

代谢组学简介

代谢组学是一个新兴的“组学”研究领域，主要对代谢物（即细胞代谢的最终产物）进行特性识别和鉴定。代谢组学研究将会产生与成百上千种代谢物相关的复杂数据集。代谢组学数据的综合分析需要采用分析方法和数据分析策略，而这些分析方法和数据分析策略通常都是唯一的，需要专门的数据分析软件才能进行化学信息学、生物信息学和统计学分析。Agilent 产品为您提供提供了必要的数据分析工具。

Agilent 代谢组学工作流程可指导您使用 Agilent MassHunter 定性分析和 Agilent Mass Profiler Professional 来设置试验设计，收集重复的样品，评估系统适应性，提高数据质量并执行综合的代谢组学数据分析。

MassHunter 定性分析软件可让您自动从样品中查找并提取所有的质谱和色谱信息，甚至在未完全解析成份时也能进行如此操作。功能强大的数据导航能力可让您浏览单个样品中特定化合物的信息，并比较多个样品的色谱和质谱。此软件还包含一个可自定义的用户界面以及将结果保存、导出或复制到其他应用程序（如 Mass Profiler Professional）中的功能。

Mass Profiler Professional 是一个功能强大的数据简化、分析和可视化工具，可让您针对复杂数据集（如在代谢组学中遇到的生物基质的化学特征）提出的简单问题找到统计意义显著的答案。

您可以将 Agilent 代谢组学工作流程作为任何分析的线路图，这些分析需要针对复杂数据集提出的简单问题确定统计意义显著的答案。此代谢组学工作流程可用于执行以下分析：

- 比较两个或多个生物组
- 查找并确定潜在的生物标记
- 查找毒物学的生物标记
- 了解生物途径
- 发现新的代谢物
- 开发可为一组样品生成特性标记的数据挖掘和数据处理过程
- 构建用于样品分类的统计模型。

小分子调节 (LC/MS)

代谢组学包括对小分子进行分析，小分子是指分子量从 50 到 6000 amu 的分子。由于代谢组学研究的最佳结果涉及到识别分子离子的精确质量，因此，应该调整 LC/MS 仪器调节，以 (1) 提高对完整分子离子的灵敏度和 (2) 提高对小型低质量化合物的总体灵敏度。典型的自动仪器调节可优化仪器在整个质量范围内的灵敏度，但可能会导致与高破碎比代谢组学相结合的小分子物质的敏感度较低。

必需的项目

在使用以下“必需”部分中所述的硬件和软件时，代谢组学工作流程性能最佳。必需的硬件和软件用于执行图 1 中所示的数据分析任务。

Agilent 代谢组学工作流程

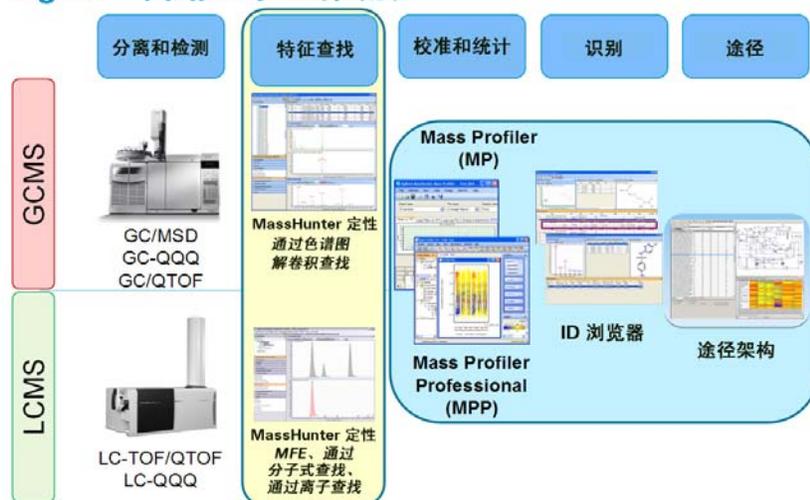


图 1 用于执行代谢组学工作流程的 Agilent 硬件和软件。

典型的 Agilent 代谢组学工作流程如图 2 所示，从数据采集开始，到涉及未定目标（发现）的 LC/MS 和已定目标（确认）的 LC/MS/MS 的分析。分子特征提取 (MFE) 和按分子式查找 (FbF) 是 MassHunter 定性分析用于查找化合物的两种不同算法。可以将 Agilent 分析平台生成的所有结果文件导入到 Mass Profiler Professional 中，用于质量控制、统计分析、可视化处理和进行解释。

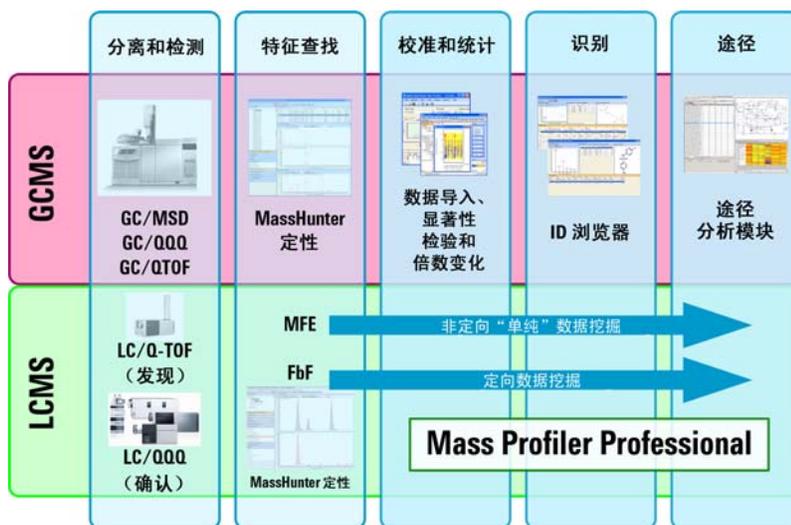


图 2 Agilent 代谢组学工作流程（从独立到途径）分析通常会涉及 GC/MS 和 / 或 LC/MS 分析。

必需的硬件

- 运行 Windows 的 PC
 - **最低配置:** XP SP3 (32 位) 或 Windows 7 (32 位或 64 位), 4 GB RAM
 - **建议的配置:** Windows 7 (64 位), 8 GB 或更大 RAM
- 硬盘驱动器的 C 分区至少要有 50 GB 的可用空间
- 来自 Agilent GC/MS、LC/MS、CE/MS 和 / 或 ICP-MS 系统中的数据或从其他仪器导入的数据。

必需的软件

- Agilent Mass Profiler Professional 软件 B.12.00 或更高版本
- Agilent MassHunter 定性分析软件 B.03.01、B.04.00、B.05.00 SP1 或更高版本
- Agilent MassHunter 数据采集软件 B.03.02、B.04.00、B.05.00 或更高版本 (这将包括 Agilent MassHunter DA Reprocessor)
- Agilent MassHunter 定量分析软件 B.03.02 或更高版本

可选软件

- Agilent ChemStation 软件
- AMDIS
- MassHunter ID 浏览器 B.03.01 或更高版本
- METLIN 个人化合物数据库和谱库
- Agilent Fiehn GC/MS 代谢组学谱库

什么是代谢组学？

代谢组学是在指定的生物状态下确定某个有机体的所有代谢物及其数量的过程。有机体的代谢物表示在完整定义状态下的化学“指纹”，这种状态是由一组特定情况定义的。

代谢组学工作流程涉及对两种类型的变量进行研究：

自变量：在采样之前，您已知有机体状态的一个或多个属性。这些属性称作自变量。

在代谢组学数据分析的各个步骤期间，代谢组学工作流程是指已知的有机体状态或影响有机体的外部因素，如参数值、条件或属性值。在统计分析中，已知状态和外部因素表示自变量。

因变量：自变量发生变化时可观测到的生物反应。在代谢谱中，此生物反应变化很明显。浓度发生明显变化的每个代谢物称作因变量。

在代谢组学数据分析中，样品中的代谢物可分别指化合物、特征、元素或实体。在统计分析中，代谢物表示因变量。

如果存在成百上千种代谢物，研究人员将采用化学计量数据分析来找出自变量和代谢谱之间的准确且有统计意义的相关性。研究人员使用来自临床诊断的代谢物反应中的有用信息，以更好地了解人类疾病的发生和进展，从而更好地进行治疗。此分析是图3中所述的代谢组学工作流程。

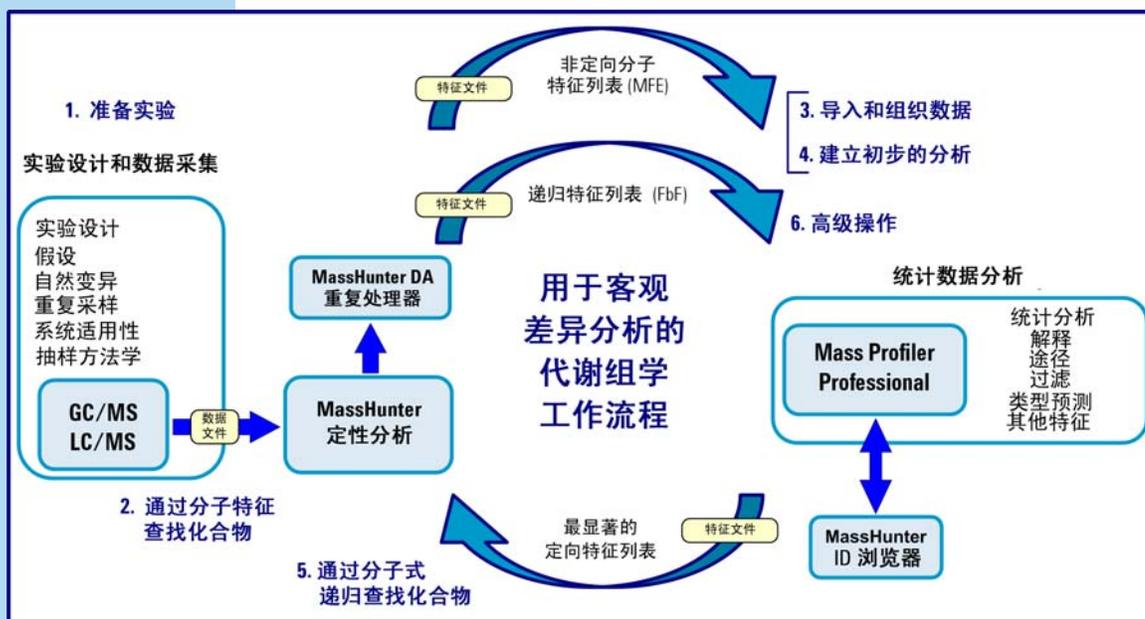


图3 Agilent 代谢组学工作流程图

假设

代谢组学流程的首要步骤是提出由分析回答的相关性问题，即假设。此问题是提出关于一组自变量和结果代谢谱之间可能存在相关性（例如，因果关系）的陈述。代谢组学工作流程用于证明或驳斥假设。

自然变异

在代谢组学工作流程开始之前，了解任意样品如何在整体上表示总体，这非常重要。由于存在与测量数据和总体关联的自然变异和不确定性，不能保证某个总体中的任何单个样品都可表示总体的均值。这样，大幅度提高样品数量可以提高样品组描述总体特性的准确度。

重复采样

因为受时间、资源和资金限制，对整个总体采样通常不可行。另一方面，样品越少，得出误报或漏报相关性的概率越高。对于您的代谢组学研究中每种条件的每个属性值，建议您的代谢组学分析至少包含十 (10) 个或更多重复样品。

系统适应性

系统适应性涉及收集数据以提供用于评估和补偿漂移和仪器误差的方式来确保高质量的结果。代谢组学工作流程在以下方面采用的技术可生成最高质量的结果：(1) 保留时间一致、(2) 强度归一化、(3) 色谱解卷积和 (4) 基线。但是，此过程无法补偿采集参数中的过度漂移。可通过维护仪器并使用优质的色谱来获得最佳结果。

采样方法

要提高代谢组学分析的数据质量，需要将采样方法与试验设计相匹配，以便收集重复的数据来获得每种条件下的属性值。被研究的总体的适用样品数量越多，对假设的回答就越准确。了解在采样中使用的方法和使用多种样品收集方法会对结果的显著性带来正面影响。

工作流程简介

步骤 1

Agilent 代谢组学工作流程包括六个步骤，涉及使用 Agilent MassHunter 定性分析、Agilent MassHunter DA Reprocessor 和 Agilent Mass Profiler Professional。

准备试验

代谢组学工作流程的首要步骤是提出假设，即要由分析回答的相关性问题。您的准备工作包括试验定义，该定义考虑了自然变异性和重复采样，检查系统适应性和采样方法，以提高分析的意义。

步骤 2

查找特征

按分子特征查找化合物 (MassHunter 定性分析)

根据质谱和色谱特性，从您的数据中提取化合物 (也称作分子特征)。此流程是指分子特征提取 (MFE)。分子特征提取可自动快速生成完整、准确的化合物列表，其中包括分子量、保留时间、 m/z 和丰度。

步骤 3

导入并整理数据

整理、导入和准备您的数据 (Mass Profiler Professional)

在创建一个项目和试验后，“MS 试验创建”将指导您完成整理试验、导入数据、定义试验变量和准备分析数据所必需的步骤。数据准备包括过滤、调整一致、归一化和确定基线。

步骤 4

创建初始分析

质量控制和初始差值表达式 (Mass Profiler Professional)

“显著性检验和倍数变化向导”将指导您完成输入参数和数值的必要步骤，以提高结果质量并生成分析的初始差值表达式。

步骤 5

递归查找特征 (可选)

按分子式查找化合物 (MassHunter 定性分析)

将最显著的特征作为目标特征导回 MassHunter 定性分析中，可提高在样品中查找特征的可靠性。此重复特征查找称作递归。提高查找特征的可靠性也会提高分析的准确度。

步骤 6

高级操作

自定义分析并解释结果 (Mass Profiler Professional)

数据中最显著的特征由 Mass Profiler Professional 进行处理，以用于最终统计分析和解释。最终解释的结果可用于证明或驳斥您的假设，也可以用于创建样品类预测模型。

使用 Mass Profiler Professional

Mass Profiler Professional 可导入由 MassHunter 定性分析创建的 CEF 文件。Mass Profiler Professional 可利用色谱 / 质谱数据的丰富信息，并轻松导入、分析和可视化较大样品组集中的 GC/MS、LC/MS、CE/MS 和 ICP-MS 数据和复杂的 MS 数据集。

Mass Profiler Professional 可向您提供一组分步式指南，以 (1) 设置项目和试验、(2) 导入数据文件、(3) 对数据文件进行排序和分组、(4) 对样品数据进行过滤、调整一致和归一化处理以及 (5) 创建初始分析。

由于工作流程浏览器中的高级操作不会指导您完成数据导入和差值分析的初始步骤，因此，建议不要跳过代谢组学工作流程的“导入和整理数据”或“创建初始分析”步骤。可以在代谢组学工作流程结束时，通过使用工作流程浏览器中的操作（图 4（第 12 页））编辑所有参数（包括在 MS 试验创建向导期间使用的默认参数）。

注释：可随时按 **F1** 或参考《Mass Profiler Professional 用户手册》获得有关各种字段和统计学处理的帮助和详细信息。

导入和整理您的数据

在创建项目和试验后，“MS 试验创建向导”可指导您完成整理试验、导入数据、定义试验变量和准备分析数据所必需的步骤。创建试验需要完成四个主要步骤：(1) 设置项目和试验，(2) 导入数据文件，(3) 对数据文件进行排序和分组，(4) 对样品数据进行过滤、调整一致和归一化处理。

a 启动 Mass Profiler Professional

双击位于桌面上的 Mass Profiler Professional 图标 ()，或者

单击开始 > 所有程序 > Agilent > MassHunter Workstation > Mass Profiler Professional > Mass Profiler Professional

b 设置项目和试验

项目是试验集合的容器。项目可以有多个关于不同样品类型和有机体的试验。项目设置包含四个步骤：

启动：单击**创建新项目**。

创建新项目：键入有关项目的描述性信息。

试验选择对话框：单击**创建新试验**或**打开现有试验**。

新试验：键入和选择指导您创建试验的信息。

c 将您的数据文件导入到此试验中。

在“MS 试验创建向导”的步骤 1 和步骤 2 期间，您的数据文件将导入到 Mass Profiler Professional 中。

步骤 1. 选择数据源：选择生成了您要用于试验的分子特征的数据源。

步骤 2. 选择要导入的数据：选择分子特征样品文件。

d 对数据文件进行排序和分组。

在“MS 试验创建向导”的步骤 5 中对数据文件进行排序，在步骤 6 中进入试验分组过程。

步骤 5. 样品重新排序：整理样品，方式是：选择和取消各个样品，对选定的内容进行重新排序，以根据自变量对样品进行分组。

步骤 6. 试验分组：根据自变量（包括试验的重复结构）定义样品分组。

e 对样品数据进行过滤、调整一致和归一化处理。

您可以在“MS 试验创建向导”的步骤 7 到步骤 11 中对样品数据进行过滤、调整一致和归一化处理。在此流程的每个步骤中，您都可以查看您的进度，并返回到之前的步骤以调整结果。

步骤 7. 过滤：按丰度、质量范围、每个特征的离子数和电荷态过滤分子特征。

步骤 8. 调整一致：根据保留时间和质量确定的容差，将样品特征调整一致。在试验类型为“已标识”时，省略此步骤，因为已标识的化合物将作为标识调整的化合物。

步骤 9. 样品摘要：显示样品中已调整和未调整实体分布的质量与保留时间图、电子表格和复合频率。复合频率图提供了未识别试验类型的调整效率的便捷视图。向导中的“上一步”和“下一步”按钮可让您轻松查看不同调整和过滤选项的效应。

步骤 10. 归一化标准：将样品特征的信号强度按比例缩放到指定的算法或外部标量计算的值。

步骤 11. 基线选项：将每个样品的信号强度与根据所有样品或控制样品计算的代表值进行比较。

现在，您已完成了数据的导入和整理过程。在下一个工作流程步骤中，您将使用 Mass Profiler Professional 根据数据创建初始差值表达式。

创建初始分析

“显著性检验和倍数变化向导”将指导您完成输入参数和数值的必要步骤，以提高结果质量并生成分析的初始差值表达式。

“显著性检验和倍数变化”工作流程可帮助您根据数据创建初始差值表达式，并根据以前使用分子特征提取功能找到的所有特征标识最显著的特征。创建初始差值表达式所必需的步骤是预先确定的，并且基于在创建项目和设置试验时输入的试验类型、试验分组和条件。

工作流程会在左侧导航器上显示步骤序列，并突出显示当前步骤。可能会自动跳过试验的某些步骤。可以在分析结束时编辑所有参数：“显著性检验和倍数变化”工作流程（通过使用工作流程浏览器中可用的操作）。

步骤 1. 摘要报告：根据您在“导入数据”向导中提供的参数，显示试验的摘要视图。显示一个谱图，其中包含 x 轴上的样品和 y 轴上的对数归一化丰度值。如果样品数量大于 30，则数据用电子表格视图（而不是谱图）表示。

步骤 2. 试验分组：必须指定自变量和自变量的属性值才能定义样品分组。自变量称作参数名称。自变量内的属性值称作参数值。参数名称内具有相同参数值的样品将被当作是重复样品。

步骤 3. 过滤标记：在试验创建期间创建的化合物现在称作实体。将根据实体在样品和参数值（现在称作条件）中的出现情况，在进一步分析中过滤（删除）这些实体。

步骤 4. 按频率过滤：将根据实体在指定样品和条件中的出现频率，进一步过滤实体。此过滤会删除不能复制的实体。

步骤 5. 样品质量控制：通过分组和当前的主成份分析法 (PCA) 表示样品。PCA 可计算所有可能的主成份，并在 3D 散点图中以可视方式表示它们。按轴刻度显示的分數可用于检查数据质量。散点图中，每个由试验分组进行颜色编码的样品用一个点表示。组内的重复点应该群集化显示在一起，与其他组中的样品分离。

步骤 6. 显著性分析：基于从统计分析计算的 p 值，过滤实体。所执行的统计分析取决于样品和试验分组。

步骤 7. 倍数变化：基于化合物的丰度比或处理与控制之间的差值（大于指定的截止值或阈值），进一步过滤化合物。

步骤 8.ID 浏览器标识：最终实体列表将直接导入 ID 浏览器（以便于标识）并返回到 Mass Profiler Professional。

“显著性检验和倍数变化”工作流程可让您使用**下一步**按钮完成每一个步骤。每个后续步骤中均会提供分析摘要。在查看分析进度后，您可以使用**上一步**按钮返回到任何以前的步骤，并进行更改。要更加熟悉这些分析参数以及参数对数据的影响，建议常用**上一步**和**下一步**按钮。

要退出此向导并跳过此向导中的后续步骤，请在任何步骤中单击**完成**。在单击**完成**后，将保存实体列表，并且您可以使用工作流程浏览器中可用的高级操作开始进行分析。

高级操作

数据中最显著的特征由 Mass Profiler Professional 进行处理，以用于最终统计分析和解释。最终解释的结果可用来证明或驳斥您的假设，也可用来创建样品类预测模型或评估途径。

工作流程浏览器中可用的操作可为您提供根据您的质谱数据分析特征所必需的工具，具体取决于您的分析的需求和目标、试验设计和研究焦点。这可帮助您创建不同的解释，以根据不同的过滤、归一化和标准统计方法执行分析。

由于工作流程浏览器中的高级操作不会指导您完成数据导入和差值分析的初始步骤，因此，建议不要跳过代谢组学工作流程的“导入和整理数据”或“创建初始分析”步骤。可以在代谢组学工作流程结束时，通过使用工作流程浏览器中的操作（图 4（第 12 页））编辑所有参数（包括在 MS 试验创建向导期间使用的默认参数）。

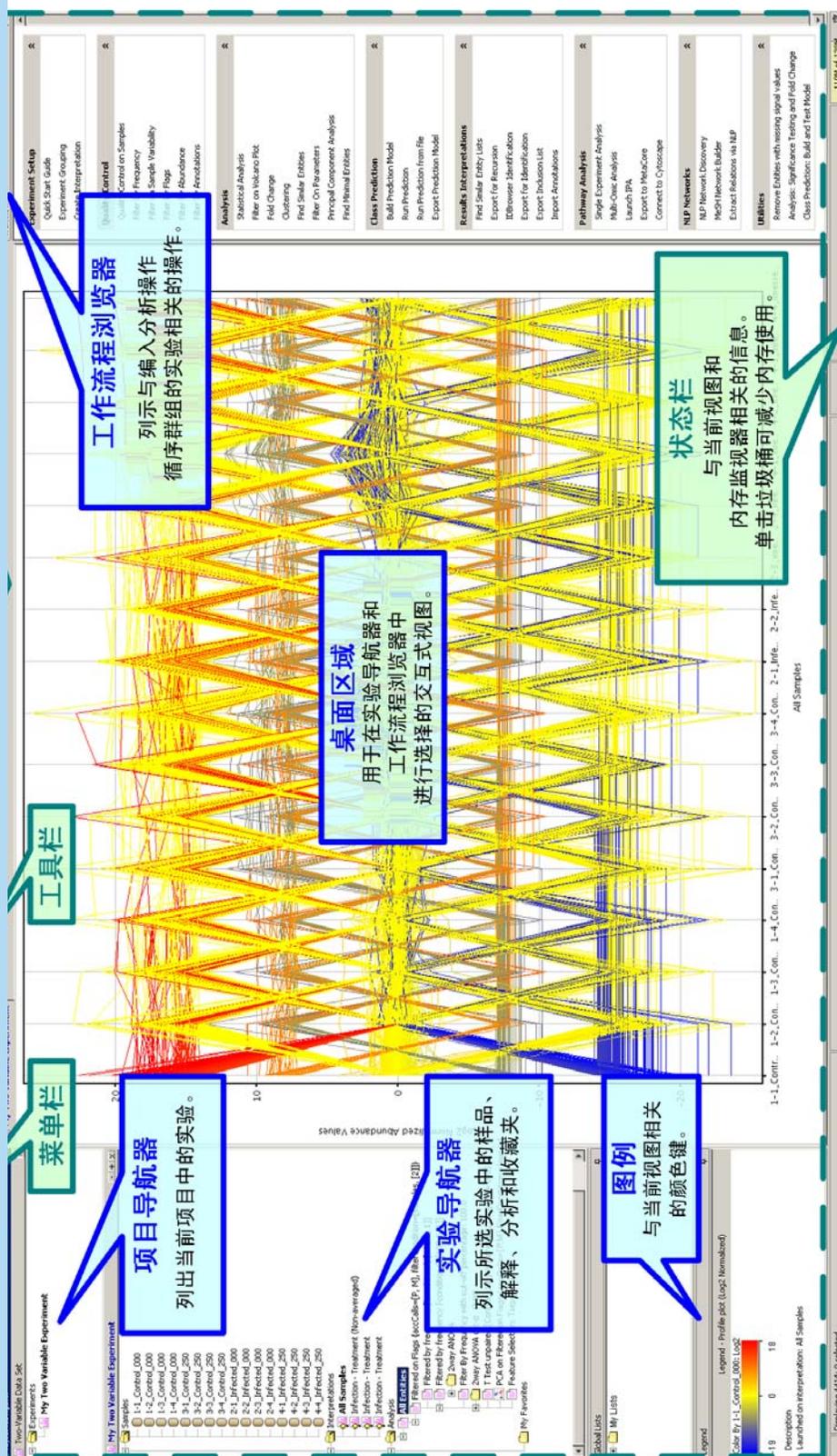


图4 Mass Profiler Professional 的主要功能区

详细信息

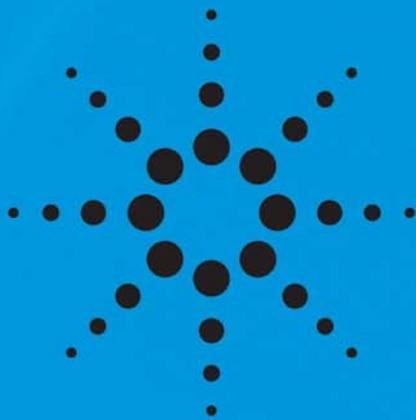
有关这些过程的详细信息，请参见《Agilent 代谢组学发现工作流程指南》。请访问我们的网站，网址为 www.agilent.com/chem。

代谢组学发现工作流程是 Agilent 手册、帮助、应用程序注释和培训视频的集合中的一部分。最新的手册和帮助集合对于了解代谢组学工作流程和需要熟悉 Agilent 软件工具的用户很有价值。培训视频提供了使用软件工具的分步式说明，以减少 GC/MS 和 LC/MS 数据示例，但需要投入大量的时间和精力来推测示例过程。此工作流程提供了使用 Agilent MassHunter 定性分析和 Agilent Mass Profiler Professional 执行代谢组学数据分析的分步式概述。

为了提供与代谢组学和 Agilent Mass Profiler Professional 相关的材料，精选了以下出版物：

- **手册：** Agilent 代谢组学工作流程 - 发现工作流程指南
(Agilent 出版物 [5990-7067EN](#)，修订版 B，2012 年 10 月)
- **宣传册：** Agilent 代谢组学解决方案
(Agilent 出版物 [5990-6048EN](#)，2012 年 4 月 30 日)
- **宣传册：** Agilent Mass Profiler Professional 软件
(Agilent 出版物 [5990-4164EN](#)，2012 年 4 月 27 日)
- **应用程序：** Mass Profiler Professional、个人化合物数据库和谱库软件，有助于识别化合物，以获得酵母代谢物的谱图
(Agilent 出版物 [5990-9858EN](#)，2012 年 4 月 25 日)
- **应用程序：** 使用 Agilent GeneSpring 11.5 分析平台的多组学分析
(Agilent 出版物 [5990-7505EN](#)，2011 年 3 月 25 日)
- **演示文稿：** 用于关键生物途径的目标标识的多组学分析软件
(Agilent 出版物 [USHUPO_IB_March_2012.pdf](#)，2012 年 3 月)
- **应用程序：** 使用 Mass Profiler Professional 软件和 LC 三重四极杆 MRM 确认的 LC/MS 代谢组学发现工作流程（针对疟疾感染红细胞）
(Agilent 出版物 [5990-6790EN](#)，2010 年 11 月 19 日)
- **宣传册：** Agilent 开发的集成生物技术：未来的新兴技术
(Agilent 出版物 [5990-6047EN](#)，2010 年 9 月 1 日)
- **基础知识：** 代谢组学：使用质谱法测定的方法
(Agilent 出版物 [5990-4314EN](#)，2009 年 10 月 27 日)

可以在 Agilent 代谢组学工作流程中的“参考资料”中找到完整的参考资料列表。



www.agilent.com



© Agilent Technologies, Inc. 2012
Revision B, October 2012



5990-7068CH



Agilent Technologies

