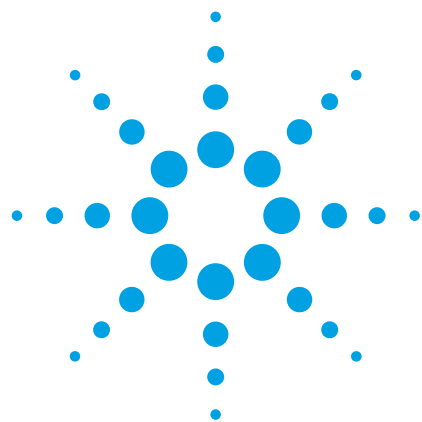




计算机辅助复杂天然产物提取物分析

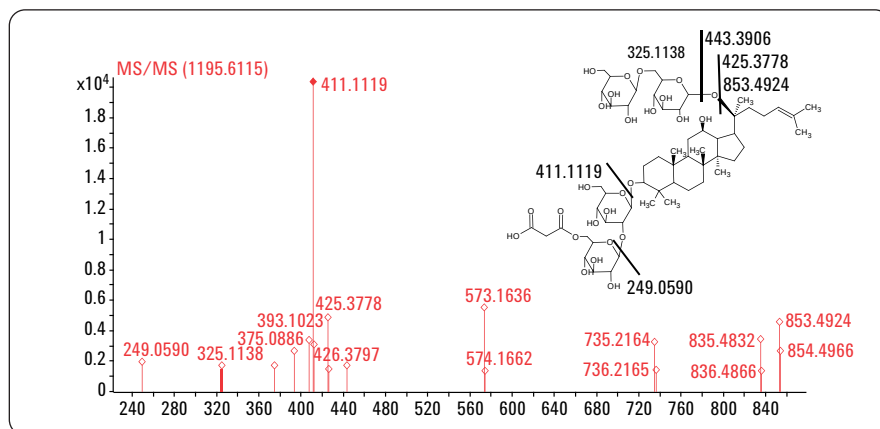
采用 Q-TOF 质谱和 Agilent MassHunter 代谢产物鉴定软件进行已知化合物的检测和未知化合物的鉴定

应用简报

药物发现

作者

Edgar Nägele
安捷伦科技公司
Waldbronn, 德国



摘要

本应用简报介绍了：

- 天然产物提取物中已知化合物和未知新化合物的计算机辅助分析
- Agilent MassHunter 代谢产物鉴定 (MetID) 软件的使用。该软件带有用于在复杂数据中定位化合物的分子特征提取 (MFE) 算法，以及其它几种用于分析天然产物提取物的算法
- 采用 Agilent 1200 系列快速高分离度液相色谱系统 (RRLC) 分离天然产物提取物中的成分
- 采用电喷雾离子化四极杆飞行时间质谱 (Q-TOF) 测定精确分子量
- 对两个人参亚种提取物的数据进行比较分析，寻找只属于某一亚种的化合物，以及在不同亚种间浓度存在差异的化合物



Agilent Technologies

引言

在全球各地的史前古文明中，草药粗提取物就已被用于治疗各种疾病。草药对各种疾病的疗效是经过几百年来反复试验发现的，关于草药的知识也得以代代相传。以草药为基础的传统中医学（TCM）作为这一优化过程的结晶，正是一个很好的例子。

由于这些药物往往是含有成百上千种不同化合物的复杂混合物，且这些化合物具有不同的功效或具有协同作用，因此，很难将药物的疗效归因于其中的某一个化合物。在西方医学中，天然来源的药物由于它们的巨大潜力正变得越来越重要。但是西方药品质量标准要求对基于天然产物的药物成分有深入的了解。在现代工作流程中，计算机辅助鉴定技术可用于植物和草药复杂提取物中已知化合物的鉴定和未知化合物的检测。该工作流程包括复杂的 MS 和 MS/MS 数据的数据库比对，以期找到已知化合物；还包括采用基于计算机的算法进行新化合物的鉴定。

人参根（*Panax* 种属）是一种著名的亚洲草药，已有 5000 多年的药用历史。其主要活性成分人参皂甙属于三萜皂甙类化合物。在过去，通过电喷雾质谱已分离并鉴定了其中的 80 多种人参皂甙¹。复杂天然产物提取物（如人参根提取物）分析的首选方法是高效液相色谱法（HPLC）²。目前采用可进行精确质量数测定、MS/MS 和 MSⁿ 分析进行结构鉴定的现代 LC/MS 系统来确定人参皂甙的复杂且相似的结构³。

本应用简报介绍了一个计算机辅助工作流程，以 Agilent MassHunter 代谢产物鉴定软件为基础，进行复杂天然产物提取物中已知化合物的检测和未知化合物的鉴定。该流程以人参根提取物的各种组分为例，分析基于 Agilent 6520 Accurate-Mass Q-TOF LC/MS 系统的精确质量数测定。

实验部分

仪器设备

- Agilent 1200 系列快速高分离度液相色谱系统（RRLC）。包括带脱气机的 Agilent 1200 系列二元泵 SL、带温控的 Agilent 1200 系列高性能自动进样器 SL、Agilent 1200 系列柱温箱（TCC）和 Agilent 1200 系列二极管阵列检测器 SL（DAD SL）
- Agilent 6520 精确质量 Q-TOF LC/MS 系统。采用双喷雾器接口进行质谱校正
- Agilent Zorbax SB-C18 柱，21 x 150 mm，1.8 μm
- Agilent MassHunter 工作站软件。用于数据采集和数据分析（定性分析软件、MetID 软件）

样品制备

将两种冻干的人参根粉末（各 1g）分别置于 10mL 甲醇中超声提取 30 分钟，过滤，然后直接用于分析。这两种样品分别为：

- 亚洲人参（*Panax ginseng*）。购自 ILHWA 有限公司（韩国）
- 美洲人参（西洋参，*Panax quinquefolius*）购自 Sigma-Aldrich 公司（德国，Taufkirchen）

内标为利血平（C₃₃H₄₀N₂O₉），分子量（MW）为 608.2733，配制为 10 ng/mL 的溶液。天然产物提取液与内标标准溶液以 10:1 混合，利血平终浓度为 1 ng/mL。进样量为 10 μL，利血平的柱上浓度为 10 pg。

实验方法

LC 条件：

- 溶剂 A：水 + 0.1% 甲酸（FA）
- 溶剂 B：乙腈 + 0.1% 甲酸
- 流速：0.5 mL/min
- 梯度：0 min 5% B、1 min 5% B、30 min 95% B
- 停止时间：30 min
- 后运行时间：15 min
- 进样量：10 μL
- 恒温自动进样器温度：4 °C
- 自动延迟体积减少：ON
- 柱温：50 °C
- DAD SL：220 nm/4，参比 360 nm/16
- 流通池：2 μL（10 mm 光程）

Agilent 6520 精确质量 Q-TOF LC/MS 系统以下列条件运行：

- 离子源：电喷雾（ESI）正离子模式，带双喷雾器，用于参比标样校正（121.05087 m/z 和 922.00980 m/z）
- 干燥气流速：12 L/min
- 干燥气温度：200 °C
- 雾化器：60 psi
- 扫描范围：200–1300
- 碎裂电压：150 V
- 锥孔电压：60 V
- 毛细管电压：3000 V
- 数据相关 MS/MS 分析：2 MS 和 2 MS/MS 谱图/秒，每个 MS 谱图包含 2 个化合物以进行 MS/MS 分析，2 个 MS / MS 谱图后进行 0.25 分钟动态排除

数据分析

数据分析的第一步是将亚洲人参 (*Panax ginseng*) 的 Q-TOF 数据文件在 Agilent MassHunter 代谢产物鉴定软件中打开, 采用分子特征提取算法 (MFE) 进行化合物提取。该分子特征提取条件采用低浓度水平的加标化合物利血平进行了优化。经过分子特征提取, 可得到所有 MFE 化合物的提取离子流色谱图 (EIC)。将这些化合物与定制的 Agilent METLIN 个人天然产物代谢物数据库中的化合物进行比较。该数据库包括精确质量数和保留时间。最后, 基于精确质量数计算得到所有化合物的化学式。

数据分析的第二步是更为细致地检查样品中的未知化合物, 以鉴定新的天然产物化合物。首先假设给定的天然样品中的大量天然产物化合物属于某已知化合物的衍生物, 可通过与该已知化合物的比较鉴定出来。由此, 我们采用了一个已知化合物作为参考化合物, 对剩下的化合物就同位素分布、MS/MS 裂解模式和特征质量数变化进行相似性检索。这些检索采用的算法包括在 Agilent MassHunter 代谢产物鉴定软件中。

第三步, 将以前分析的样品与一种同其相关的新样品 (西洋参, *Panax quinquefolius*) 进行比较。经过比较, 新样品中的化合物被分为以下三种:

- 1) 全新的化合物
- 2) 响应增加的化合物
- 3) 响应降低的化合物

接着, 采用上述检索方法对西洋参中的新化合物进行检索, 寻找与亚洲参中给定的参考化合物相关的化合物。

结果与讨论

第 1 步:

已知化合物的鉴定

首先, 采用 Agilent 1200 系列 RRLC 系统通过 Agilent ZORBAX 超高压快速高分离度 (RRHT) 色谱柱对亚洲参 (*Panax ginseng*) 根提取物的各种成分进行分离, 接着使用 Agilent 6520 精确质量 Q-TOF LC/MS 系统测定洗脱化合物的精确质量数。高分辨率的 LC 系统对于天然产物提取物中的高丰度和低丰度组分均可实现出色的分离。得到的结果数据文件采用实验部分描述的数据分析方法, 用 Agilent MassHunter 代谢产物鉴定软件进行分析。该数据库以 Agilent METLIN 个人代谢产物数据库为基础, 包含最重要的人参皂甙

化合物的精确质量数、保留时间 (RT) 以及结构信息。

在图 1 所示的结果中, 可得到 EIC (提取离子流色谱)、ECC (提取化合物色谱)、紫外检测、所有提取出的分子特征的分子量以及它们与数据库的匹配结果等信息。具体到本例中, 报道的化合物 #2 与已知化合物人参皂甙 Rb1 ($C_{54}H_{92}O_{23}$, MW = 1108.6029, RT = 16.447) 相匹配。

以测得的分子离子峰 (m/z 为 1109.6117) 及其同位素分布为基础, 计算得到了分子量, 证实该分子量与数据库中的分子式具有 -1.37 ppm 的非常低的质量相对偏差。此外, 经计算, 本例中的保留时间与数据库中的保留时间相比偏差为 -0.007 分钟。

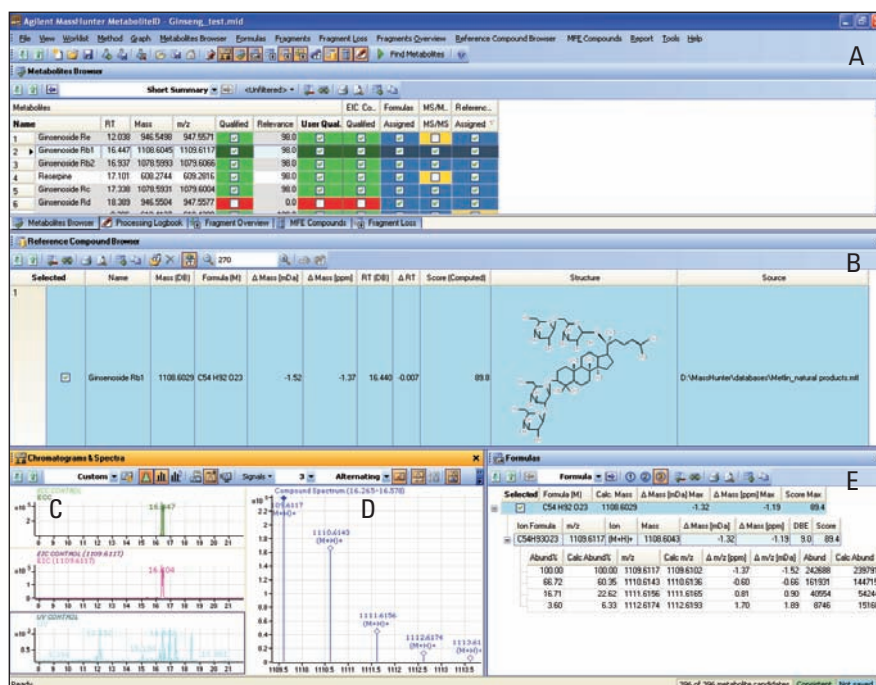


图 1

A) 结果一览表给出了通过 METLIN 数据库检索和其他搜索标准过滤的所有化合物。B) 所选化合物的详细的数据库检索结果, 包括分子式和相对质量偏差、保留时间和保留时间偏差以及结构信息。C) 包括所选化合物的 ECC 和 EIC, 以及 UV 谱图在内的所有的详细色谱图。D) 详细质谱图, 表明所选化合物的同位素分布。E) 选定化合物的计算化学式和质量精度, 包括同位素分布。

第 2 步：未知化合物的详细分析

进行天然产物提取物中的未知化合物分析时，首先假设这些化合物与那些已经鉴定的化合物类似。例如，它们可能与某些已知化合物属于一类化合物，或者是某些已知化合物的衍生物。进行这一比较分析时，已鉴定的化合物人参皂甙 Rb1 被用作参考化合物。

用于比较的第一个特征是各个化合物测得的同位素分布。在这个比较中，将所有测得的同位素分布与参考化合物人参皂甙 Rb1 ($C_{54}H_{93}O_{23}$, $m/z = 1109.6108$, $RT = 16.447$) 的计算同位素分布 (CIP) 进行了比较。Agilent MassHunter 代谢产物鉴定软件中的同位素分布匹配算法会标记超过预设匹配得分值的所有化合物。其中一个化合物的分子量为 1194.6034，保留时间在 16.612 分钟到 16.890 分钟之间 (图 2)。与计算化学式 $C_{57}H_{94}O_{26}$ 匹配得最好，相对质量偏差仅为 0.09ppm

(图 3)。相对于参考化合物，该未知化合物增加了 $C_3H_2O_3$ 部分。

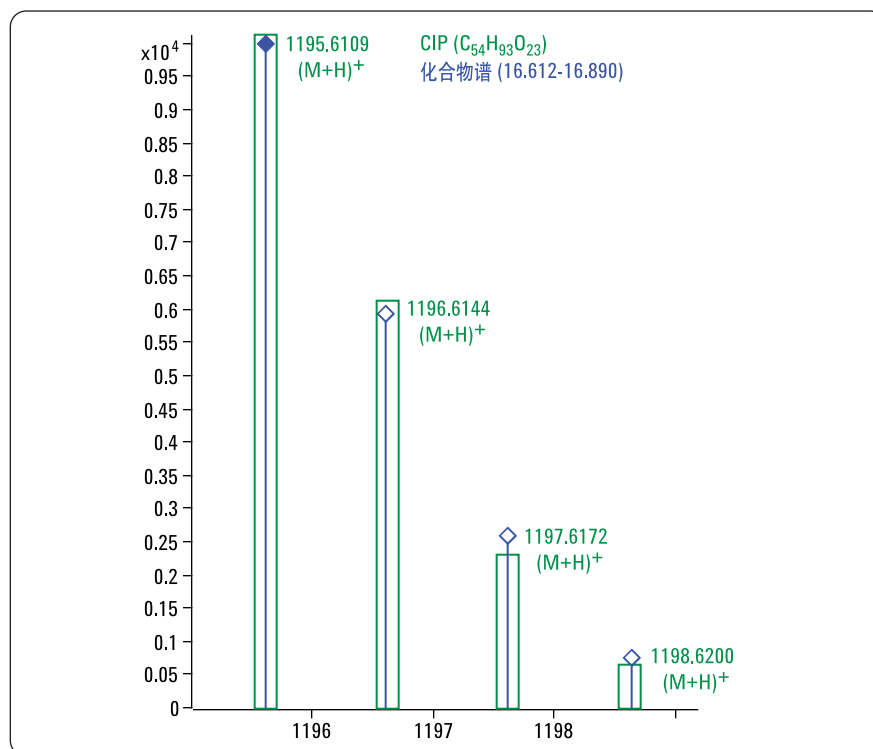


图 2
测得的 m/z 为 1195.6109 的化合物 (RT 16.612 分钟到 16.890 分钟) 的同位素分布与人参皂甙 Rb1 的计算同位素分布 (CIP, 绿色) 的匹配结果。

Formulas									
Formula									
Selected	Formula (M)	Calc. Mass	Δ Mass [mDa] Max		Δ Mass [ppm] Max		Score Max		
☒	C57 H94 O26	1194.6033	-0.10		-0.09		97.6		
	Ion Formula	m/z	Ion	Mass	Δ Mass [mDa]	Δ Mass [ppm]	DBE	Score	
☒	C57H95O26	1195.6109	(M+H)+	1194.6034	-0.10	-0.09	11.0	97.6	
	Abund%	Calc Abund%	m/z	Calc m/z	Δ m/z [ppm]	Δ m/z [mDa]	Abund	Calc Abund	
	100.00	100.00	1195.6109	1195.6106	-0.28	-0.33	10115	9680	
	57.49	63.73	1196.6131	1196.6140	0.80	0.96	5815	6169	
	24.34	25.32	1197.6152	1197.6169	1.42	1.70	2462	2451	
	6.26	7.51	1198.6327	1198.6197	-10.88	-13.05	634	727	

图 3
 m/z 为 1195.6109 的化合物的计算化学式，包括同位素分布和计算出的绝对和相对质量偏差。

进行相似性检索的第二个特征是 MS/MS 谱。出于这一目的，将所有采集的 MS/MS 谱图与选定的参考化合物人参皂甙 Rb1 的 MS/MS 谱进行了比较（图 4）。在该谱图内，发现了相似化合物的典型碎片。在本例中，这些碎片是典型的糖基和甾体化合物的碎片。对于所有的碎片离子，计算了分子式和质量精度（表 1）。

为了实现手动比较，通过碎裂模式匹配算法生成了一个特殊的碎片离子概览图（图 5）。在本概览图中，列出了与人参皂甙 Rb1 MS/MS 谱相关的所有化合物的 MS/MS 碎片。图中显示了碎片离子，x 轴为碎片质量数，y 轴为化合物保留时间。选择与人参皂甙 Rb1 的碎片离子匹配的化合物进行进一步分析。在图中表示为灰色条带。对具有质量数已知、 m/z 为 325 的碎片离子的新化合物也继续进行进一步分析。发现一种保留时间在 16.7 分钟的化合物还包含了质量数已知、 m/z 为 407、425 和 443 的典型碎片。

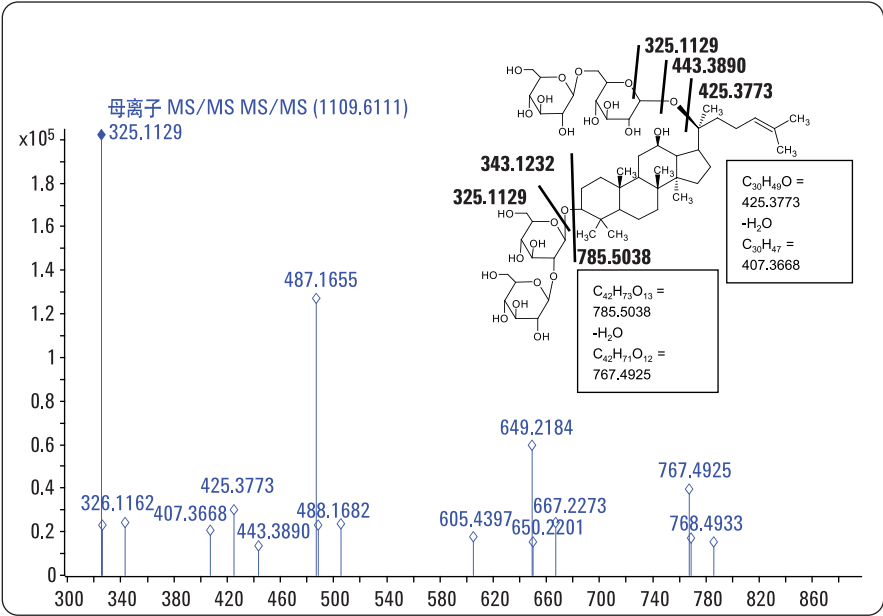


图 4
人参皂甙 Rb1 的 MS/MS 谱图和碎片离子归属。（另见表 1）。

m/z	离子化学式	计算值 m/z	$\Delta m/z$ [mDa]	$\Delta m/z$ [ppm]	中性丢失基团	丢失基团的 化学式	丢失基团的 质量数
163.0596	$C_6H_{11}O_5$	163.06010	0.48	2.92	946.5515	$C_{48}H_{82}O_{18}$	946.5501
325.1129	$C_{12}H_{21}O_{10}$	325.11292	0.01	0.02	784.4982	$C_{42}H_{72}O_{13}$	784.4973
343.1232	$C_{12}H_{23}O_{11}$	343.12349	0.24	0.71	766.4878	$C_{42}H_{70}O_{12}$	766.4867
407.3668	$C_{30}H_{47}$	407.36723	0.45	1.10	702.2443	$C_{24}H_{46}O_{23}$	702.2430
425.3773	$C_{30}H_{49}O$	425.37779	0.51	1.21	684.2338	$C_{24}H_{44}O_{22}$	684.2324
443.3890	$C_{30}H_{51}O_2$	443.38836	-0.69	-1.56	666.2220	$C_{24}H_{42}O_{21}$	666.2219
487.1655	$C_{18}H_{31}O_{15}$	487.16575	0.26	0.54	622.4456	$C_{36}H_{62}O_8$	622.4445
505.1757	$C_{18}H_{33}O_{16}$	505.17631	0.65	1.30	604.4354	$C_{36}H_{60}O_7$	604.4339
605.4397	$C_{36}H_{61}O_7$	605.44118	1.50	2.48	504.1714	$C_{18}H_{32}O_{16}$	504.1690
649.2184	$C_{24}H_{41}O_{20}$	649.21857	0.16	0.24	460.3927	$C_{30}H_{52}O_3$	460.3916
667.2273	$C_{24}H_{43}O_{21}$	667.22913	1.87	2.80	442.3838	$C_{30}H_{50}O_2$	442.3811
767.4925	$C_{42}H_{71}O_{12}$	767.49400	1.53	1.99	342.1186	$C_{12}H_{22}O_{11}$	342.1162
785.5038	$C_{42}H_{73}O_{13}$	785.50457	0.75	0.96	324.1073	$C_{12}H_{20}O_{10}$	324.1056
929.5454	$C_{48}H_{81}O_{17}$	929.54683	1.43	1.54	180.0657	$C_6H_{12}O_6$	180.0634
1109.6118	$C_{54}H_{93}O_{23}$	1109.61020	-1.32	-1.19			

表 1
人参皂甙 Rb1 的 MS/MS 碎片、计算化学式和质量精度。

对该化合物的 MS/MS 谱进行进一步考察时, 通过比较参考化合物人参皂甙 Rb1 的 MS/MS 谱图, 发现 m/z 为 325.1138、407.3666、425.3778 和 443.3906 的碎片可解释为葡萄糖组分和甾体分子母体的碎片离子 (图 6)。

分析至此, 以所有数据进行了参考化合物人参皂甙 Rb1 的衍生物检索。为此, 将母离子分子式变化引入转化搜索列表。软件计算得到了质量数和分子式, 并对鉴定出的化合物划分了归属。这样, m/z 为 1195.6109、保留时间为 16.7 分钟的化合物被归为人参皂甙 Rb1 的丙二酰衍生物 (mRb1), 与 Rb1 的区别为一个 $C_3H_2O_3$ (Δm 为 86.0004) 基团⁴。有了这些信息, MS/MS 谱图中的 m/z 为 411.1119 和 249.0590 的碎片离子可进一步得到解释, 结构也可得以确定 (图 6)。对于所有 MS/MS 碎片离子的全部信息, 如质量数、分子式、母离子中性丢失基团、质量精度、衍生物质量数变化值、与参考化合物相比的质量数变化值等所有信息均列于表 2 中。

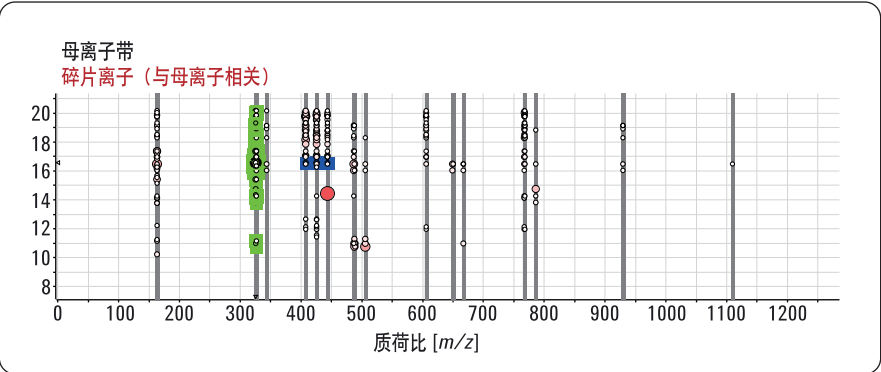


图 5
碎片离子概览图。对含 m/z 为 325 的已知质量数碎片的化合物 (以绿色突出显示) 进行进一步研究。一种保留时间为 16.7 分钟的化合物含有质量数已知、 m/z 为 407、425 和 443 的典型碎片 (以蓝色突出显示)。

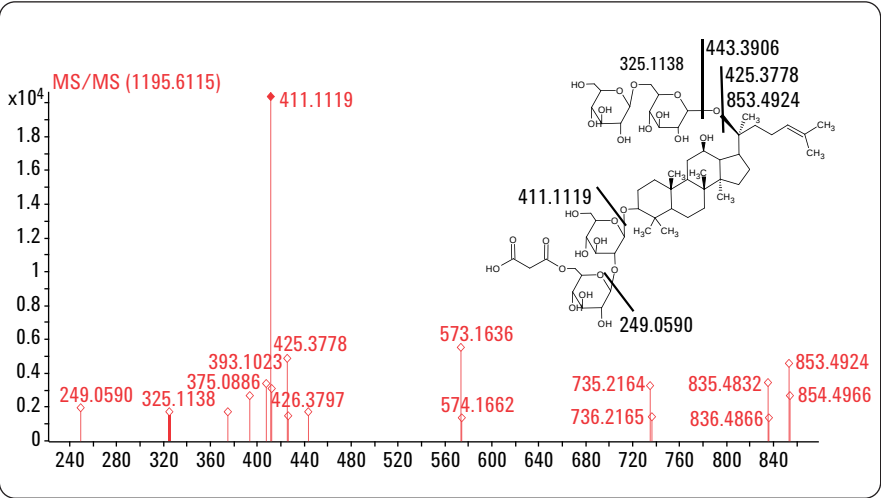


图 6
 m/z 为 1195.6115、保留时间为 16.7 分钟的化合物的 MS/MS 谱图, 及其推测的结构和主要碎片归属。

m/z	离子化学式	计算值 m/z	$\Delta m/z$ [mDa]	$\Delta m/z$ [ppm]	中性丢失基团	丢失基团的 化学式	丢失基团的 质量数	FPM m/z^*	m/z 变化值 **	$\Delta [mDa]$ 变化值***	化学式 变化
127.1110	$C_8H_{15}O$	127.11174	0.74	5.80	1068.50045	$C_{49}H_{80}O_{25}$	1068.49887				
249.0590	$C_9H_{13}O_8$	249.06049	1.53	6.14	946.55249	$C_{48}H_{82}O_{18}$	946.55012				
325.1138	$C_{12}H_{21}O_{10}$	325.11292	-0.90	-2.77	870.49763	$C_{45}H_{74}O_{16}$	870.49769	325.1129		0.91	
375.0886	$C_{15}H_{19}O_{11}$	375.09219	3.57	9.53	820.52284	$C_{42}H_{76}O_{15}$	820.51842				
393.1023	$C_{15}H_{21}O_{12}$	393.10275	0.43	1.08	802.50912	$C_{42}H_{74}O_{14}$	802.50786				
407.3666	$C_{30}H_{47}$	407.36723	0.58	1.42	788.24480	$C_{27}H_{48}O_{26}$	788.24338	407.3668		0.13	
411.1119	$C_{15}H_{23}O_{13}$	411.11332	1.44	3.50	784.49957	$C_{42}H_{72}O_{13}$	784.49729	325.1129	86.0004	1.43	$C_3H_2O_3$
425.3778	$C_{30}H_{49}O$	425.37779	-0.05	-0.11	770.23361	$C_{27}H_{46}O_{25}$	770.23282	425.3773		0.56	
443.3906	$C_{30}H_{51}O_2$	443.38836	-2.25	-5.07	752.22085	$C_{27}H_{44}O_{24}$	752.22225	443.3890		1.56	
573.1636	$C_{21}H_{33}O_{18}$	573.16614	2.50	4.36	622.44781	$C_{36}H_{62}O_8$	622.44447	487.1655	86.0004	2.24	$C_3H_2O_3$
735.2164	$C_{27}H_{43}O_{23}$	735.21896	2.51	3.42	460.39500	$C_{30}H_{52}O_3$	460.39165	649.2184	86.0004	2.36	$C_3H_2O_3$
835.4832	$C_{45}H_{71}O_{14}$	835.48383	0.68	0.82	360.12830	$C_{12}H_{24}O_{12}$	360.12678				
853.4924	$C_{45}H_{73}O_{15}$	853.49440	2.02	2.37	342.11907	$C_{12}H_{22}O_{11}$	342.11621	767.4925	86.0004	0.49	$C_3H_2O_3$

*FPM = 原始 (参考) 化合物的类似碎片的 m/z 值。 ** m/z 变化值 = 原始 (参考) 化合物的碎片离子与其衍生物碎片离子相比发生的质量数变化 = 由于衍生而发生的质量数变化。
*** $\Delta [mDa]$ 变化值 = $[(FPM\ m/z + m/z\ 变化值) - m/z] \times 1000$ = 与参考化合物相比的质量数变化值。

表 2
丙二酰人参皂甙 Rb1 (mRb1) 的 MS/MS 碎片离子、质量数、通过精确质量数计算得到的化学式、与母离子相比的中性丢失基团、通过精确质量数得到的中性丢失基团的化学式、由于衍生而产生的质量数变化值以及参与参考化合物相比的质量数变化值。

对质量偏差为 2.22 ppm。该化学式和分子量与已知化合物拟人参皂甙 F11 相匹配 (图 8)。MS/MS 谱图分析得到的碎片与该化合物的结构一致 (图 9)，这些碎片的计算化学式表明质量相对偏差很低 (表 3)。

结论

本应用简报报道了一种通过计算机软件辅助，在高度复杂的植物提取物中鉴定天然产物的方法。在本例中，采用 Agilent 6520 精确质量 Q-TOF LC/MS 系统对极其复杂的人参根提取物进行了分析。数据文件采用 Agilent MassHunter 代谢产物鉴定软件进行数据处理。首先，通过软件辅助分析进行数据库检索鉴定已知化合物。第二步，通过软件辅助搜索与已知化合物的相似性来鉴定未知化合物。为此，采用了同位素分布匹配和 MS/MS 碎裂模式匹配算法。人参皂甙 (提取物中的主要成分) 的复杂结构可通过对碎裂模式匹配结果和基于精确质量数的化学式计算结果的解释进行解析。第三步，将已经解析的样品与另一个人参品种的新样品进行比较，鉴定了新化合物和表达量增加的化合物。

m/z	离子化学式	计算值 m/z	$\Delta m/z$ [mDa]	$\Delta m/z$ [ppm]	中性丢失基团	丢失基团的化学式	丢失基团的质量数
143.1071	$C_6H_{15}O_2$	143.10666	-0.44	-3.09	658.39290	$C_{34}H_{58}O_{12}$	658.39283
309.1186	$C_{12}H_{21}O_9$	309.11801	-0.58	-1.88	492.38141	$C_{30}H_{52}O_5$	492.38147
421.3470	$C_{30}H_{45}O$	421.34649	-0.49	-1.16	380.15302	$C_{12}H_{26}O_{13}$	380.15299
439.3580	$C_{30}H_{47}O_2$	439.35706	-0.94	-2.13	362.14200	$C_{12}H_{26}O_{12}$	362.14243
457.3686	$C_{30}H_{49}O_3$	457.36762	-0.99	-2.17	344.13138	$C_{12}H_{24}O_{11}$	344.13186
801.5004	$C_{42}H_{72}O_{14}$	801.49950	-0.92	-1.14			

表 3

拟人参皂甙 F11 的 MS/MS 碎片质量数、计算化学式和计算质量精度。

参考文献

1. Liu, S., Cui, M., Liu, Z., Song, F., Mob, W., "Structural analysis of saponins from medical herbs using electrospray ionization tandem mass spectrometry," *J. Am. Soc. Mass Spectrom.* 15:133-144, **2004**.
2. Fuzzati, N., "Analysis methods of ginsenosides," *J. Chrom. B* 812:119-133, **2004**.
3. Wang, X., Sakuma, T., Asafu-Adjaye, E., Shiu, G. K., "Determination of ginsenosides in plant extracts from *Panax ginseng* and *Panax quinquefolius* L. by LC/MS/MS," *Anal. Chem.* 71:1579-1584, **1999**.
4. Kite, G. C., Howes, M. J. R., Leon, C. J., Simmonds, M. S. J., "Liquid chromatography/mass spectrometry of malonyl-ginsenosides in the authentication of ginseng," *Rapid Commun. Mass Spectrom.* 17:238-244, **2003**.
5. Li, W., Gu, C., Zhang, H., Awang, D. V. C., Fitzloff, J. F., Fong, H. H. S., van Breemen, R. B., "Use of high performance liquid chromatography-tandem mass spectrometry to distinguish *Panax ginseng* C. A. Meyer (Asian ginseng) and *Panax quinquefolius* L. (North American ginseng)," *Anal. Chem.* 72:5417-5422, **2000**.
6. Chan, T. D. W., But, P. P. H., Cheng, S. W., Kwok, I. M. Y., Lau, F. W., Xu, H. X., "Differentiation and authentication of *Panax ginseng*, *Panax quinquefolius*, and ginseng products by using HPLC/MS," *Anal. Chem.* 72:1281-1287, **2000**.

www.agilent.com/chem/metid:cn

© 安捷伦科技 (中国) 有限公司, 2009
2009 年 1 月 1 日出版
出版号 5990-3234CHCN



Agilent Technologies