

全球天然产物社会分子网络系统（GNPS）结合SCIEX高分辨质谱系统进行天然产物研究

Global Natural Products Social Molecular Networking system (GNPS) is combined with SCIEX High Resolution Mass Spectrometry system to study natural products

谢亚平，司丹丹，龙志敏

Yaping Xie, Dandan Si, Zhimin Long

SCIEX 应用支持中心，中国

Key words: Global Natural Products Social Molecular Networking system, SCIEX High Resolution Mass Spectrometry system, Natural Products

引言

全球天然产物分子网络系统（Global Natural Products Social Molecular Networking，简称GNPS）是近年发展起来的天然产物可视化鉴定技术，结构类似的天然产物，在相同的质谱裂解条件下，会产生相似的特征碎片离子，通过计算机计算这些碎片离子的相似度，按相似度大小，合成一张可视化的网络图谱，其中每一个节点表示一个化合物的MS/MS 质谱图，结构类似的化合物分子会在一个分子网络中聚集成簇。目前，分子网络技术作为研究工具广泛在天然产物领域使用。应用的主要方向包括中药化学成分定性表征、药物代谢方面的研究、化学成分定量表征及质量控制、活性成分的分离与制备、培养方法和提取工艺的优化、疾病诊断和个性化治疗等方向。GNPS数据库主要来源包括第三方数据库MassBank、HMDB、ReSpec、NIST（2014）等，Dorrestein实验室和Sirenas MD单位采集的化合物和谐图，全球多个科研群体提供的化合物和谐图等。GNPS质谱数据库不仅具有鉴定已知化合物（dereplication）、类似物（variable dereplication）和自动分析MN（Molecular Networking）中化合物的功能，还具有对多来源的MS/MS数据进行存储管理及更新数据的功能等。

本文本研究采用分子网络技术，将SCIEX高分辨质谱系统采集样品数据上传至GNPS，构建忍冬四个不同部位（S1-S4）化合物的分子网络关系，可获得四个部位最终成分鉴定结果、结构类似

物分子网络关系以及不同化合物成分在四个部位中的分布规律，为忍冬中不同部位的有效成分的研究提供依据。

技术特点

1. SCIEX高分辨质谱系统超快的扫描速度（最高133HZ或100HZ）保证采集到天然产物中更多的MS和MS/MS质谱信息，保证数据采集的全面性
2. SCIEX高分辨质谱系统独有的IDA+DBS采集模式，保证采集到更高质量的MS/MS质谱信息，保证数据采集的有效性
3. GNPS分子网络分析主要依赖高质量的MS/MS进行峰对齐及相似度分析，结合SCIEX高分辨质谱更全面有效的MS/MS信息，可获得更可靠全面的分析结果

仪器设备

SCIEX ExionLC™ AD系统和SCIEX 高分辨质谱系统



图1. ExionLC™ AD系统和SCIEX 高分辨质谱系统

液相条件

液相系统: SCIEX ExionLC™ AD系统

色谱柱: Phenomenex Kinetex C18(100 × 2.1mm, 2.6 μm)

流动相: A: 水 (含0.025%甲酸)

B: 甲醇:乙腈 =1:1 (含0.025%甲酸)

流速: 0.3 mL/min

柱温: 40 °C

液相梯度:

时间(min)	A(%)	B(%)
0	98	2
2	98	2
20	40	60
22	5	95
25	5	95
25.1	98	2
30	98	2

质谱条件

扫描模式: ESI+/- TOF MS-IDA-15TOF MS/MS

扫描范围: 一级m/z 60-1300 二级m/z 40-1300

CDS自动校正

动态背景扣除 (DBS) 开启

离子源参数:

喷雾电压: 5500V/-4500 V

气帘气 CUR: 35 psi

雾化气 GS1: 55 psi

雾化气 GS2: 55 psi

源温度: 550 °C

DP电压: ± 60 V

碰撞能量: ± 40 ± 20

实验流程

SCIEX高分辨质谱系统IDA+DBS采集模式获得的高质量数据上传至GNPS网站, 在IDA+DBS采集模式下可获得大量的高质量MS/MS质谱信息, GNPS通过计算机算法计算这些MS/MS质谱数据的相似度, 然后根据相似度的大小将这些质谱图整合成一张可视化的网络图谱, 如图2所示。

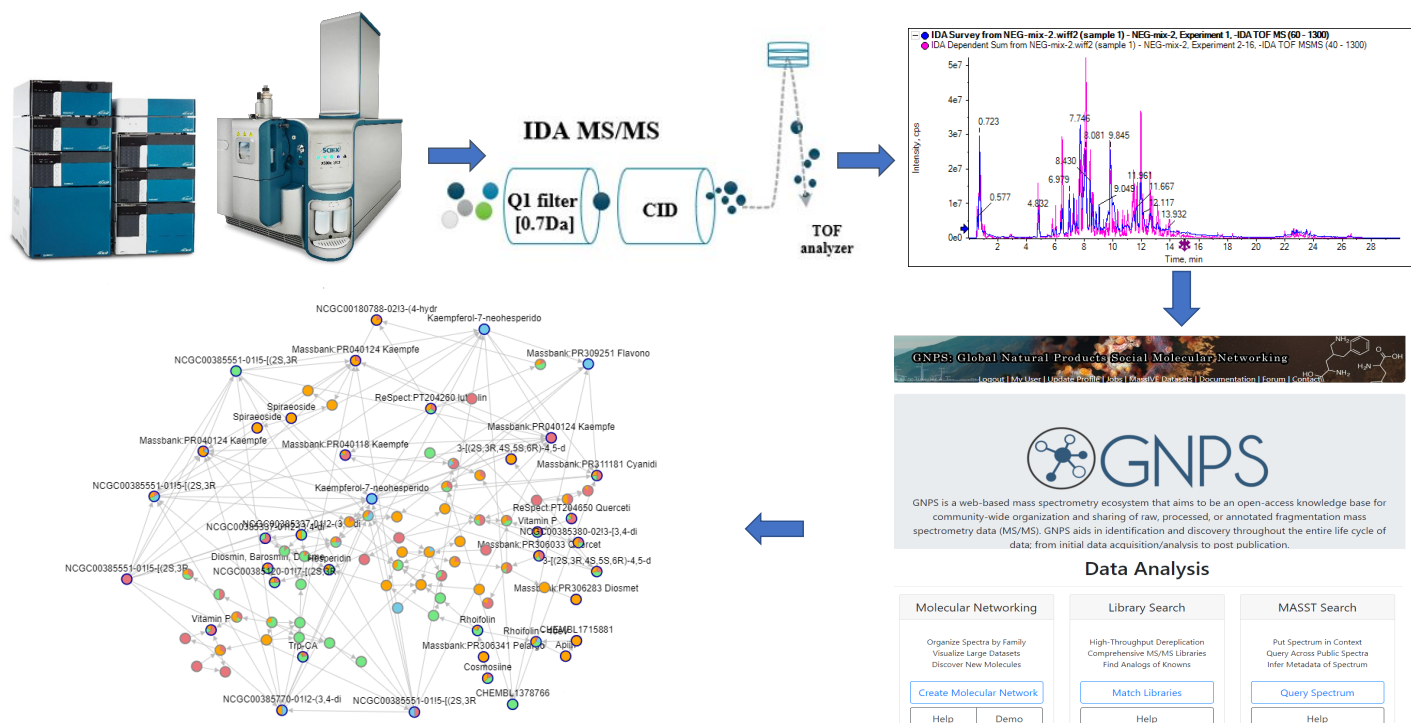


图2. 实验流程

实验结果

1. 天然产物成分鉴定结果

上传数据至GNPS网站进行数据处理之后可获得天然产物成分鉴定结果表，如图3所示

SpectrumID	Compound Name	Ion Source	Instrument	Ion Mode	Precursor	Exact Mass	RT	Query	molecular formula	superclass	class	subclass	Compound
CCMSLIB01	Chlorogenic Acid	N/A	ESI-QTOF	negative	353.088	0	461.255	C16H18O9	Organic oxygen c	Organooxygen c	Alcohols and	isolated	
CCMSLIB01	O-Phospho-SERINE	ESI	qToF	negative	184.002	0	50.524	C3H8NO4P	Organic acids and	Carboxylic acids	Amino acids	peptidol	
CCMSLIB01	Massbank PR30979 Flavon	ESI	qToF	Negative	333.093	0	799.325	C24H22O14	Phenylpropanoids	Flavonoids	Flavonoid glyco	isolated	
CCMSLIB01	ALPHA-N-ACETYL-L-GLU	ESI	qToF	negative	187.072	0	17.965	C7H12NO6	Organic acids and	Carboxylic acids	Amino acids	peptidol	
CCMSLIB01	MassNA 2224599 Lepargyl	LC-ESI	qToF	Negative	187	188.105	738.513	C9H16O4	Lipids and lipid-like	Fatty Acids	Fatty acids and	isolated	
CCMSLIB01	Massbank PR30956 Benz	ESI	qToF	Negative	315.071	0	228.584	C13H16O9	Organic oxygen c	Organooxygen c	Carbohydrates	isolated	
CCMSLIB01	ISOCITRIC ACID	ESI	qToF	negative	191.02	0	49.812	C6H8O7	Organic acids and	Carboxylic acids	Tricarboxylic	isolated	
CCMSLIB01	Tyr-UDCA	ESI	qToF	Negative	554.349	555.356	71.426	C33H49NO6	N/A	N/A	N/A	Crude	
CCMSLIB01	Massbank PR309158 LPC	ESI	qToF	Negative	264.33	0	1388.4	N/A	Lipids and lipid-like	Glycerophosphol	Glycerophosphol	isolated	
CCMSLIB01	Massbank PR309160 LPC	ESI	qToF	Negative	566.345	0	1398.99	N/A	Lipids and lipid-like	Glycerophosphol	Glycerophosphol	isolated	
CCMSLIB01	Massbank PR309160 LPC	ESI	qToF	Negative	566.345	0	1405.18	N/A	Lipids and lipid-like	Glycerophosphol	Glycerophosphol	isolated	
CCMSLIB01	Massbank PR309160 LPC	ESI	qToF	Negative	566.345	0	1405.48	N/A	Lipids and lipid-like	Glycerophosphol	Glycerophosphol	isolated	
CCMSLIB01	MassNA 2276521 D-Gluc	LC-ESI	qToF	Negative	195	196.058	39.744	C6H12O7	Organic oxygen c	Organooxygen c	Carbohydrates	isolated	
CCMSLIB01	CAMP	LC-ESI	qToF	Negative	328.04	0	130.464	N/A	N/A	N/A	N/A	isolated	
CCMSLIB01	Massbank PT000746 Tric	ESI	qToF	Negative	329.067	0	989.213	C17H14O7	Phenylpropanoids	Flavonoids	O-methylated	isolated	
CCMSLIB01	Massbank PR304595 S(S)	ESI	qToF	Negative	573.209	0	667.256	C28H34N2O	Lipids and lipid-like	Terpenol lipids	Terpene glycos	isolated	
CCMSLIB01	Massbank PT204580 4,8-E	ESI	qToF	Negative	577.135	0	416.13	C30H26O12	Phenylpropanoids	Flavonoids	Biflavonoids and	isolated	
CCMSLIB01	Massbank PR309158 FFA	ESI	qToF	Negative	239.231	0	1018.91	C18H34O6	Lipids and lipid-like	Fatty Acids	Fatty acids and	isolated	
CCMSLIB01	Massbank PR306189 Form	ESI	qToF	Negative	591.172	0	889.48	C38H32O14	Phenylpropanoids	Flavonoids	Flavonoid glyco	isolated	
CCMSLIB01	Tiboside	DI-ESI	qToF	Negative	593.116	594.137	928.506	C30H26O13	Phenylpropanoids	Flavonoids	Flavonoid glyco	isolated	
CCMSLIB01	Massbank PR040118 Kam	ESI	qToF	Negative	593.2	0	740.566	C27H30O15	Phenylpropanoids	Flavonoids	Flavonoid glyco	isolated	
CCMSLIB01	Massbank PR307061 NP	ESI	qToF	Negative	593.151	0	561.543	C27H30O15	Phenylpropanoids	Flavonoids	Flavonoid glyco	isolated	

图3. GNPS谱库鉴定结果

2. 天然产物中结构类似物分析结果

上传数据至GNPS网站进行分子网络数据处理之后可获得可视化分子网络图，如图4所示。其中可以看到蓝色标记区域内为化合物m/z 447.096及其结构类似物，与m/z 447.096结构类似的化合物包括m/z 593.151、m/z 579.135、m/z 493.099、m/z 531.148、m/z 483.837，MS/MS结构相似度采用余弦值表示，五个化合物与化合物m/z 447.096的相似度分别为0.92、0.90、0.90、0.86、0.88（如图中蓝色字体标注余弦值）。其中m/z 447.096 GNPS谱库匹配结果为木犀草苷（luteolin-7-O-glucoside），m/z 593.151 GNPS谱库匹配结果为山奈酚-3-O-芸香糖苷（Kaempferol-3-o-rutinoside），以这两种化合物为例展示其MS/MS及结构相似性，可见两者具有多个相同的主要碎片离子（图中黄色区域标记）并且具有相似化合物结构，图5所示。

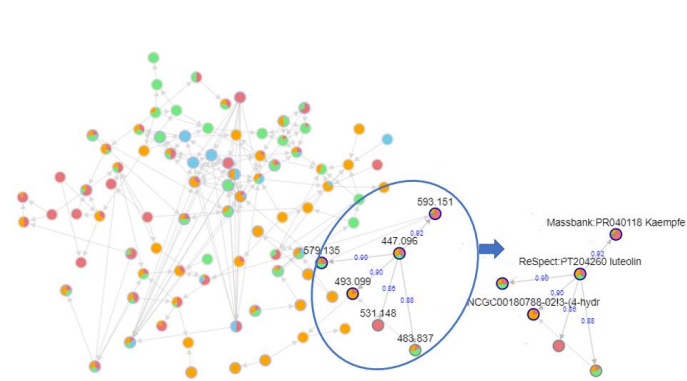


图4. GNPS结构类似物分析结果

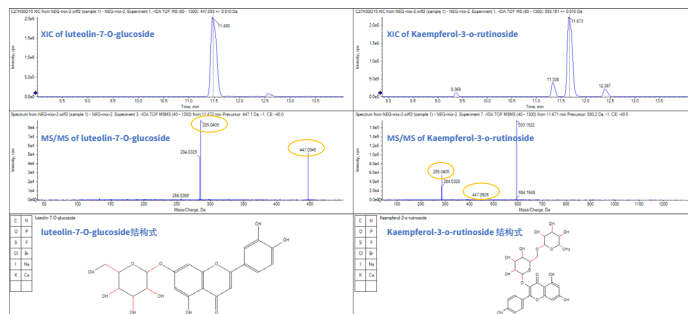


图5. luteolin-7-O-glucoside和Kaempferol-3-o-rutinoside相似性对比

3. 天然产物成分差异分析

上传数据至GNPS网站进行分子网络数据处理之后可获得可视化分子网络图，如图6所示。其中可以看到蓝色标注区域为绿原酸（Chlorogenic acid）及其结构类似物分子网络图，在该区域可以看到，每种成分（即每个节点）均含有一种甚至几种不同颜色，例如绿原酸（Chlorogenic acid）含有红蓝绿黄四种不同颜色，这四种颜色分别代表不同部位忍冬样品S1-S4，因此我们可以快速直观看到忍冬S1-S4四个不同部位均含有中药成分绿原酸（Chlorogenic acid）。每个节点的不同颜色以饼状图显示，其中所占饼图的比例越大代表该成分在样品中含量越高，因此可以直观看到绿原酸（Chlorogenic acid）在S1（红色）样品中含量最高，在S2（蓝色）样品中含量最低。通过可视化的分子网络图可以快速发现不同样品之间的成分差异，与目前应用较多的PCA、PLS-DA、OPLS-DA等方法相比，分子网络技术具有数据分析能力强、可视化程度高、数据兼容性强、结果准确可靠等优势。



图6. 绿原酸在不同样品中（S1-S4）含量分布差异（其中红蓝绿黄四种颜色分别代表S1-S4）

总结

本文使用SCIEX 高分辨质谱系统获得高质量的MS及MS/MS信息，SCIEX高分辨质谱系统在超快的扫描速度下保证了数据的高灵敏度，全面性和有效性。GNPS利用SCIEX高质量的MS/MS数据可快速进行天然产物的鉴定、结构类似物分析以及差异成分分析，为天然产物的研究提供了更加快速便捷可视化的分析方法，可很好地应用于天然产物的多方向研究分析。

SCIEX临床诊断产品线仅用于体外诊断。仅凭处方销售。这些产品并非在所有国家地区都提供销售。获取有关具体可用信息，请联系当地销售代表或查阅<https://sciex.com.cn/diagnostics>。所有其他产品仅用于研究。不用于临床诊断。本文提及的商标和/或注册商标，也包括相关的标识、标志的所有权，归属于AB Sciex Pte. Ltd. 或在美国和/或某些其他国家地区的各权利所有人。

© 2022 DH Tech. Dev. Pte. Ltd. RUO-MKT-02-14562-ZH-A



SCIEX中国

北京分公司
北京市朝阳区酒仙桥中路24号院
1号楼5层
电话：010-5808-1388
传真：010-5808-1390
全国咨询电话：800-820-3488, 400-821-3897

上海公司及中国区应用支持中心
上海市长宁区福泉北路518号
1座502室
电话：021-2419-7200
传真：021-2419-7333
官网：sciex.com.cn

广州分公司
广州国际生物岛星岛环北路1号
B2栋501、502单元
电话：020-8510-0200
传真：020-3876-0835
官方微信：[SCIEX-China](#)