酮类化合物的主要食物来源为洋葱（鲜重为 210 mg/kg）和苹果（鲜重为 30 mg/kg）。有趣的是，大多数在人类和细胞培养中得到的槲皮素衍生物在食物中却未得到发现（图 1）。

遗传学是影响食物中黄酮类化合物补充的主要因素。然而，其含量水平也会受农艺和环境因素（如土壤改良剂、生长区域和紫外线照射量）的影响。

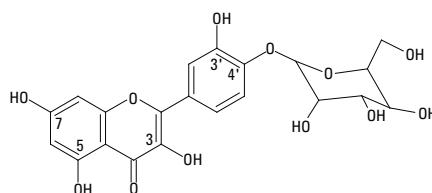
目前，对品种、生长季节和区域、加工、存贮及包装等如何影响生物活性化学物质的组成都知之甚少。同样地，这些因素对食物营养价值的作用也不是很清楚。这种认识上的欠缺，使这些食物在医药或个体化营养的使用方面，或者相同生物活性含量的食物基质产品的生产，都具有很大难度。

本文以洋葱为例，探讨如何采用精确质量四极杆-飞行时间液质联用，以及 Agilent MassHunter 个人化合物数据库和谱库 (PCDL) 管理器、MassHunter 定性分析和 MassHunter Mass Profiler Professional 软件，在无需标准对照的情况下，使用非目标和未知化合物分析方法，来鉴定种类差异。



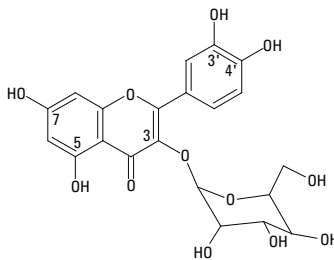
槲皮素糖苷配基 (I)

在细胞培养研究和临床试验中使用，并且可作为一种配料出售。



槲皮素-4'-O-β-D-葡萄糖苷 (II)

洋葱中存在的主要形式。



槲皮素-3-O-β-D-半乳糖苷 (III)

苹果中存在的主要形式。

图 1. 槲皮素存在形式的实例

实验部分

实验过程的详细描述可参见 ACS 研讨会系列刊物 *Physical Methods in Food Analysis* 上发表的补充报告 [1]。

样品前处理

本文评价了来自美国加利福尼亚州奥克斯纳德 Gills Onion 公司的四种黄色洋葱 (Cowboy、Chief、Vaquero 和 Sommerset)，以及三种红色洋葱 (Red Rock、Salsa 和 Merenge)。

使用常用提取步骤对没有特定针对的化合物进行分析。由于洋葱内层只含有很少的花青素（红色），因此将它们与外层剥离并冻干，再用 80% 的甲醇提取 20 min。提取条件不考虑花青素在辨别红色与黄色洋葱品种中的作用。重复三次提取操作。

仪器

采用 Agilent 1290 Infinity 液相色谱系统和配备安捷伦喷射流技术电喷雾离子化的 Agilent 6530 精确质量四极杆飞行时间 (Q-TOF) 液质联用系统，对提取样品进行分析。1290 Infinity 液相色谱系统配备有一个集成了真空脱气机 (G4220A) 的二元泵、带恒温器 (G1330B) 的自动进样器 (G4226A)，以及一个柱温箱 (G1316C)。1290 Infinity 液相色谱系统参数见表 1。

表 1. Agilent 1290 Infinity 液相色谱系统参数

仪器	Agilent 1290 Infinity 液相色谱系统														
流动相	(A) 0.1% 甲酸水溶液, 初始 5% B (B) 0.1% 甲酸的乙腈溶液														
梯度程序	线性 <table border="1"> <thead> <tr> <th>时间 (min)</th> <th>%B</th> </tr> </thead> <tbody> <tr> <td>0-5</td> <td>5-10</td> </tr> <tr> <td>5-8</td> <td>10-12</td> </tr> <tr> <td>8-10</td> <td>12-15</td> </tr> <tr> <td>10-15</td> <td>15</td> </tr> <tr> <td>15-18</td> <td>15-55</td> </tr> <tr> <td>18-20</td> <td>55-90</td> </tr> </tbody> </table>	时间 (min)	%B	0-5	5-10	5-8	10-12	8-10	12-15	10-15	15	15-18	15-55	18-20	55-90
时间 (min)	%B														
0-5	5-10														
5-8	10-12														
8-10	12-15														
10-15	15														
15-18	15-55														
18-20	55-90														
流速	0.4 mL/min														
色谱柱	Agilent Poroshell 120 EC-C18, 2.1 × 100 mm, 2.7 μm (部件号 695775-902)														
后运行时间	初始流动相时 4 min														
温度	30 °C														
进样量	5 μL														

对样品进行四极杆-飞行时间液质联用分析。为了鉴定出所有可能存在的黄酮类化合物，需要将仪器方法制订得更为宽泛。在负离子和正离子两种模式下，采集质量范围 m/z 100–1000 内的总离子流谱图。除了那些极性相对较低的化合物，电喷雾离子化通常可以电离很多种化合物。虽然黄酮类化合物在预期中具有很好的响应，但是它们的相对响应强度受其自身化学结构的影响非常大（例如，共轭位置、糖的种类和酰化作用等）。

一些化合物只在正离子模式下有响应，而另一部分仅在负离子模式下有响应。那些在两种模式下均有响应的化合物，在其中一种模式下的离子化效率可能高于另一种模式。四极杆-飞行时间质谱仪 (Q-TOF MS) 参数见表 2。

表 2. Q-TOF MS 参数

仪器	Agilent 6530 精确质量 Q-TOF LC/MS
离子化模式	正离子和负离子电喷雾，配备安捷伦喷射流离子聚焦技术
采集速率	1.0 张谱图/秒
质量范围	100–1000 m/z
干燥气温度	225 °C
干燥气流速	8.0 L/min
鞘气温度	300 °C
鞘气流速	10.0 L/min
雾化气压力	45 psi
截取锥电压	65 V
八极杆 RF	750 V
碎裂电压	125 V
毛细管电压	2.5 kV

在负离子模式下，连续进样分析参照离子 m/z 119.0363（去氢嘌呤）和 966.0007（甲酸盐加合的六（ ^1H , ^1H , ^3H -四氟丙氧基）磷嗪或 HP-921）来校正每张谱图，以实现优于 2 ppm 的精确质量测定。

数据分析

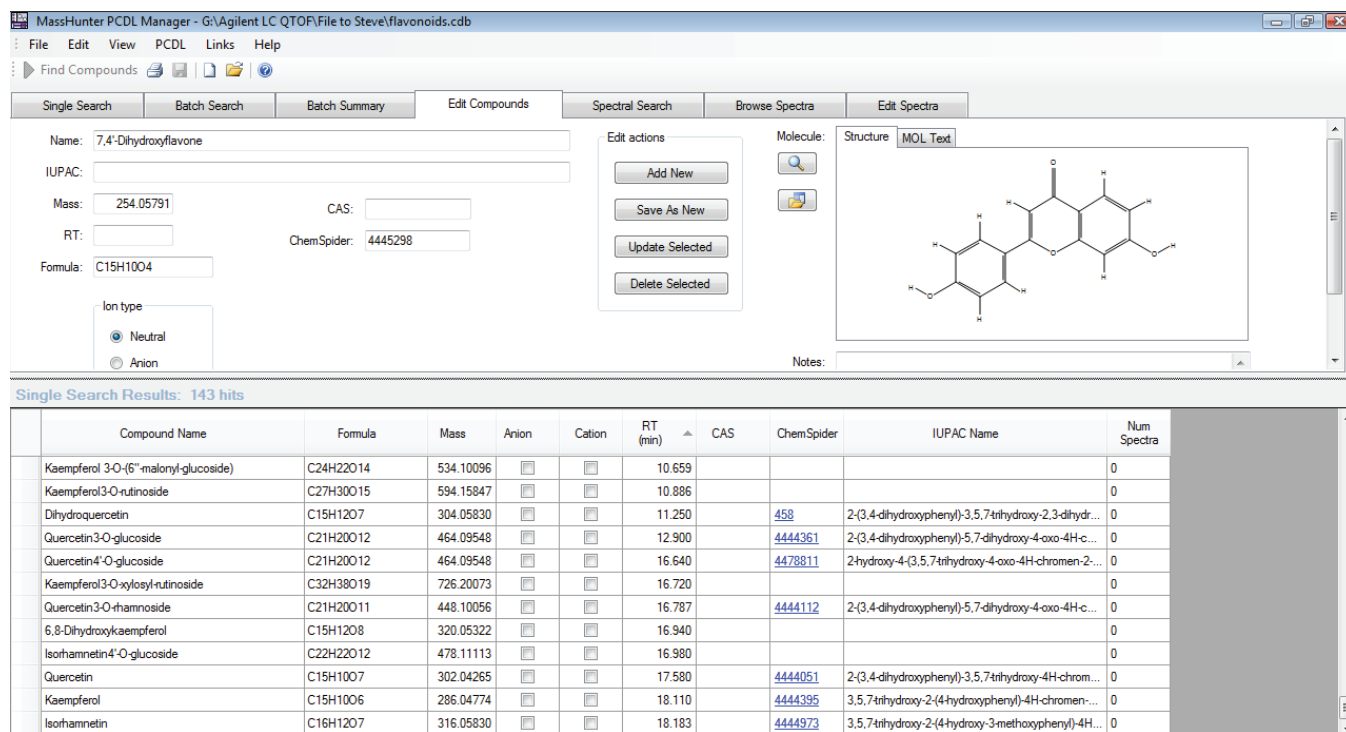
按照如下四个步骤进行数据分析：

1. 创建自定义黄酮类化合物数据库用于洋葱分析
2. 检索从四极杆-飞行时间液质联用单独进行质谱分析所获得的数据，与黄酮类化合物数据库进行匹配并记录匹配得分
3. 根据化合物的匹配结果（非目标性结果，即不使用标准品进行鉴定或确认），进行主成分分析 (PCA)，以区分洋葱的不同品种和颜色
4. 为了进行比较，还对不同种类洋葱中所有可能存在的化合物（未知的，无需匹配）进行 PCA 分析

数据库创建

使用公共文献，包括数据库如 Phenol-Explorer (www.phenol-explorer.eu/) 和 ChemSpider (www.chemspider.com)，鉴定了 150 种可能存在的黄酮类化合物，包括洋葱中可能存在的共轭结构黄酮。将每个黄酮的分子式输入 MassHunter PCDL Manager 软件（图 2）来创建数据库。MassHunter PCDL Manager 软件可以利用输入的分子式，自动计算精确质量。然后在 MassHunter 定性分析软件中使用该数据库来检索每个样品的数据文件，以便与这些化合物进行匹配。

MassHunter PCDL Manager 软件允许输入化合物名称、IUPAC 名称、相关注释、链接、结构，以及其他有用的信息。假如结构是 .mol 格式，软件便会利用导入的 .mol 文件计算分子式和精确质量。MassHunter PCDL Manager 软件的数据库组件包含的是分子而不是离子。对于数据库中的任何化合物，均可以向谱库组件中添加精确质量串联质谱图。可以使用 MassHunter PCDL Manager 软件进行简单的检索，但是要检索完整的数据文件就需要使用 MassHunter 定性分析软件。



The screenshot displays the MassHunter PCDL Manager software interface. The top window shows the 'Single Search' tab with the following input fields: Name: 7,4'-Dihydroxyflavone, IUPAC: (empty), Mass: 254.05791, CAS: (empty), RT: (empty), ChemSpider: 4445298, Formula: C15H10O4, and Ion type: Neutral (selected). The 'Edit actions' panel includes buttons for 'Add New', 'Save As New', 'Update Selected', and 'Delete Selected'. The 'Molecule' section shows the chemical structure of 7,4'-Dihydroxyflavone. Below the search input is a table titled 'Single Search Results: 143 hits'.

Compound Name	Formula	Mass	Anion	Cation	RT (min)	CAS	ChemSpider	IUPAC Name	Num Spectra
Kaempferol 3-O-(6"-malonyl-glucoside)	C24H22O14	534.10096	<input type="checkbox"/>	<input type="checkbox"/>	10.659				0
Kaempferol 3-O-rutinoside	C27H30O15	594.15847	<input type="checkbox"/>	<input type="checkbox"/>	10.886				0
Dihydroquercetin	C15H12O7	304.05830	<input type="checkbox"/>	<input type="checkbox"/>	11.250		458	2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxy-2,3-dihydro...	0
Quercetin 3-O-glucoside	C21H20O12	464.09548	<input type="checkbox"/>	<input type="checkbox"/>	12.900		4444361	2-(3,4-dihydroxyphenyl)-5,7-dihydroxy-4-oxo-4H-c...	0
Quercetin 4'-O-glucoside	C21H20O12	464.09548	<input type="checkbox"/>	<input type="checkbox"/>	16.640		4478811	2-hydroxy-4-(3,5,7-trihydroxy-4-oxo-4H-chromen-2-...	0
Kaempferol 3-O-xylosyl-rutinoside	C32H38O19	726.20073	<input type="checkbox"/>	<input type="checkbox"/>	16.720				0
Quercetin 3-O-rhamnoside	C21H20O11	448.10056	<input type="checkbox"/>	<input type="checkbox"/>	16.787		4444112	2-(3,4-dihydroxyphenyl)-5,7-dihydroxy-4-oxo-4H-c...	0
6,8-Dihydroxykaempferol	C15H12O8	320.05322	<input type="checkbox"/>	<input type="checkbox"/>	16.940				0
Isorhamnetin 4'-O-glucoside	C22H22O12	478.11113	<input type="checkbox"/>	<input type="checkbox"/>	16.980				0
Quercetin	C15H10O7	302.04265	<input type="checkbox"/>	<input type="checkbox"/>	17.580		4444051	2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxy-4H-chrom...	0
Kaempferol	C15H10O6	286.04774	<input type="checkbox"/>	<input type="checkbox"/>	18.110		4444395	3,5,7-trihydroxy-2-(4-hydroxyphenyl)-4H-chromen-...	0
Isorhamnetin	C16H12O7	316.05830	<input type="checkbox"/>	<input type="checkbox"/>	18.183		4444973	3,5,7-trihydroxy-2-(4-hydroxy-3-methoxyphenyl)-4H...	0

图 2. 使用 Agilent MassHunter PCDL Manager 软件建立用于洋葱分析的精确质量数据库

数据库检索

使用 MassHunter 定性分析软件 (图 3), 在数据分析步骤 1 创建的 PCDL 中检索由四极杆-飞行时间液质联用仪分析得到的数据。特别是, 通过软件的分子式查找功能, 我们可以设定精确质量的允差, 然后将它应用到数据文件的检索中, 以搜寻目标的同位素离子、加合物 (如加 H⁺ 和加 Na⁺)、二聚体和三聚体等。

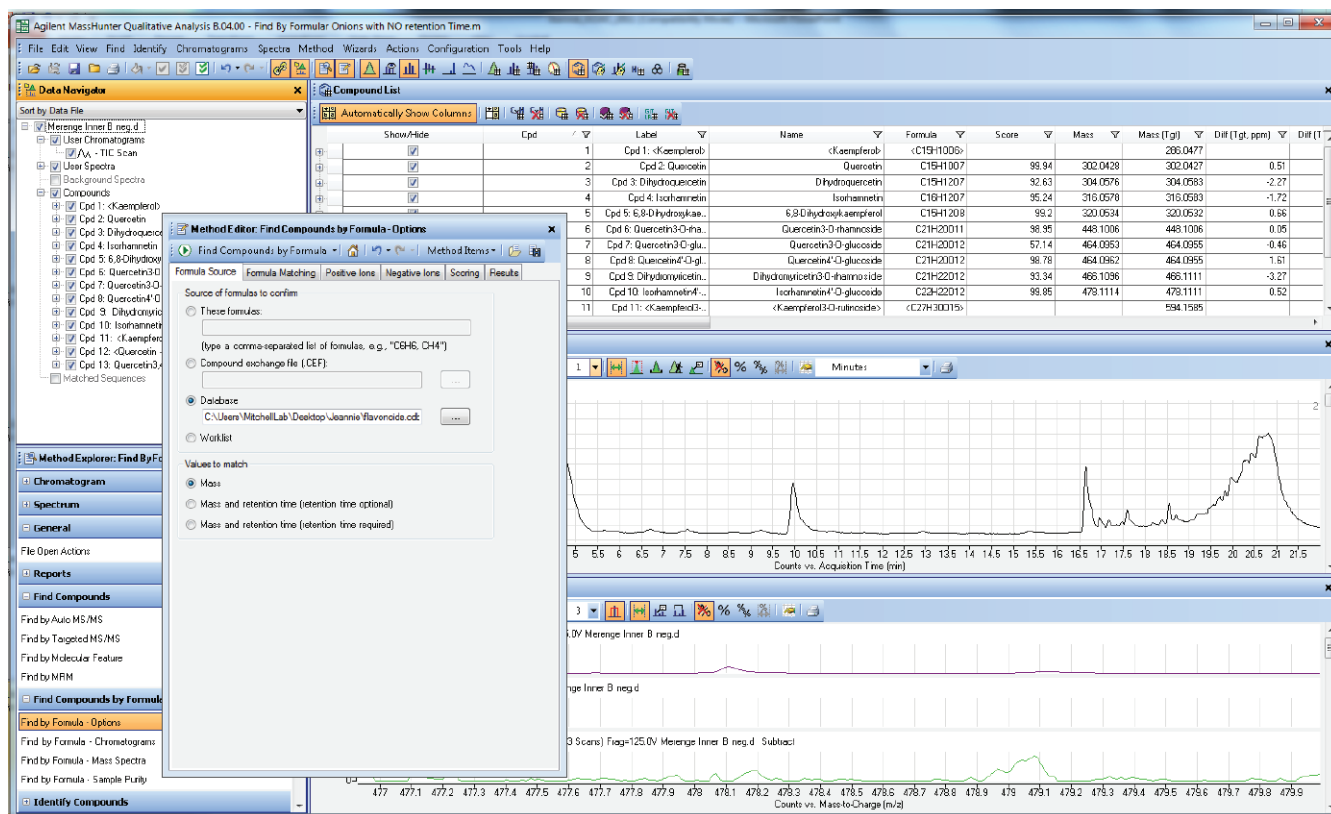


图 3. 使用 Agilent MassHunter 定性分析软件对数据进行检索, 以便与黄酮类化合物数据库中的分子式进行匹配

根据分子式的精确质量和预期的分子式同位素模式，可评价每个实验测定的离子质量。利用该数据获得此离子匹配数据库的得分(图4)。这种方法无需使用标准品，只通过软件即可对黄酮类化合物进行初步鉴定。我们称这种方法为非目标化合物的定向检索。

质量、同位素质量和同位素质量间距的精确匹配仅能确定分子式，而不能确定是否是数据库中的实际化合物。当不使用标准品时，可以利用串联质谱提供的更多信息进行化合物鉴定。若标准品可用，其保留时间和串联质谱图也可以加到数据库中，并将其用于分值匹配。需要在样品相同的条件下分析标准品，比较获得的结果以便进行验证。

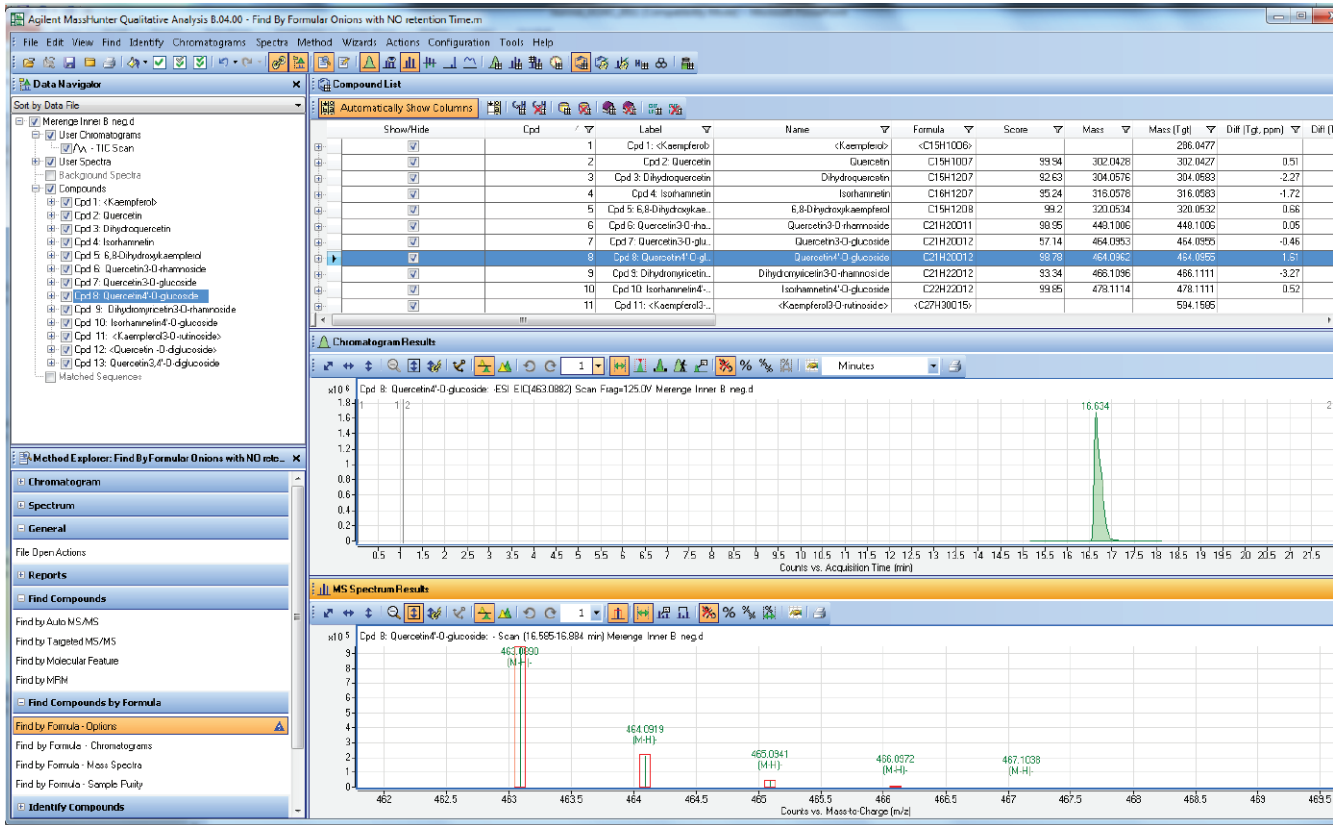


图4. 使用 Agilent MassHunter 定性分析软件对测定质量、同位素丰度和质量间距与其理论计算值进行比较，以生成黄酮类化合物匹配得分

主成分分析

一旦不同品种洋葱中初步鉴定的黄酮类化合物的名单生成以后，可以使用 MassHunter Mass Profiler Professional 软件中的 PCA 分析对它们进行进一步的检验，以区分品种和颜色。进行 PCA 分析有两种方法。这两种方法均被设计用于提供一个化合物列表：一个是从数据库检索出的初步鉴定的化合物列表，另一个是未知的化合物列表。

根据与数据库匹配得到的化合物进行第一次 PCA 分析。第二种方法使用分子特征提取从所有种类洋葱的 TIC 中提取所有可能的化合物。由于它们在一种洋葱中存在而在另一种中不存在，基于此，对发现的未知化合物进行过滤。MassHunter 定性分析软件中的分子特征提取器可以提取代表色谱峰的所有离子（从而消除了本底离子），并且按照加合物种类、可能的同位素、二聚体和三聚体等将它们进行分组，所有这些离子无需任何确证即可形成其分子特征。每个特征随后回算到一个分子质量，仍然不需要任何确证。这种未报道的数据挖掘方法可无需鉴定品种特有的化合物而直接提供种类差异。

结果与讨论

精确质量四极杆-飞行时间液质联用分析

虽然洋葱提取样品的 TIC 看起来特征性不太显著，但它们包含了许多代表多种化合物的离子。Merenge 品种内层的 TIC 见图 5。

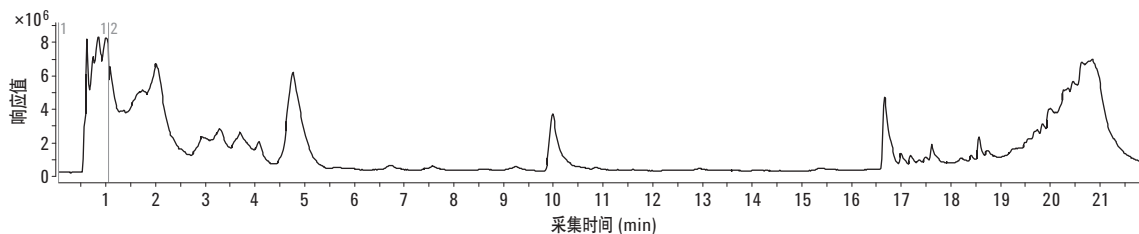


图 5. Merenge 品种的 TIC

数据库检索结果

创建的黄酮类化合物数据库的 MassHunter 定性分析软件通过分子式进行查找，在 7 种洋葱中初步鉴定了 19 种黄酮类化合物（表 3）。

表 3. 通过 Agilent MassHunter 定性分析软件按分子式查找功能，检索创建的数据库初步鉴定的黄酮类化合物

飞燕草素 3-O-(6''-丙二酰基-葡萄糖苷)
二羟基山奈酚
二氢杨梅素 3-O-鼠李糖苷
二氢槲皮素
异鼠李素
异鼠李素 4'-O-葡萄糖苷
山奈酚
山奈酚 3-O-(6''-丙二酰基-葡萄糖苷)
山奈酚 3,7-O-二葡萄糖苷
山奈酚 3-O-乙酰-葡萄糖苷
山奈酚 3-O-芸香糖苷
山奈酚 3-O-木糖-芸香糖苷
槲皮素
槲皮素 3,7,4'-三葡萄糖苷
槲皮素-O-二葡萄糖苷
槲皮素 3,4'-O-二葡萄糖苷
槲皮素 3-O-葡萄糖苷
槲皮素 3-O-鼠李糖苷
槲皮素 4'-O-葡萄糖苷

主成分分析结果

PCA 得分图展现了样品间数据的差异。例如，从这 7 种洋葱的、基于数据库检索初步鉴定的 19 种黄酮类化合物的 PCA 得分图中可以看出，不同种类间分离良好（图 6）。由于在洋葱内层进行分析，且提取溶剂已排除花青素，那么该品种特征并不取决于色素相关化合物。Cowboy 和 Sommerset 组彼此很接近（分离较少），这种情况也存在与 Chief、Vaquero 和 Salsa 之间。红色品种（Red Rock 和 Merenge）与其他品种显著分离。

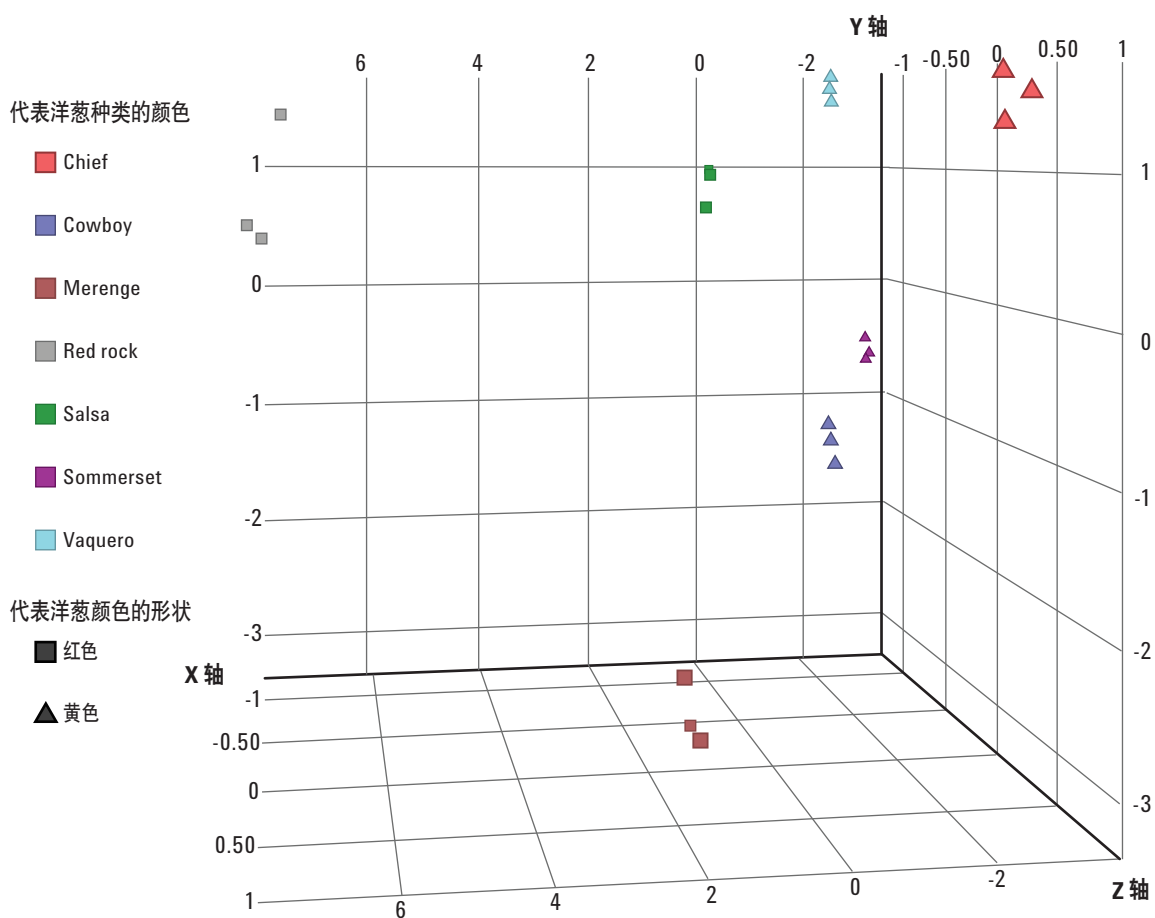
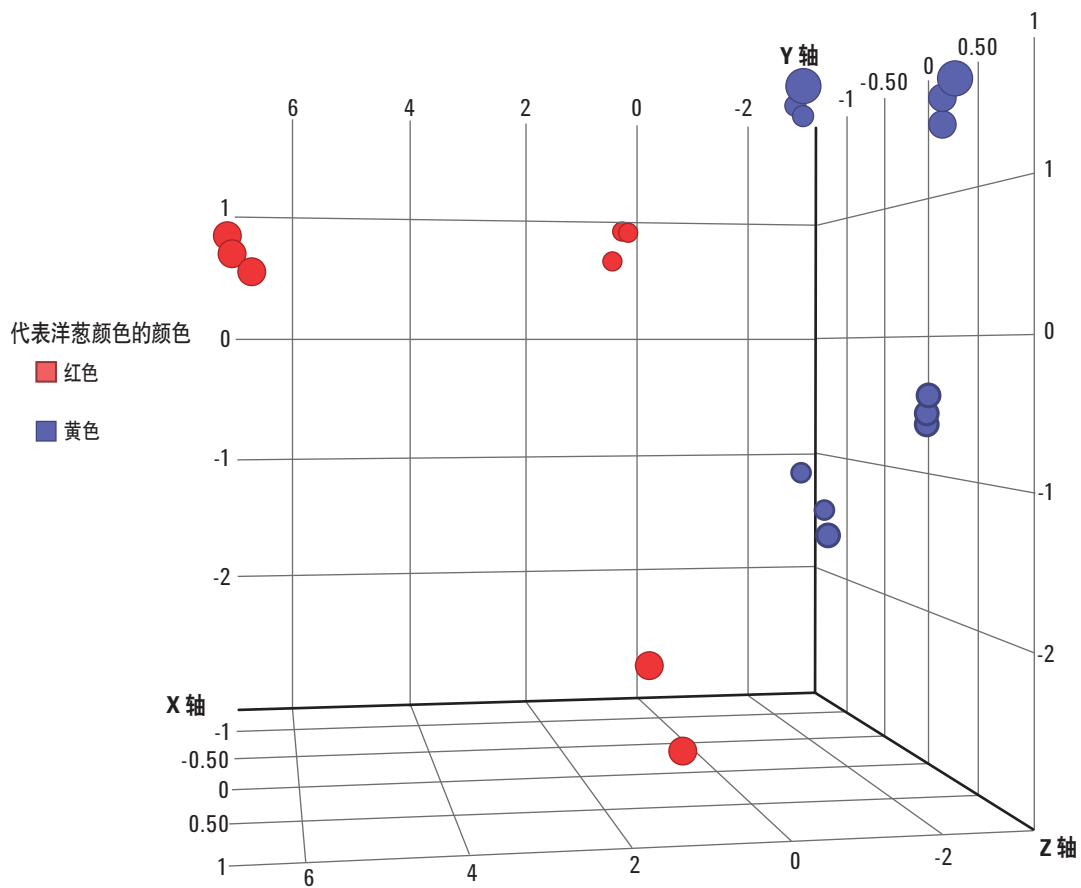


图 6. 根据黄酮类化合物数据库中检索发现的目标未知化合物，进行洋葱种类区分的 PCA 分析

颜色差异的 PCA 得分图 (图 7) 显示, 尽管沿 x 轴有一些相关性, 但种类间距离很短。



在对数据进行分子特征提取后，利用所有未知化合物得到了7种洋葱的PCA得分图（图8），该图表明此方法也可以用来评价不同品种的常见组成和颜色。

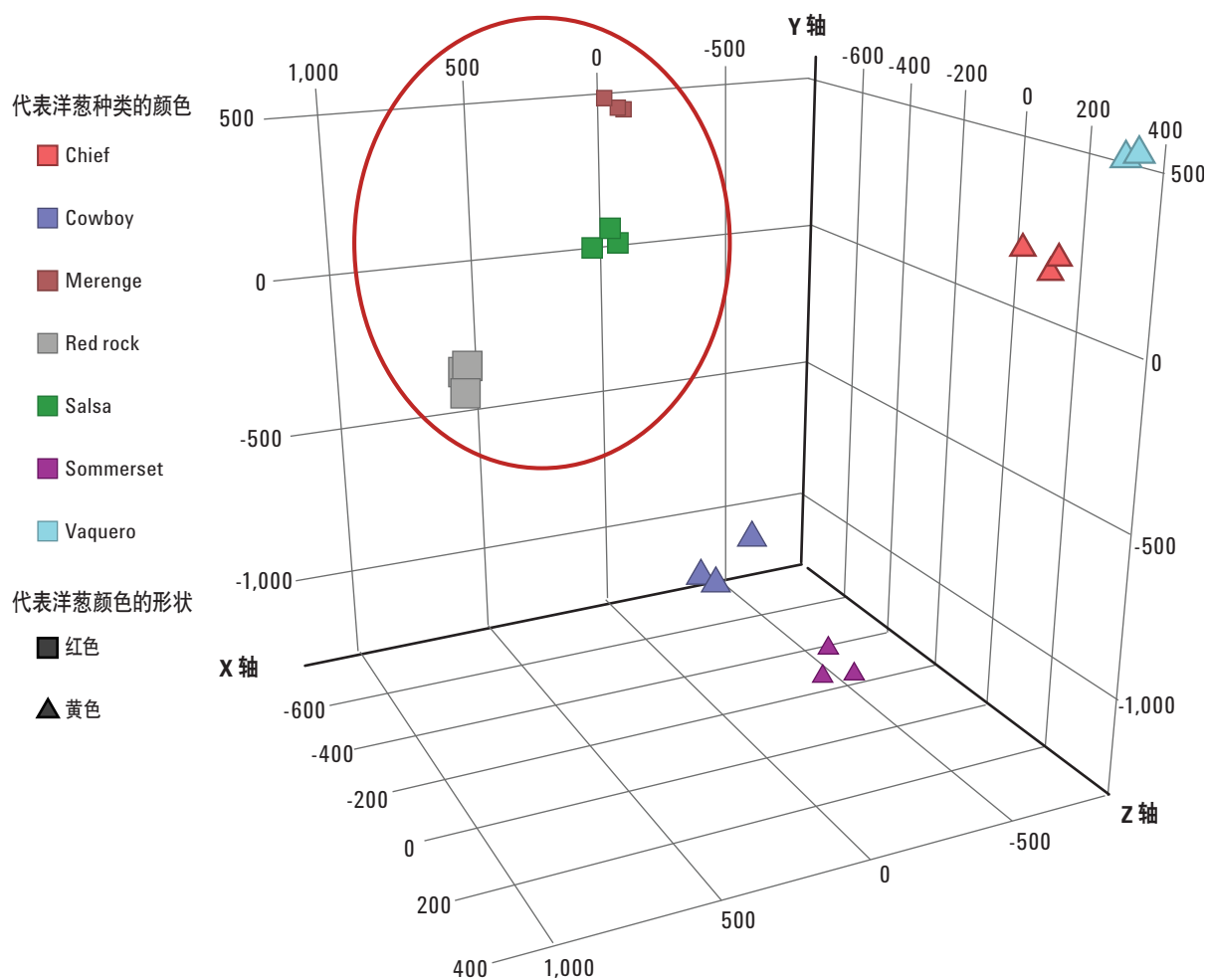


图8. 对数据进行分子特征提取后，利用所有未知化合物得到的PCA得分图

PCA 载荷图展现了在协方差空间发现的能够代表那些样品的实体或化合物。这些载荷是化合物的列表。由于 Chief 和 Merenge 可在 PCA 中得到良好分离，因此可将其用于评价载荷并生成化合物列表，以区分两个品种。

表 4 列出了初步鉴定的可以表征颜色差异的前 10 种化合物。软件发现的每种分子式都对应许多可能的化合物，因此更多的串联质谱信息有助于鉴定出实际存在的化合物，并且可以利用它们来进行区分。

表 4. 前 10 种初步鉴定的非花青素类化合物与颜色差异性相关

6,8-二羟基山奈酚

山奈酚 3-O-(6"-丙二酰基-葡萄糖苷)

山奈酚二葡萄糖苷-1

山奈酚二葡萄糖苷-2 (异构体, 不同的保留时间)

山奈酚 3-O-乙酰-葡萄糖苷槲皮素

槲皮素 3,4'-O-二葡萄糖苷

槲皮素二葡萄糖苷-1

槲皮素二葡萄糖苷-2 (异构体, 不同的保留时间)

槲皮素 3-O-鼠李糖苷

二氢槲皮素

结论

使用公共文献和数据库，以及 Agilent MassHunter PCDL Manager 软件，创建了一个用于定向检索特定精确质量的洋葱黄酮类化合物数据库。利用该数据库，对来自 7 种洋葱提取物精确质量四极杆-飞行时间液质联用仪分析的数据进行定向检索，鉴定了 19 种潜在的黄酮类化合物。基于潜在黄酮类化合物的 PCA 分析，根据品种和颜色的差异进行分离。数据经分子特征提取后，二次 PCA 分析利用所有未知化合物得到了类似的结果。利用二次 PCA 分析，在初步鉴定出的黄酮类化合物中，鉴定出 10 种与品种和颜色关联最强的化合物。这些结果表明，精确质量四极杆-飞行时间液质联用仪与任何一种 PCA 分析方法联用后，都可以用来建立洋葱中黄酮类化合物的品种差异分析。当然，我们还需要对不同环境下生长的更多品种进行深入分析，以建立它们之间明晰的相关性。

致谢

作者衷心感谢 Gills Onions 有限公司（奥克斯纳德，加利福尼亚州）为本工作提供洋葱样品。我们还要感谢加州大学戴维斯分校的 Susan Ebeler 博士和 Tom Collins 博士，以及安捷伦科技的 Steven Fischer 先生在高级数据处理中所提供的非常宝贵的科研投入和帮助。

参考文献

1. J.S.E. Lee, J.A. Zweigenbaum, A.E. Mitchell "Nontargeted Unknown LC(ESI-)-Q/TOF MS Approaches for Food Verification" *Physical Methods in Food Analysis*, M. Tunick, *et al.* ACS Symposium Series; American Chemical Society. Washington, DC. 2013.

更多信息

这些数据仅代表典型的结果。如需了解更多有关我们产品和服务的信息，请访问我们的网站 www.agilent.com/chem/cn。

www.agilent.com/chem/cn

安捷伦对本材料可能存在的错误，或由于提供、展示或使用本资料所造成的间接损失不承担任何责任。

本文中的信息、说明及技术指标如有变更，恕不另行通知。

© 安捷伦科技（中国）有限公司，2013
2013年11月13日，中国印刷
5991-3419CHCN



Agilent Technologies