

Biodeep™

全球领先的代谢组学云分析平台

Biodeep Online为每位代谢组用户提供生物信息数据的存储、分析

谢桂纲 于建东

20190118

- 科学存储、解卷积与鉴定、生物信息挖掘于一体的代谢组学云分析平台



01

解卷积与科学数据存储



02

小分子代谢物鉴定



03

非靶向代谢组学分析





收费方式



- 免费存储（永久免费空间30G，超出需购买）；
- 免费解卷积



- 付费：根据最终代谢物鉴定数量，10个米克币/每个代谢物。



- 付费：根据比对组别数量，单次分析（首组800米克币，每增加一组，另付400米克币）。



关于“米克币”

- 米克币是什么？
Biodeep付费功能所需要的“货币”；
- 怎么获得米克币？
 - 1) 免费获得：线下客户赠送、首次注册、提出改进意见等；
详情参考：<http://help.biodeep.cn/#platform/3.0getMoney.md>
 - 2) 购买：按照RMB：米克币=1:1，量多从优。
充值入口：http://my.biodeep.cn/index.php?app=balance#account_recharge

上传数据

- 新建项目
- 编辑分组[例如：pos , neg]
- 上传数据与文件管理

创建任务

- 参数选择
- 创建任务

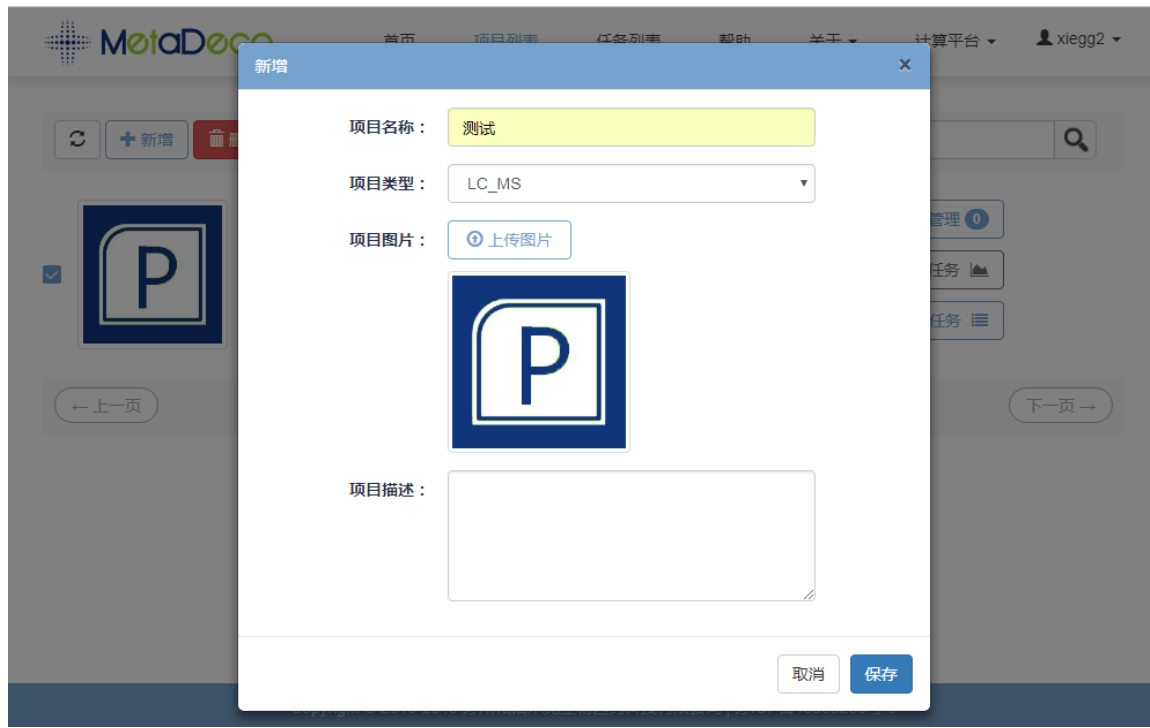
查看结果

- 在线查看
- 结果下载

详情参考: <http://help.biodeep.cn/>

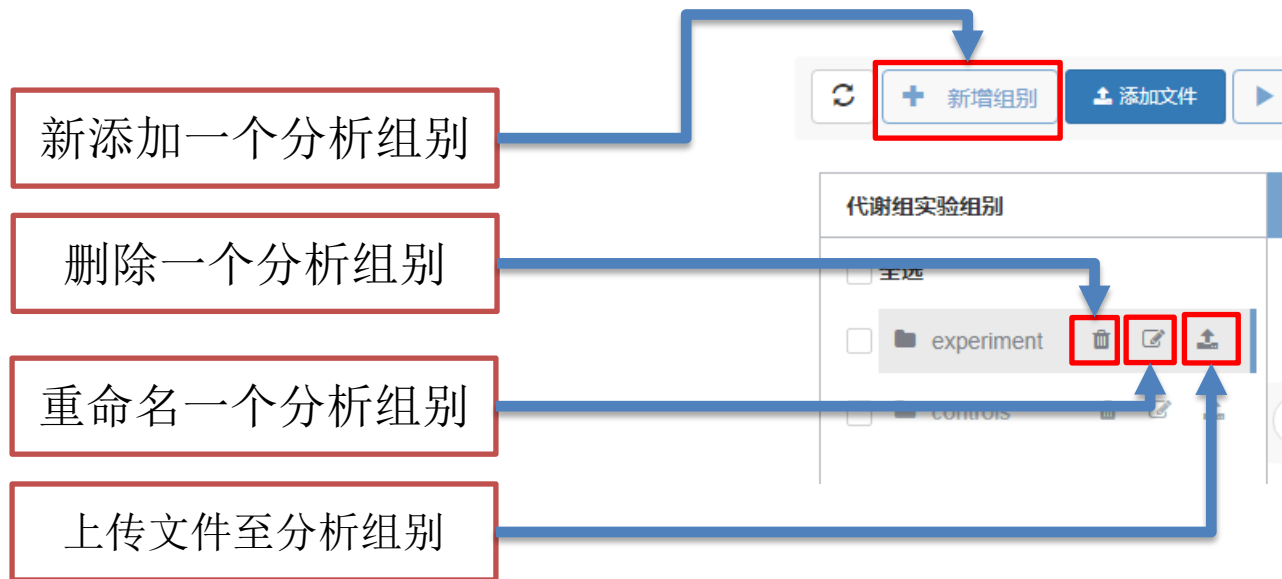
上传数据

- 1. 新建分析项目
- http://mz.biodeep.cn/index.php/index/project_home



上传数据

- 2. 编辑样本的分组信息
- 会需要在文件管理器之中编辑分组信息，以方便进行后续的比较分析



上传数据

- 3. 上传原始数据文件
- LCMS: 支持mzXML, Raw(Thermo)等
- GCMS: 支持CDF
- 因为服务器后台比较繁忙, 建议首先转换为mzXML格式之后再上传, 可以缩短队列等待时间



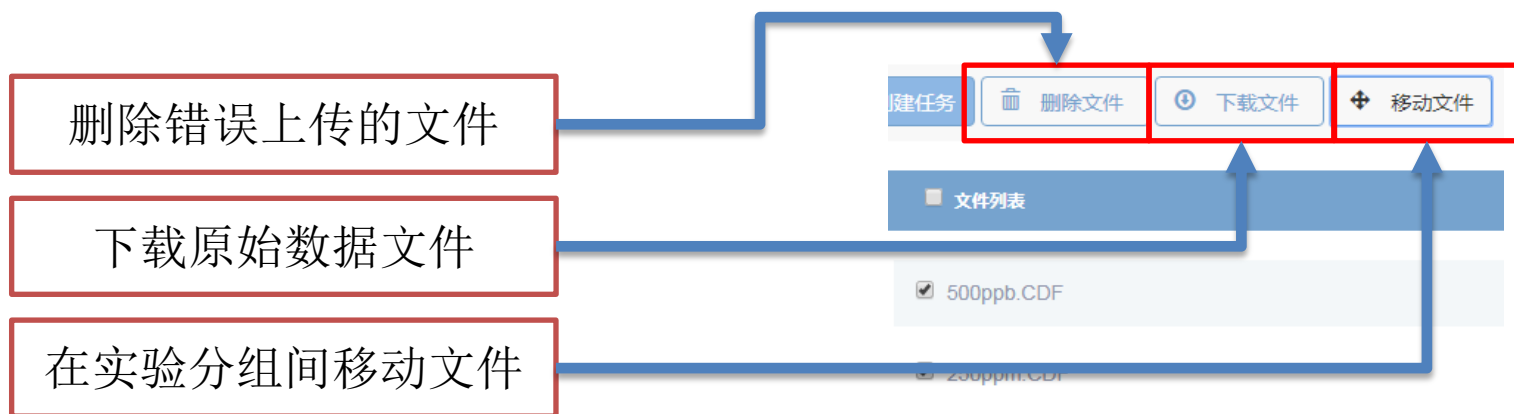
格式转换软件: <http://proteowizard.sourceforge.net/>

参考文献: Meta-analysis of untargeted metabolomic data from multiple profiling experiments, Nature protocols.



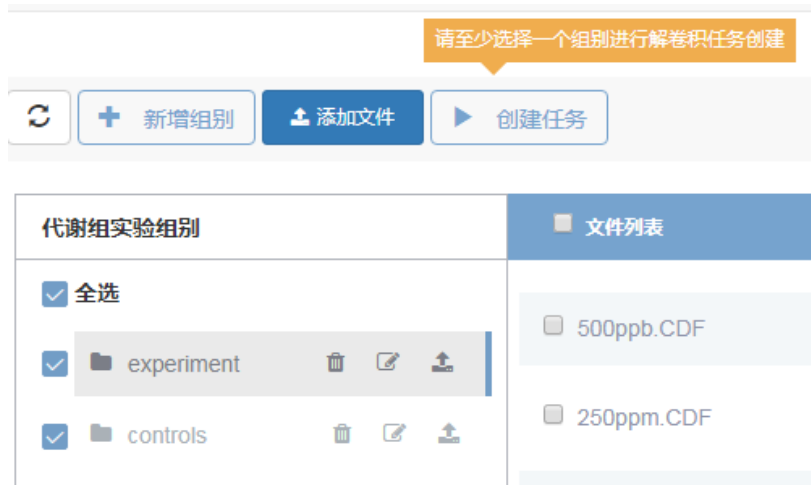
上传数据

- 4. 文件管理操作
- 在文件列表中选中文件之中，可以进行下面的三个操作：



创建任务

- 在实验组别列表之中选中需要进行分析的组别，之后点击创建任务按钮，即可开始新建一个分析任务



上传数据

设定解卷积参数

- 点击创建任务之后下一步会需要设定解卷积参数，一般直接使用默认参数即可，也可以针对自己的实验自行调整最佳参数



新建解卷积计算任务

选择预设参数: GC / Single Quad (centWave)

任务基本信息 | 特征检测 | 保留峰归组 | 保留时间校正

Xcms方法: centWave

mzCenterFun: wMean integrate: 1 fitgauss: ☐

noise: 0 verbose.columns: ☐ firstBaselineCheck: ☒

snthresh: 6 mzdiff: 0.01 ppm: 100

peakwidth: 5 — 10

prefilter: 0 — 0

创建任务

查看结果

- 任务执行成功之后，可以点击任务标题的链接在线查看结果数据，也可以点击下载结果链接下载结果至本地查看

查看任务结果： 流程示意 / T20190111-&xv-gjVi

RT 图
RT alignment

MS1
标准品库
一级质谱
搜索结果

表达量矩
阵热图
MS1
peaktable

结果下载
下载质谱
原始文件
解卷积结
果

下载质谱原始文件解卷积结果

您可以点击下面的链接进行结果文件的下载操作。您将会下载得到一个以您的任务名称命名的zip压缩包文件，在这个压缩包之中可能会包含有下列数据文件：

align.png：原始数据文件的rt比对的结果图
ms1.search.csv：一级质谱标准品库的搜库注释结果，这个表格文件之中包含有物质注释的名字以及数据库编号
heatmap.png：对样品之中的表达量数据进行热图可视化
peaktable.csv：样品文件进行解卷积操作得到的原始表格文件，里面包含有待鉴定化合物的mz/rt信息以及在不同样品间的表达量的值
data.txt：数字矩阵文件，如果您需要使用MetaDiscovery计算平台进行下一步的多元统计计算分析，那么您将会需要这个文件作为输入数据。
sampleinfo.txt：样品的分组信息，在使用MetaDiscovery计算平台进行统计分析的时候您会需要使用这个文件来进行分组信息的输入。

下载结果文件 →

T20190111-
&xv-gjVi

执行成功

2019-01-11 09:48:59

2019-01-11 09:49:03

2019-01-11 09:50:09

删除

查看参数

查看日志

下载结果

上传数据

- 选择数据
- 选择标准品数据库

创建任务

- 创建任务

结果

- 查看结果

详情参考: <http://help.biodeep.cn/>

上传数据

- http://msms.biodeep.cn/new_task.php?app=new_task

创建注释任务

第一步：设置任务属性	[-]
注释任务名称：	<input type="text" value="如果您不输入任务名称，系统将会为您自动生成！"/>
第二步：选择样本数据 *	[+]
第三步：选择peaktable *	[+]
第四步：样本选择	[+]
第五步：选择标准品库 *	[+]

创建任务

上传数据

- 选择样本数据
- 为了进行小分子化合物鉴定，首先会需要在MetaDeco之中上传实验的原始数据
- 在这里目前只能够进行LCMS项目的注释
- 也可以切换到右边创建新的
- MetaDeco项目上传原始数据

第二步：选择样本数据 *

选择样本数据

上传样本数据

Entries per Page: 10 ▼

☒ QE

☐ 流程示意

上传数据

- 选择peaktable
- 如果目标分析项目已经在MetaDeco之中成功的进行了解卷积任务，那么可以在第三步选择peaktable做全注释
- 也可以上传自己的peaktable，进行选择性的物质注释，例如进行差异代谢物注释鉴定

第三步：选择peaktable *

选择peaktable 上传peaktable

您可以在MetaDeco对本原始数据项目做完解卷积分析之后，在这里选择解卷积成功的任务作为您的样本数据之中的二级碎片的母离子信息

文件夹名称	文件数量
experiments_p	1
controls_p	1
QC_p	3
experiments_n	1
controls_n	1
QC_n	3

选择peaktable文件

☐ T20190111-AiZpLh|L

创建任务

• 样本极性设置

- 会需要在第四步中为样本分组进行设置，告诉鉴定程序每一个样本的极性模式以及加和物的类型。
- 在这一步也可以不做设置，物质鉴定程序会尝试自动识别极性模式。但是对于一些原始数据文件可能会出现自动识别失败的问题
- 用鼠标拖拽分组到对应的极性模式
- 之中即可

第四步：样本选择

Tips：您可以使用鼠标拖动所有分组中的数据拖动至阳离子组分组或者阴离子组中，以选择性分析您的样本数据；

如果您在这里不设置任何分组则程序默认自动进行离子化模式分组并分析所有的样本数据；

因为原始数据处理的原因可能有些样本原始数据文件无法得到正确的离子化模式，这种情况下您可能需要在这里手动选择分组来解决这个问题。

所有分组	阳离子组	阴离子组
<div>QC_p</div> <div>QC_n</div>	<div>experiments_p</div> <div>controls_p</div>	<div>experiments_n</div> <div>controls_n</div>
	默认加和物类型([M+H] ⁺ , [M] ⁺)	默认加和物类型([M-H] ⁻ , [M] ⁻)
	Choose option	Choose option

创建任务

- 选择标准品库

- 可以使用自己所构建的标准品库以及其他用户共享出来的标准品库进行物质鉴定操作。
- 如果使用共享标准品库，除了免费的MoNA数据库之外，其他的共享标准品库的使用可能会产生
- 相应的费用

第五步：选择标准品库 *

Tips：使用共享标准品库可能会需要向标准品库的拥有者付费，默认使用免费的MoNA标准品库。

自建标准品库

共享标准品库

☐ 使用Biodeep超级标准品库

☒ 使用MoNA公共库

按照使用人数排序

按照使用价格排序

按照收藏数排序

我的收藏

bb

创建时间：2018-10-1
4

0人使用了该标品库

正负模式测试

创建时间：2018-10-1
9

0人使用了该标品库

ceshi9

创建时间：2018-10-2
3

0人使用了该标品库

创建任务

- 创建任务
- 完成注释任务的参数设置之后可以点击最下方的按钮创建注释任务
- 可以点击任务标题的链接进行结果的查看

创建注释任务

第一步：设置任务属性

第二步：选择样本数据 *

第三步：选择peaktable *

第四步：样本选择

第五步：选择标准品库 *

创建任务



任务名称	状态	创建时间
T20190111-sNbzIvfB	执行成功	2019-01-11 11:11:11

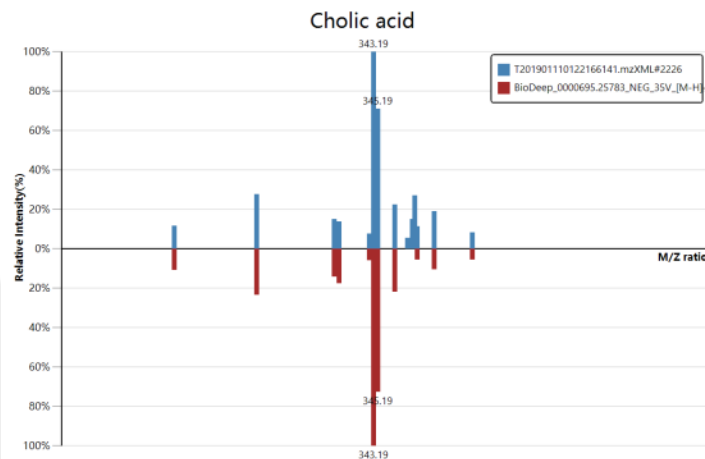
结果

- 查看注释结果
- 在结果表格之中，可以查看鉴定结果小分子化合物的基本属性，BioDeep数据库编号和鉴定的结果判定依据

#编号	BioDeep编号	分子质量	分子式	母离子	质谱图编号	标准品库	加合物	准确度	算法名称	得分
	BioDeep_000 0158	192.027	C6H8O7	191.019@16 8	BioDeep_000 0158.21537_ NEG_35V_[M -H]-	MoNA.rda	[M-H]-	MSMScomfir med	SSM	SSM: 1/1 HITS+: 2/2 PPM: 7
	BioDeep_000 5598	193.0739	C10H11NO3	192.066@43 3	BioDeep_000 5598.49242_ NEG_20V_[M -H]-	MoNA.rda	[M-H]-	MSMScomfir med	SSM	SSM: 1/1 HITS+: 1/1 PPM: 1
	BioDeep_003 2745	195.053159	C9H9NO4	194.043@461	BioDeep_003 2745.40338_ NEG_35V_[M	MoNA.rda	[M-H]-	MSMScomfir med	SSM	SSM: 0.94/1 HITS+: 1/1 PPM: 14

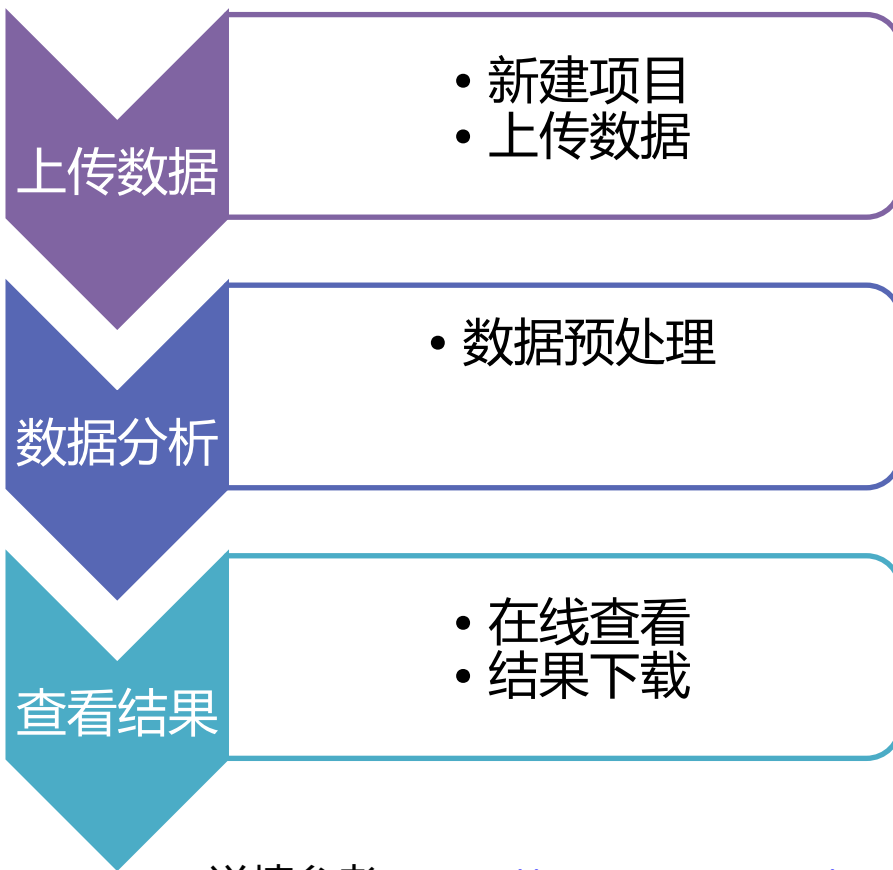
点击链接将会进行BioDeep数据库查询

点击行数据将会显示详细的二级质谱图比对结果



鉴定结果依据

#编号	BioDeep编号	分子质量	分子式	母离子	质谱图编号	标准品库	加合物	准确度	算法名称	得分
	BioDeep_0000158	192.027	C6H8O7	191.019@168	BioDeep_0000158.21537_NEG_35V_[M-H]-	MoNA.rda	[M-H]-	MSMScomfirmmed	SSM	SSM: 1/1 HITS+: 2/2 PPM: 7
	BioDeep_0005598	193.0739	C10H11NO3	192.066@433	BioDeep_0005598.49242_NEG_20V_[M-H]-	MoNA.rda	[M-H]-	MSMScomfirmmed	SSM	SSM: 1/1 HITS+: 1/1 PPM: 1
	BioDeep_0032745	195.053159	C9H9NO4	194.043@461	BioDeep_0032745.40338_NEG_35V_[M-H]-	MoNA.rda	[M-H]-	MSMScomfirmmed	SSM	SSM: 0.94/1 HITS+: 1/1 PPM: 14



详情参考: <http://help.biodeep.cn/>

上传数据

- 1. 新增分析项目
- 点击新增按钮，
- 在弹出框之中填写相应信息
- 之后保存，完成项目的创建

A screenshot of the "新建" (New) dialog box. It has a title bar with "新建" and a close button. The form contains:

- 项目编号 (Project ID): 列: HT201511271001
- 项目名称 (Project Name): 列: 小麦代谢组学 (GC-MS)
- 项目类型 (Project Type): --请选择-- (dropdown menu)
- 项目图片 (Project Image): 上传图片 (upload image button) and a green square icon with a white 'P'.
- 项目描述 (Project Description): 项目描述 (text area)

At the bottom right are "关闭" (Close) and "保存" (Save) buttons.

上传数据

- 上传计算数据
- 依次点击分析文件，新增按钮，可以打开分析用的数据文件的上传窗口
- 之后还会需要依次上传样本的分组信息和代谢物注释结果信息（可选）



数据格式参考：<http://help.biodeep.cn/#metadiscovery/1.7dataformat.md>

数据分析

- 数据预处理
- 所上传的计算数据需要进行预处理才能够得到比较好的分析结果，在BioDeep之中需要首先完成预处理才能够进入后续的统计分析。
- 一般情况下，预处理阶段直接使用系统的默认值即可。

步骤1. 去除太多缺失值的变量

☒ 去除超过 %缺失值的变量

步骤2. 预测缺失值

☒ 替换成一个小值 (原始数据中最小阳性值的一半)

确定

批次归一化

☒ 不进行归一化

☐ loess ?

归一化

☐ 不进行归一化

☒ 总峰面积归一化 ?

☐ 中位数归一化 ?

数值转化

☒ 不转化

☐ Log转化 ?

☐ 立方根变换 ?

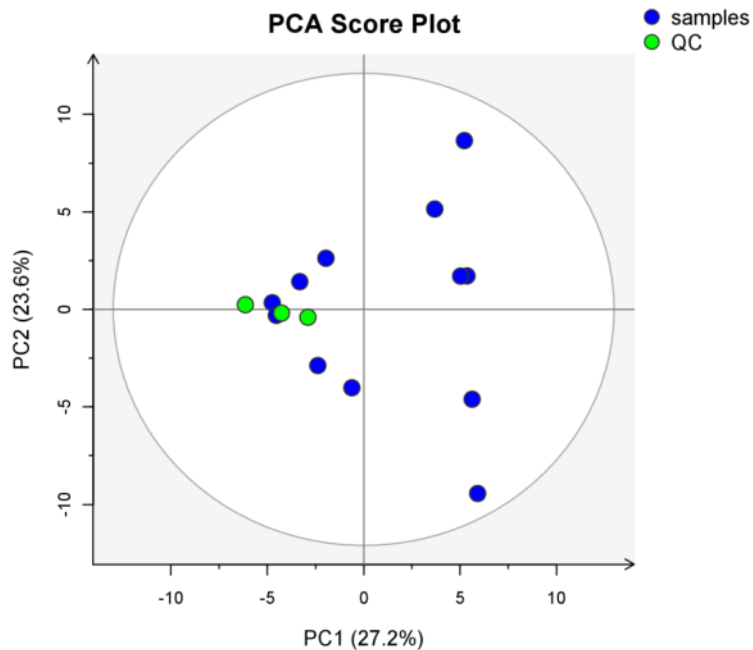
数据分析

- 质控分析
- 质控分析可以查看仪器是否稳定

☐ 不进行QA分析 ☒ 进行QA分析

预览

确定



数据分析

- 在后续的数据分析之中，大部分的项目都是实验组之间的比较统计分析操作，完成了数据的预处理之后，会首先需要选择需要进行分析的分组，才能够进入后续的统计分析之中

3. 矩阵操作



矩阵操作



4. 多元统计分析



5. 差异代谢物分析



6. 可视化分析



7. 聚类分析



8. 功能分析



9. 生物标记物分析



我的选择组别列表

拖动组内排序



A

CK

确定

点击选择需要分析的比较组别

点击确定进行保存

数据分析

- 多元统计分析
- 多元统计分析中PCA，PLS-DA，OPLS-DA分析都需要首先进行自动拟合，然后就可以切换标签页进行结果图以及表格的下载操作



数据分析

- 差异代谢物分析
- 差异代谢物分析包括一个比对组内的差异代谢物计算（火山图）和不同的比对组间的差异代谢物分析（文氏图）。文氏图分析会要求所上传的数据之中至少具有两个比对组别。

5. 差异代谢物分析



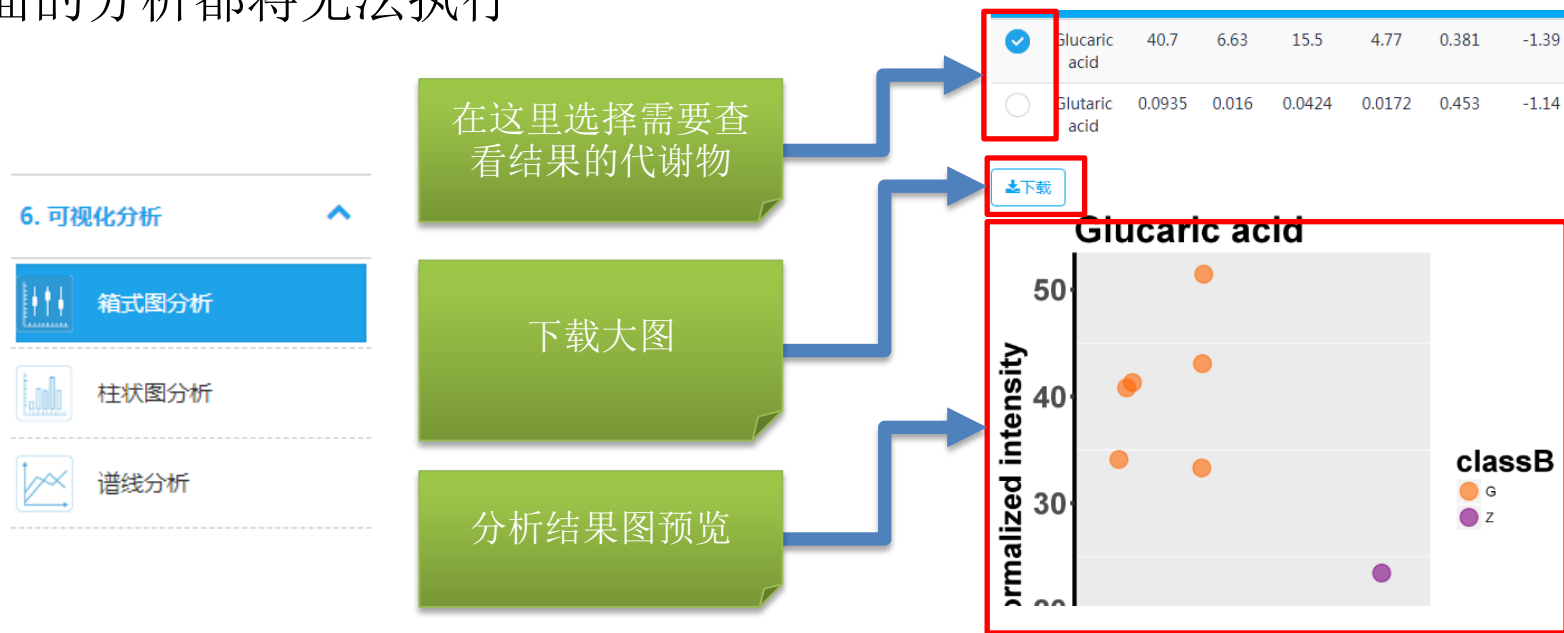
一般情况下直接使用
系统默认参数计
算即可

检验方式 学生氏T检验 配对检验 ☐ 比较方式 A/CK

确定

数据分析

- 差异代谢物结果可视化分析
- 进一步的可视化分析是建立在差异代谢物分析结果的基础上的。如果没有在上一步执行差异代谢物分析，或者没有差异代谢物结果，则后面的分析都将无法执行



数据分析

- 聚类分析
- 聚类分析则可以查看了解样本之间以及代谢物之间的关联
- 在这个小节的分析中可以选择分析所有的代谢物还是前面所计算出来的差异代谢物之间的关联

用户参数选择：

全部代谢物

差异代谢物

下载

7. 聚类分析

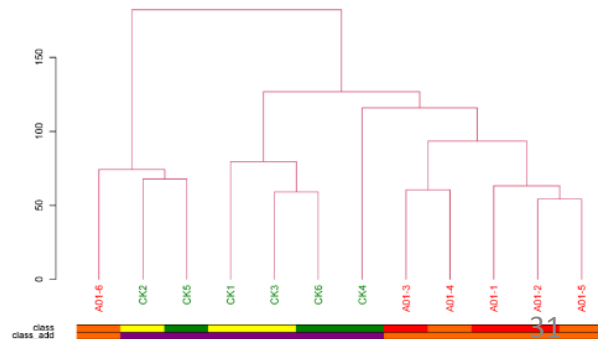
热图分析



代谢物关联分析



层次聚类分析



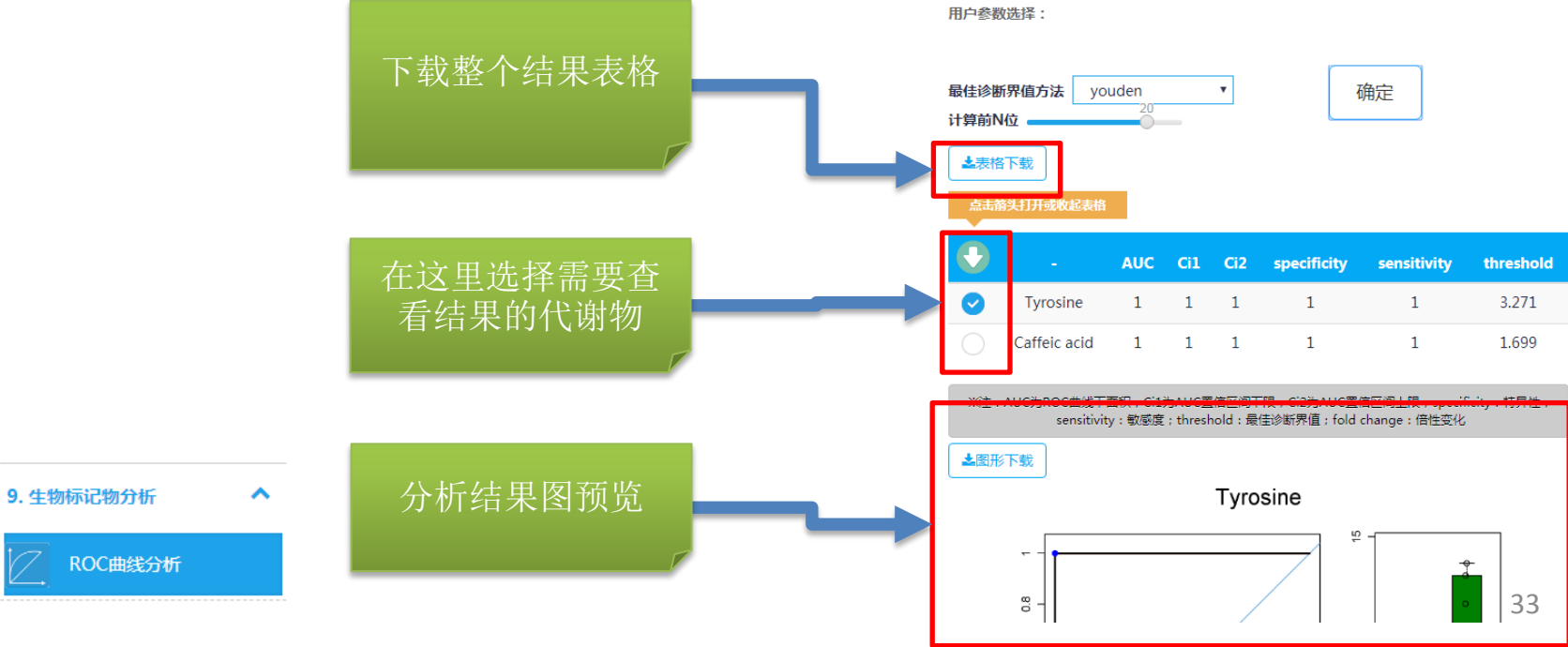
数据分析

- 生物学功能分析
- 生物学功能分析，主要是进行和KEGG代谢通路，以及KEGG代谢反应网络相关的分析操作。（生物学功能分析需要在上传可选的代谢物meta信息文件，如果没有上传，将无法进行这部分的分析操作。）



数据分析

- 生物标记物分析
- 计算能够特异性的指示某种生物性状变化的代谢物。



数据分析

<div>任务名称搜索</div>				
任务名称	状态	开始时间	结束时间	操作
公共流程1-->metabolome.txt	执行完毕	2018-11-20 03:43:32	2018-11-20 03:47:48	<div><div>删除</div><div>下载</div><div>查看日志</div><div>查看在线报告</div></div>
<div>← 上一页 第1页/共1页 下一页 →</div>				

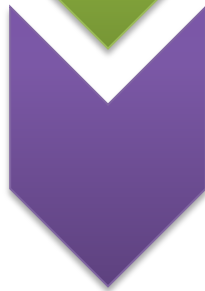
- 下载结果，查看所有分析内容



- 创建标准品库项目



- 上传标准品库原始数据文件
- 上传对应的注释信息
- 编辑构建任务参数

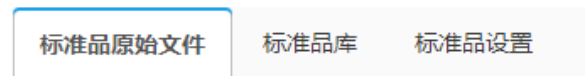


- 标准品库自动化构建成功

- 标准品库操作
- 可以点击标准品库的标题进入标准品库的详细信息查看页面，在这里你可以进行：
 1. 编译标准品库：从原始数据自动化构建标准品库
 2. 删除标准品库
 3. 查看标准品库原始文件列表
 4. 查看标准品库详细内容
 5. 设置标准品库属性信息



• 上传标准品库原始数据文件



您可以点击文件名查看您的标准品在BioDeep数据库之

□如果您不希望进行此人工确认操作，则标准品库编译和



上传标品库文件

注释信息文件:

浏览...

开始上传

要求您所上传的注释信息文件中，应该要包含有您所上传的原始文件名以及其所对应的物质注释信息。您可以下载这个物质的注释信息Excel模板文件进行注释信息的填写。

上传原始数据文件

+ 添加文件...

开始批量上传

- 上传标准品库原始数据文件
- 需要上传两部分的数据文件用来进行自动化构建标准品库
- 1. 注释信息文件：用来指明该物质是什么
- 2. 原始数据文件：用来指明该物质的具体的二级质谱碎片结构



- 当所有填写的信息都检查通过之后，就可以开始上传标准品的原始数据文件了。
- 上传完成后，可以点击最下方的查看标准品库进行进一步的编辑操作。

上传标准品文件

上传原始数据文件

+ 添加文件... ① 开始批量上传

137.37 Mbit/s | 00:00:00 | 100.00 % |
14.51 MB / 14.51 MB

21.mzXML	2.67 MB	取消
37.mzXML	1.61 MB	取消
38.mzXML	2.94 MB	取消
86.mzXML	2.96 MB	取消
87.mzXML	4.34 MB	取消

查看标准品库

注释信息文件: metainfo-template.xlsx

浏览...


开始上传

要求您所上传的注释信息文件中，应该要包含有您所上传的原始文件名以及其所对应的物质注释信息。 您可以下载这个物质的注释信息Excel模板文件进行注释信息的填写。

原始数据文件名	化合物名称	CAS	ChEBI	HMDB	KEGG	检查结果
21.mzXML	gamma-aminobutyric acid zwitterion	56-12-2	59888	HMDB0000112	C00334	格式检查通过
37.mzXML	glycine	56-40-6	15428	HMDB00123	C00037	格式检查通过
38.mzXML	L-methionine	63-68-3	16643	HMDB00696	C00073	格式检查通过
86.mzXML	thiourea	62-56-6	36946	HMDB34155	C14415	格式检查通过
87.mzXML	diethanolamine	111-42-2	28123	HMDB04437	C06772	格式检查通过

- 确定物质注释信息
- 点击原始文件名链接，将会执行BioDeep标准品数据库查询
- 可能会存在多个相似的注释信息，这个时候会需要点击弹出框的物质名称列的选择框来确定最佳的注释信息
- 最后点保存完成操作

原始文件名	已分配BioDeepID
37.mzXML	BioDeep_0000037
38.mzXML	未分配
21.mzXML	未分配
86.mzXML	未分配
87.mzXML	未分配



匹配标准品注释信息

name:glycine

BioDeepID:

CAS:56-40-6

HMDB:HMDB00123

KEGG:C00037

ChEBI:15428

如果您确认您所上传的物质为新物质，并且BioDeep数据库之中没有您的物质的注释信息的话，请点击[保存为新物质]按钮进行保存。

物质名称	分子质量	BioDeep编号	CAS	KEGG	HMDB	ChEBI	匹配度
<input type="checkbox"/> Glycine(C2H5NO2)	75.032028409	BioDeep_0000037	56-40-6	C00037	HMDB0000123	15428	5

保存为新物质

取消

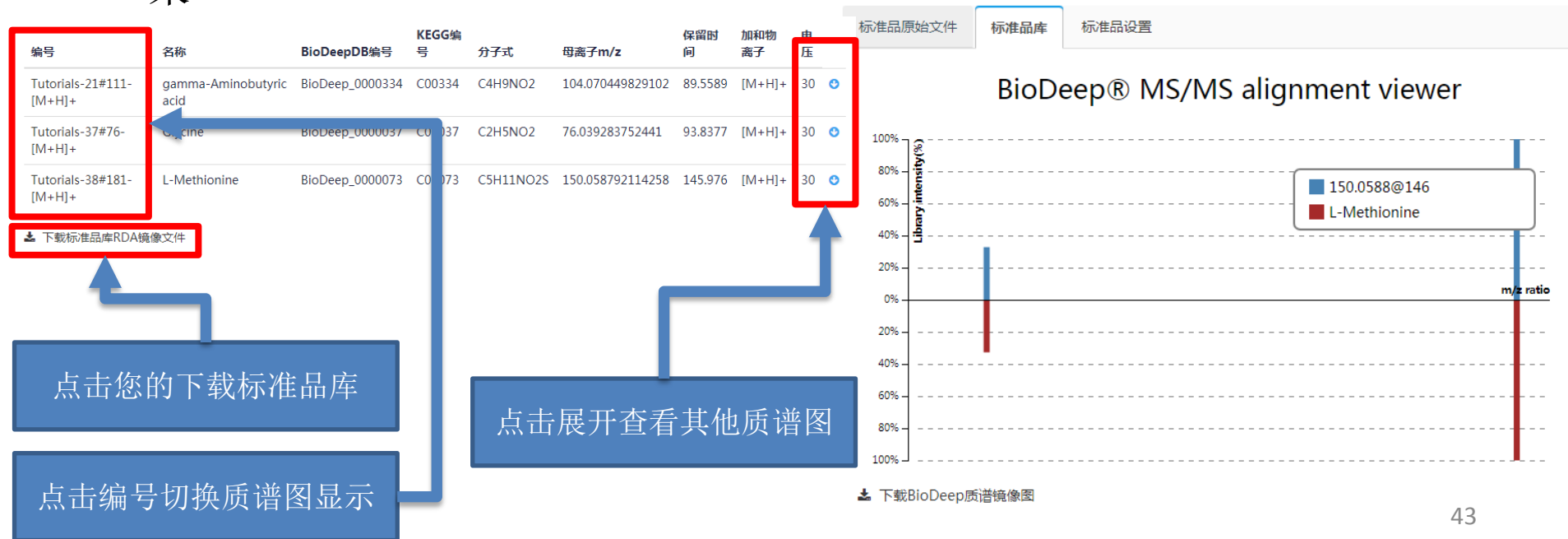
保存

- 自动化构建标准品库
- 点击编译标准品库按钮，将会跳转到下一步设置一些参数信息
- 之后需要填写任务参数，最后点击开始编译执行标准品库的自动化构建操作

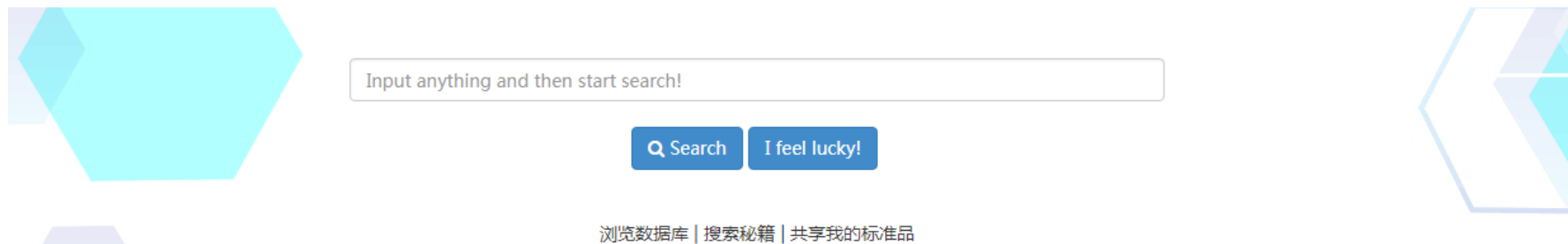


查看标准品库

- 完成自动化构建任务之后，可以至标准品库详细信息页面查看构建结果



- <http://msms.biodeep.cn/biodeepDB.php?app=index>



详情参考: <http://help.biodeep.cn/>

- 搜索入口介绍



- 搜索帮助可以点击搜索秘籍链接了解详情

点击链接可以打
开标准品质谱图
查看页面

返回 | **Cholic acid**

点击链接可以按
照分子质量或者
分子式查询

BioDeep_0000695 (81标准品质谱图谱, 查看所有质谱图)

► Synonym:

点击链接可以跳
转到外部数据库
查看详细信息

MF: C₂₄H₄₀O₅

MW: 408.2876 g/mol (m/z Calculator)

CAS: 81-25-4

InChI: InChI=1S/C₂₄H₄₀O₅/c1-13(4-7-21(28)29)16-5-6-17-22-18(12-20(27)24(16,17)3)23(2)9-8-15(25)10-14(23)11-19(22)26/h13-20,22,25-27H,4-12H2,1-3H3,(H,28,29)/t13-,14+,15-,16-,17+,18-,19-,20+,22+,23+,24-/m1/s1

InChIKey: BHQCFFYRZLCQQ-ELDTZBJS-A-N

HMDB: HMDB0000619

chebi: 16359

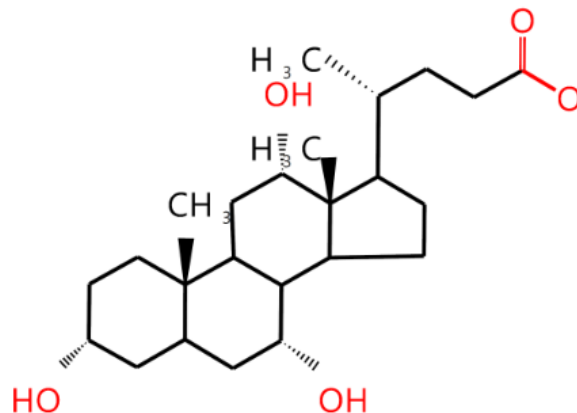
KEGG: C00695

metlin: 206

pubchem: 3963

source: 3

class: Compounds with biological roles/Steroid



点击链接可以打
开m/z计算器页面



WWW.BIODEEP.CN

Search

THANK YOU !



苏州工业园区星湖街218号
生物纳米科技园A4楼201~206室



0512-62959105

