

使用第三方开源包增强 MPP 的分析能力：在 MPP 的 R 脚本编辑器中进行数据读取、操作和分析

作者

于擎
安捷伦科技（中国）有限公司

引言

Agilent Mass Profiler Professional 软件凭借先进的数据处理功能、强大的统计方法和数学模型，适用于对复杂的质谱数据进行统计分析。同时，MPP 内置有脚本编辑器，用户能够编写自己的脚本。借助这些脚本，可以在 MPP 中应用目前 MPP 软件本身尚不支持的算法，对数据进行一系列复杂操作，甚至可以每次使用略有不同的参数来重复运行并评估算法。POMA^[1] 是一个比较受欢迎的代谢组学数据处理 R 开源包，它引入了结构化、可重复且易于使用的工作流程，适用于质谱数据的可视化、预处理、探索性和统计分析。利用 MPP 软甲的 R 脚本编辑器，可以在 MPP 中使用 POMA 的部分功能，对导入 MPP 的数据进行插补、变换、去除离群点以及统计分析等一系列操作；并通过索引对应每种化合物的精确质量数和同位素分布等信息，方便利用安捷伦提供的庞大的 PCDL 数据库和谱库对经过处理的数据中的相应化合物进行定性分析。

数据及分析软件

- 1. Agilent Mass Profiler Professional 15.1
- 2. R 4.2.3
- 3. RStudio 2023.03.0+386 "Cherry Blossom" Release（可选）
- 4. 数据：来自 MPP 培训用数据，分为两组，每组三个样本。两组样本本身无差异，差异来自于人为的外源性化合物添加

原始数据的读取和操作

- 1. 在 MPP 中，与 R 对接的函数是 `getDataset()`。这个函数可以从选中的实体列表 (Entity List) 中读取数据，用于后续的数据操作和分析。需要注意的是，此函数仅在 MPP 的 R 脚本编辑器中可用。在进行数据操作之前，加载 `tidyverse` 包^[2] 以便后续的数据操作。

```
library(tidyverse)
```

使用 MPP 中的内置函数 `getDataset()` 读取数据集：

```
df = getDataset()
```

如果在 MPP 的 R 脚本编辑器中运行 `df`，即可看到 `df` 的输出。这里使用的是从 All Entities 中直接导出的数据：

```
MPP_data = read_csv("MPP_raw_data.csv")
```

表 1. 存储在 MPP 中的数据

化合物	Group1_	Group1_	Group1_	Group1_	Group1_	Group1_
	urine_	urine_	urine_	urine_	urine_	urine_
	Mix_A_1	Mix_A_2	Mix_A_3	Mix_B_1	Mix_B_2	Mix_B_3
263.191@6.162	21.90	21.85	21.89	21.90	21.87	21.88
439.2208@5.7410007	21.51	21.54	21.46	21.51	21.47	21.47
314.1397@5.781	21.56	21.58	21.57	21.56	21.56	21.56
113.0592@0.664	20.62	20.57	20.58	20.53	20.58	20.59
93.0579@0.82399994	20.64	20.02	20.59	19.92	20.59	20.60
288.2089@8.294	0.00	0.00	0.00	20.21	20.20	20.23

表 1 展示了读取到的存储在 MPP 中的数据。此时读取的数据并非“原始”数据，而是经过 $\log_2$ 变换后的数据。这种变换发生在归一化步骤中，且不能取消。有时研究人员可能想要使用原始数据进行分析，这时只需要对读取的数据进行幂运算，即可将其还原为原始数据。需要注意的是，MPP 的 R 脚本编辑器中没有 NA，所有缺失值都以 0 表示（即使在 Profinder 的导出选项中选择将缺失作为缺失处理时也是如此）。实施幂运算后，缺失值将会被 1 取代，这时 R 不再认为此处存在缺失。为了避免这种情况，可以在幂运算后将所有数值为 1 的值替换为 0，以保证数据的原始性。为方便起见，仅显示数据的前六行，后面也采用相同的显示方式。

除此之外，此处使用的练习数据还存在一个明显的问题，即样本数过少（每组中仅包含三个样本）。为提高流程的通用性，这里基于原始数据中每种化合物的均值和标准差创建了一套虚拟数据。虚拟数据与原始数据组合后，将总样本数提升至 40，即每组包含 20 个样本。

```
df_raw = read_csv("MPP_raw_data.csv")
df_t = df_raw %>%
  column_to_rownames(var = "Compound") %>%
  t() %>%
  as.data.frame() %>%
  rownames_to_column(var = "Group") %>%
  mutate(Group = str_sub(Group, 1, 18)) %>%
  as_tibble()
df_group = df_t %>%
  group_split(Group)
df_grp_A = df_group[[1]] %>%
  select(-1)
df_grp_B = df_group[[2]] %>%
  select(-1)

vir_A = data.frame(matrix(nrow = 17, ncol = 212))
for (i in seq_along(df_grp_A)) {
  vir_A[[i]] = rnorm(17, mean = mean(df_grp_A[[i]]), sd = sd(df_grp_A[[i]]))
  vir_A
}
colnames(vir_A) = colnames(df_grp_A)
group_vir_A = tibble("Group" = rep("Group1_urine_Mix_A", 17))
vir_A_df = bind_cols(group_vir_A, as_tibble(vir_A))
vir_A_mix_pre = bind_rows(df_group[[1]], vir_A_df)
vir_A_mix = vir_A_mix_pre %>%
  mutate(ID = row_number()) %>%
  mutate(Group = str_c(Group, ID, sep = "_")) %>%
  select(-ID)

vir_B = data.frame(matrix(nrow = 17, ncol = 212))
for (i in seq_along(df_grp_B)) {
  vir_B[[i]] = rnorm(17, mean = mean(df_grp_B[[i]]), sd = sd(df_grp_B[[i]]))
  vir_B
}
colnames(vir_B) = colnames(df_grp_B)
group_vir_B = tibble("Group" = rep("Group1_urine_Mix_B", 17))
vir_B_df = bind_cols(group_vir_B, as_tibble(vir_B))
vir_B_mix_pre = bind_rows(df_group[[2]], vir_B_df)
vir_B_mix = vir_B_mix_pre %>%
  mutate(ID = row_number()) %>%
  mutate(Group = str_c(Group, ID, sep = "_")) %>%
  select(-ID)

virtual_data = bind_rows(vir_A_mix, vir_B_mix) %>%
  column_to_rownames("Group") %>%
  t() %>%
  as.data.frame() %>%
  rownames_to_column("Compound") %>%
  as_tibble()
```

新生成的虚拟数据如表 2 所示，第一列为化合物信息，其后的每一列显示的是在对应条件下各个化合物的强度信息。下面将使用此组合数据进行分析。

对组合数据进行幂运算，结果见表 3，将数据还原为原始强度。

```
df = virtual_data %>%
  mutate(across(where(is.numeric), ~2^.x)) %>%
  mutate(across(where(is.numeric), ~replace(.x, .x == 1, 0)))
```

为了使处理后的数据能够在 MPP 中形成实体列表，需要生成数据索引。此后，新生成的数据都可以通过左连接函数 left_join() 从该索引中获取它们在原始数据中对应的化合物信息，以便生成适用于定性分析的实体列表。首先生成数据的原始索引，即原始数据中化合物的位置信息，如表 4 所示。

```
raw_index = df %>%
  select(Compound) %>%
  mutate(ID = row_number())
```

表 4. 原始索引

化合物	ID
263.191@6.162	1
439.2208@5.7410007	2
314.1397@5.781	3
113.0592@0.664	4
93.0579@0.82399994	5
288.2089@8.294	6

在对原始数据进行操作以后，需要读取数据的解释信息（即，将要如何对数据进行研究）。MPP 提供了 getExperimentGroupingDataset() 函数，可以读取选中的（通过鼠标单击）解释信息。需要注意的是，此函数仅在 MPP 的 R 脚本编辑器中可用。

```
group = getExperimentGroupingDataset()
```

由于 getExperimentGroupingDataset() 函数仅能在 MPP 内部运行，因此这里编写了一套与原始数据对应的分组信息（表 5），用于数据处理：

```
group = read_csv("group.csv")
```

表 5. 分组信息

样本	条件名称
Group1_urine_Mix_A_2	A
Group1_urine_Mix_A_3	A
Group1_urine_Mix_A_4	A
Group1_urine_Mix_A_5	A
Group1_urine_Mix_A_6	A

在得到处理后的原始数据以及分组信息后，即可创建 SummarizedExperiment^[3]，以便于使用 POMA 包对数据进行处理。可以通过多种方法创建 SummarizedExperiment，此处使用 QFeatures 包^[4]中的 readSummarizedExperiment() 函数来完成这一过程：

```
library(QFeatures)
library(SummarizedExperiment)
df_sum = readSummarizedExperiment(df, ecol = 2: 41, fnames = "Compound") #需
要根据实际数据的样本数量进行更改为 2: x (x 为实际样本数 + 1)
df_sum$group = group$Condition.Name
df_sum

## class: SummarizedExperiment
## dim: 212 40
## metadata(0):
## assays(1): ''
## rownames(212): 263.191@6.162 439.2208@5.7410007 ... 121.918@0.575
## 104.9927@0.519
## rowData names(1): Compound
## colnames(40): Group1_urine_Mix_A_1 Group1_urine_Mix_A_2 ...
## Group1_urine_Mix_B_19 Group1_urine_Mix_B_20
## colData names(1): group
```

数据的预处理

缺失值插补

创建 SummarizedExperiment 对象后，便可以使用 POMA 包对数据进行分析。首先要处理的是数据中的缺失值。缺失值在代谢组学数据中十分常见，当数据通过 auto MS/MS 采集获得时，这种现象更加明显。对于代谢组学数据，一般使用 knn (K-近邻) 方法对缺失值进行插补。需要明确的是，对于本文使用的数据，其样品差异来自于基质添加，因此这里的缺失值是由外源添加导致的，并非真正的数据缺失，缺失值插补反而会使数据失真。下面仅出于展示数据分析流程的目的而进行缺失值插补。

PomalImpute() 函数可用于对 SummarizedExperiment 对象进行缺失值插补，ZerosAsNA 参数可以将 0 作为缺失值处理。缺失值插补方法可以在 method 参数中进行更改，可选方法包括 none、half_min、median、mean、min、knn 和 rf。如果使用“none”，则所有的缺失值都会被替换为 0，在本例中相当于不做处理。需要注意的是，在对缺失值进行插补之前，需要对缺失数据进行评估，以确认是否有必要以及以何种形式进行插补。

```
library(POMA)
imputed = PomaImpute(df_sum, ZerosAsNA = T, cutoff = 20, method = "knn", RemoveNA = T)
df_imputed = assay(imputed) %>%
  as.data.frame() %>%
  rownames_to_column(var = "Compound") %>%
  mutate(Compound = str_replace(Compound, "(X)(\\d+\\.\\d+\\.\\d+)\\.\\d+(\\d+\\.\\d+\\.\\d+)",
    "\\2@\\3")) #应用函数过程会更改 Compound 变量的值，需要进行文本替换
imputed_index = left_join(df_imputed, raw_index, by = "Compound") #获取索引
```

插补后的结果详见表 6，从中可以看到缺失值得到了插补。值得注意的是，由于缺失值的比例达到 50%，因此实际插补方法变成了“mean”。接下来，便要将插补后的结果生成 MPP 中可用的实体列表。这里可以使用 MPP 自带的与 R 的连接函数 `addResultColumns()` 将结果生成在 Analysis 条目下的实体列表中。此函数的主要参数是一个数据框，由通过前述方法生成的数据以及该数据框中的化合物在原始数据中的索引组成。这里同样需要注意，此函数仅在 MPP 的 R 脚本编辑器中可用。通过索引，所有化合物的保留时间、m/z 和同位素等信息都可以与其处理后的强度信息对应起来，支持直接使用 ID Browser 进行数据库/谱库检索。

通过以下操作，将在 All Entities 条目下生成一个新的实体列表。可以通过双击该实体列表，将其重命名（如 Imputation Result），以便于后续使用。

```
addResultColumns(data.frame(imputed_index$ID, df_imputed))
```

数据变换

常用的假设检验方法一般要求数据满足两个基本条件：(1) 因变量的残差符合或大致符合正态分布（可以简化地认为因变量符合正态分布）；(2) 因变量的方差基本相等。这里需要澄清的是，并非样本量超过一定的量，数据就会呈正态分布。中心极限定理并非针对原始数据，而是针对统计量。无论原始数据的分布如何，从原始数据中进行多次抽样，得到多个统计量（如样本均值），这些统计量的分布接近正态，而与原始数据无关。

一般代谢组学实验得到的数据中的各个变量可能无法满足上述假设，因此需要进行变换以大致满足假设检验的需求。由于之前已经说明了数据的特殊性，因此本文将使用未经插补的数据进行后续分析。

```
normalized = PomaNorm(df_sum, method = "log_transformation")
df_nor = assay(normalized) %>%
  as.data.frame() %>%
  rownames_to_column(var = "Compound") %>%
  mutate(Compound = str_replace(Compound, "(X)(\\d+\\.\\d+\\.\\d+)\\.\\d+(\\d+\\.\\d+\\.\\d+)",
    "\\2@\\3"))
nor_index = left_join(df_nor, raw_index, by = "Compound")
```

数据变换的结果详见表 7。关于数据变换方法的详细信息及其优缺点，在文献^[5]中有详细的论述。我们可以按照 POMA 官方教程，查看原始数据的分布以及实施数据变换之后的数据分布情况。首先查看原始数据的分布情况，如图 1 所示；然后，查看经过数据变换后的数据分布情况，如图 2 所示；两者对比，可以看到数据分布发生了明显改变。

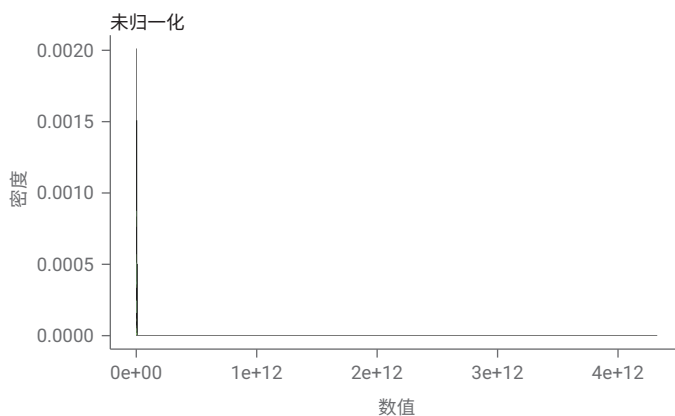


图 1. 未经变换的数据分布

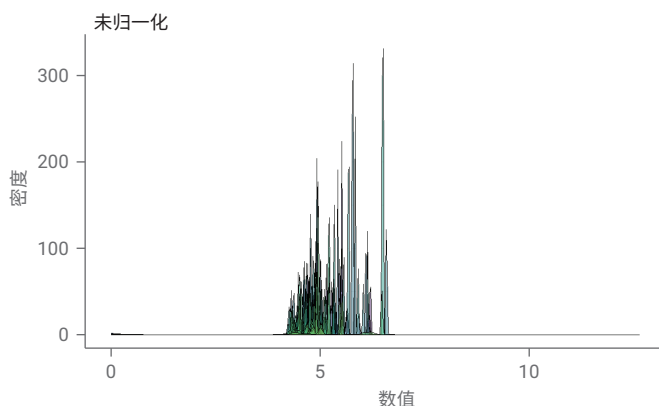


图 2. 经过变换后的数据分布

通过以下操作，将 All Entities 条目下生成一个新的实体列表。可以通过双击该实体列表，将其重命名（如 Normalized Result），以便于后续使用。

```
addResultColumns(data.frame(nor_index$ID, df_nor))
```

剔除离群点

数据预处理的最后一步是剔除离群点（异常值）。离群点被定义为与绝大多数剩余数据点的观察结果不一致的数据点，它们会对最终的统计分析产生巨大影响。首先可以通过多边形的方式查看是否存在离群点，如图 3 所示。

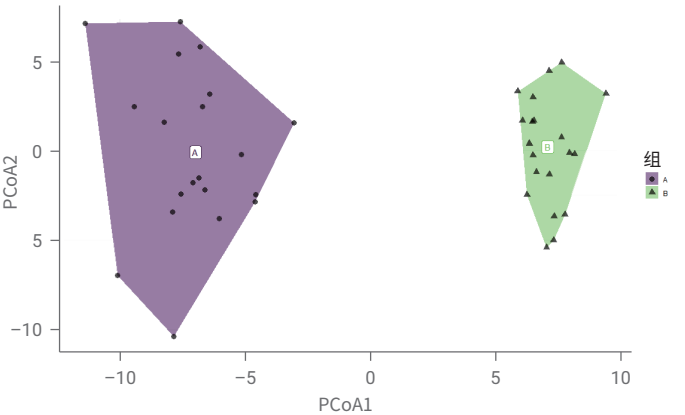


图 3. 查看离群点信息

可以看到，在这份练习数据中，由于数据构成的特殊性，不存在离群点。但是为了完整展示数据分析流程，下面通过 PomaOutliers() 函数进行离群点清除。清除离群点后的结果详见表 8（结果并无变化）。

```
pre_preprocessed = PomaOutliers(normalized, do = "clean")
df_pre_preprocessed = assay(pre_preprocessed) %>%
  as.data.frame() %>%
  rownames_to_column(var = "Compound") %>%
  mutate(Compound = str_replace(Compound, "(X)(\\d+\\.\\d+)\\.\\. (\\d+\\.\\d+)",
    "\\2@\\3")) %>%
  as_tibble()
pre_preprocessed_index = left_join(df_pre_preprocessed, raw_index, by = "Compound")
```

从结果来看，本文所使用的练习数据中无任何数据点被作为离群点删除。此结果可以作为数据预处理的最终结果，通过 addResultColumns() 函数写入实体列表中。

```
addResultColumns(data.frame(pre_preprocessed_index$ID, df_pre_preprocessed))
```

拓展统计学分析方法

MPP 中内置有常见的假设检验方法。这里补充一种在表观基因组学和转录组学中广泛使用的统计方法 — limma^[6]。limma 是一种经验贝叶斯方法，其借用跨特征的信息来评估标准误并计算 t 统计量。当样本数量较少时，limma 比标准 t 检验更为有效。表 9 展示了 limma 的分析结果。

```
limma_res = PomaLimma(pre_preprocessed, contrast = "A", adjust = "fdr") %>%
  filter(adj.P.Val < 0.05) %>%
  filter(logFC > 2|logFC < -2) %>%
  mutate(feature = str_replace(feature, "(X)(\\d+\\.\\d+)\\.\\. (\\d+\\.\\d+)",
    "\\2@\\3"))
limma_res_index = left_join(limma_res, raw_index, by = c("feature" = "Compound"))
```

表 9. limma 的分析结果

特征	logFC	AveExpr	t	P 值	调节后 P 值	B
314.1397@5.781	6.493632	6.492128	11258.527	0	0	269.1253
245.1796@6.162	5.776789	5.775282	9674.468	0	0	265.2466
277.2022@6.6889997	5.675316	5.675282	9356.306	0	0	264.3390
313.2242@7.166	5.336158	5.333308	8150.004	0	0	260.4098
453.2309@6.6409993	4.949421	4.950692	6356.260	0	0	252.6873
249.1723@6.055	5.832526	5.831154	6313.315	0	0	252.4668

同样，使用 addResultColumns() 函数将结果作为实体列表写入 All Entities 下面。

```
addResultColumns(data.frame(limma_res_index$ID, limma_res))
```

多元分析

MPP 中已经包含多种多元分析方法。这里通过 POMA 包补充一种对于特征选择很有帮助的算法 — LASSO（最小绝对收缩和选择算子）^[7]。LASSO 通过正则化（引入惩罚项），将弱相关变量删除。如果将参数 ntest 设置为 NULL，即可用于特征选择。

```
lasso_res = PomaLasso(pre_preprocessed, alpha = 1, labels = T, ntest = NULL)
lasso_coef = lasso_res$coefficients %>%
  filter(str_detect(feature, "X")) %>%
  mutate(feature = str_replace(feature, "(X)(\\d+\\.\\d+)\\.\\. (\\d+\\.\\d+)",
    "\\2@\\3"))
```

经过筛选后的特征如表 10 所示。

```
lasso_coef_index = left_join(lasso_coef, raw_index, by = c("feature" = "Compound"))
```

表 10. 经过筛选后的特征（去除截距）

特征	系数
288.2089@8.294	1.7390254
326.1644@8.3	0.0000291
614.3725@8.294	0.0000256

由于所用练习数据的特殊性，以上结果仅用于展示分析过程的目的。LASSO 的分析结果同样可以通过函数 addResultColumns() 导入实体列表中。

```
addResultColumns(data.frame(lasso_coef_index$ID, lasso_coef))
```

脚本使用方法

1. 从 R 的官方网站下载最新版 R 软件，可以直接使用 R GUI，也可以将 R GUI 与 RStudio 配合使用。在 R 控制台中输入或粘贴以下代码，以安装所需的开源包：

```
install.packages("tidyverse")
if (!require("BiocManager", quietly = TRUE))
  install.packages("BiocManager")
BiocManager::install(c("QFeatures", "SummarizedExperiment", "POMA"))
```

2. 从开始菜单中，依次单击开始 > Agilent > Mass Profiler Professional 15.1（不同版本号的 MPP 软件此处会有版本号的差异，MPP 15 以下版本需要在 http://basil.strandls.com/downloads/Script_plg_file/ 下载 ScriptEditor.plg 文件，在 <MPP Installation Directory>/bin/packages/marray/product/4.x 目录下新建名为 plugins 的文件夹，将 ScriptEditor.plg 文件复制至该文件夹。同时需要注意，在 MPP 15 以下版本中，脚本只能运行一次 addResultColumns() 函数。建议将 MPP 升级至最新版本），打开 MPP 软件。在创建新实验时，选择 Data Import Wizard，仅导入数据，如图 4 所示。

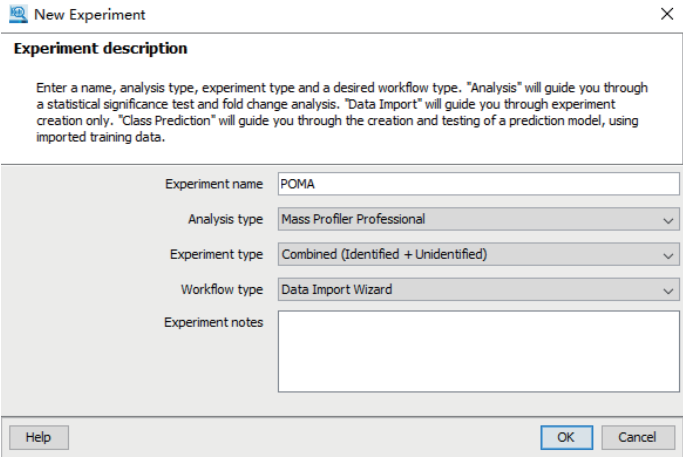


图 4. 创建新实验

3. 完成导入数据、分组、过滤和对齐操作后，转至归一化步骤。这里可以选择合适的数据归一化方法，用于消除前处理和进样中产生的误差，如图 5 所示（本文选择“none”）。

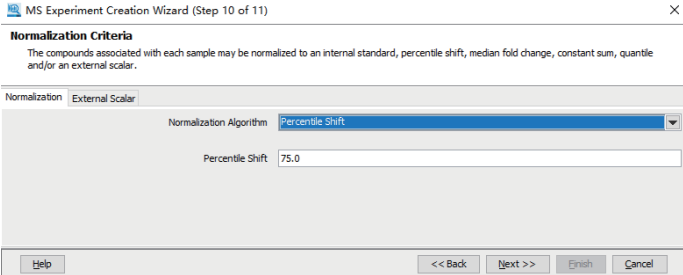


图 5. 数据归一化

4. 在数据缩放步骤中，如果希望在后续处理中使用 POMA 中的数据缩放方法，可以如图 6 所示选择“none”（即不进行数据缩放）。

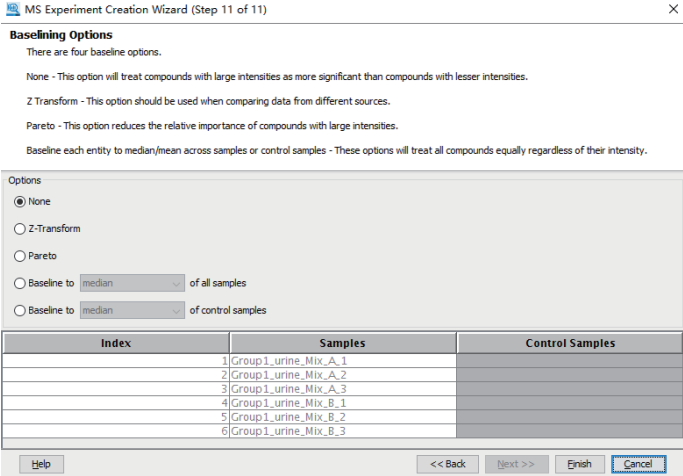


图 6. 数据缩放

5. 完成数据导入后，需要在 workflow 中创建解释信息，这些信息可以被视为 MPP 中的分组信息。单击 Workflow 中的 Experiment Setup > Create interpretation，在弹出的对话框中创建解释信息，如图 7 所示。

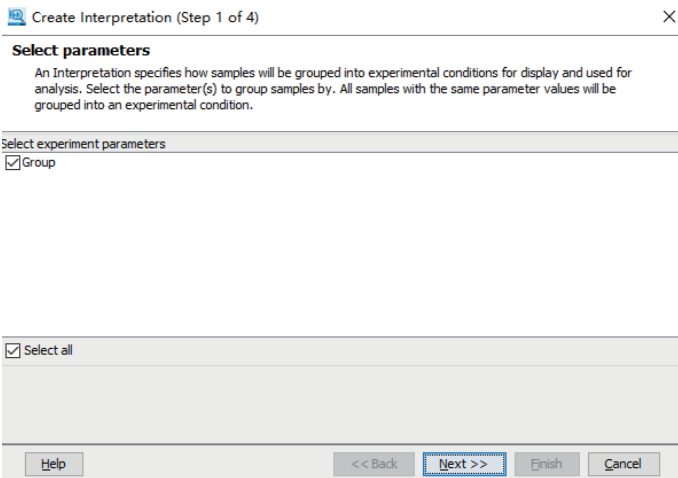


图 7. 创建解释信息

6. 在 MPP 软件中建立与 R 软件的关联。在 MPP 菜单栏中选择 Tools > Options > Miscellaneous > R Integration，然后在右侧 R Path 栏中输入 R 软件的安装目录，定位到“bin”，如图 8 所示。

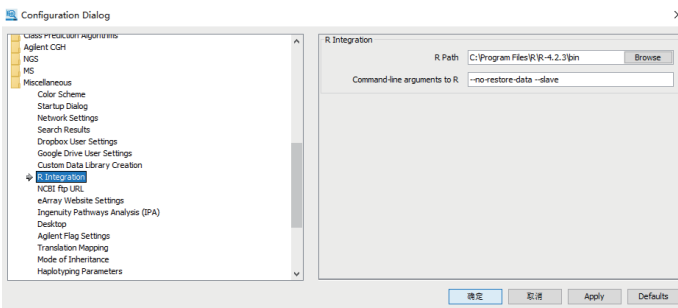


图 8. 定位 R 的位置

7. 在 MPP 中，选择 Tools > R Editor，将下面的代码复制到脚本编辑器中，如图 9 所示。

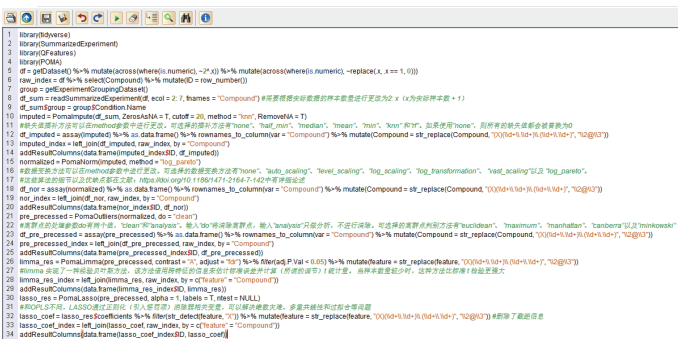


图 9. 将脚本复制到 R 脚本编辑器中

```
library(tidyverse)
library(SummarizedExperiment)
library(QFeatures)
library(POMA)

df = getDataset() %>%
  mutate(across(where(is.numeric), ~2^.x)) %>%
  mutate(across(where(is.numeric), ~replace(.x, .x == 1, 0)))

raw_index = df %>%
  select(Compound) %>%
  mutate(ID = row_number())

group = getExperimentGroupingDataset()
df_sum = readSummarizedExperiment(df, ecol = 2: 7, fnames = "Compound") #需要
#根据实际数据的样本数量进行更改为 2: x (x 为实际样本数 + 1)
df_sum$group = group$Condition.Name
imputed = PomaImpute(df_sum, ZerosAsNA = T, cutoff = 20, method = "knn", RemoveNA = T)
#缺失值插补方法可以在 method 参数中进行更改。可选择的插补方法有 "none"、"half_min"、
#"median"、"mean"、"min"、"knn" 和 "rf"。如果使用 "none"，则所有的缺失值都会被替
#换为 0
df_imputed = assay(imputed) %>%
  as.data.frame() %>%
  rownames_to_column(var = "Compound") %>%
  mutate(Compound = str_replace(Compound, "(X)(\\d+\\.\\d+)(\\.\\d+\\.\\d+)",
    "\\2@\\13"))

imputed_index = left_join(df_imputed, raw_index, by = "Compound")
addResultColumns(data.frame(imputed_index$ID, df_imputed))
normalized = PomaNorm(imputed, method = "log_pareto")
#数据变换方法可以在 method 参数中进行更改。可选择的变换方法有 "none"、"auto_scaling"、
#"level_scaling"、"log_scaling"、"log_transformation"、"vast_scaling" 以及
#"log_pareto"。
#这些算法的细节以及优缺点都在文献: https://doi.org/10.1186/1471-2164-7-142 中有详细论述
df_nor = assay(normalized) %>%
  as.data.frame() %>%
  rownames_to_column(var = "Compound") %>%
  mutate(Compound = str_replace(Compound, "(X)(\\d+\\.\\d+)(\\.\\d+\\.\\d+)",
    "\\2@\\13"))

nor_index = left_join(df_nor, raw_index, by = "Compound")
addResultColumns(data.frame(nor_index$ID, df_nor))
pre_preprocessed = PomaOutliers(normalized, do = "clean")
#离群点的处理参数 do 有两个值，"clean" 和 "analysis"。输入 "do" 将清除离群点，输入
#"analysis" 只做分析，不进行清除。可选择的离群点判别方法有 "euclidean"、"maximum"、
#"manhattan"、"canberra" 以及 "minkowski"
df_pre_preprocessed = assay(pre_preprocessed) %>%
  as.data.frame() %>%
  rownames_to_column(var = "Compound") %>%
  mutate(Compound = str_replace(Compound, "(X)(\\d+\\.\\d+)(\\.\\d+\\.\\d+)",
    "\\2@\\13"))

pre_preprocessed_index = left_join(df_pre_preprocessed, raw_index, by = "Compound")

addResultColumns(data.frame(pre_preprocessed_index$ID, df_pre_preprocessed))
limma_res = PomaLimma(pre_preprocessed, contrast = "A", adjust = "fdr") %>%
  filter(adj.P.Val < 0.05) %>%
  mutate(feature = str_replace(feature, "(X)(\\d+\\.\\d+)(\\.\\d+\\.\\d+)",
    "\\2@\\13"))
#Limma 实现了一种经验贝叶斯方法，该方法借用跨特征的信息来估计标准误差并计算 (所谓的
#调节) t 统计量。当样本数量较少时，这种方法比标准 t 检验更强大
limma_res_index = left_join(limma_res, raw_index, by = c("feature" = "Compound"))
addResultColumns(data.frame(limma_res_index$ID, limma_res))
lasso_res = PomaLasso(pre_preprocessed, alpha = 1, labels = T, ntest = NULL)
#和 OPLS 不同，LASSO 通过正则化 (引入惩罚项) 消除弱相关变量，可以解决维数灾难、多重共
#线性和过拟合等问题
lasso_coef = lasso_res$coefficients %>%
  filter(str_detect(feature, "X")) %>%
  mutate(feature = str_replace(feature, "(X)(\\d+\\.\\d+)(\\.\\d+\\.\\d+)",
    "\\2@\\13")) #删除了截距信息
lasso_coef_index = left_join(lasso_coef, raw_index, by = c("feature" = "Compound"))
addResultColumns(data.frame(lasso_coef_index$ID, lasso_coef))
```

8. 如图 10 所示，选中要进行分析的实体列表，并同时选择用于分组的解释信息（即告知 `getDataset()` 和 `getExperimentGroupingDataset()` 函数从哪里读取信息）。

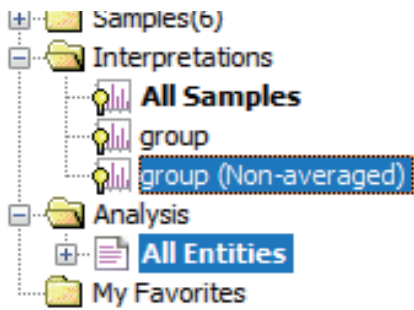


图 10. 选择用于分析的数据和解释

9. 单击 R 脚本编辑器中的“运行”键，运行脚本，将在 Analysis > All Entities 下生成子实体列表（如图 11 所示）。双击该实体列表，即可对其进行重命名。新生成的实体列表可以用于后续分析。

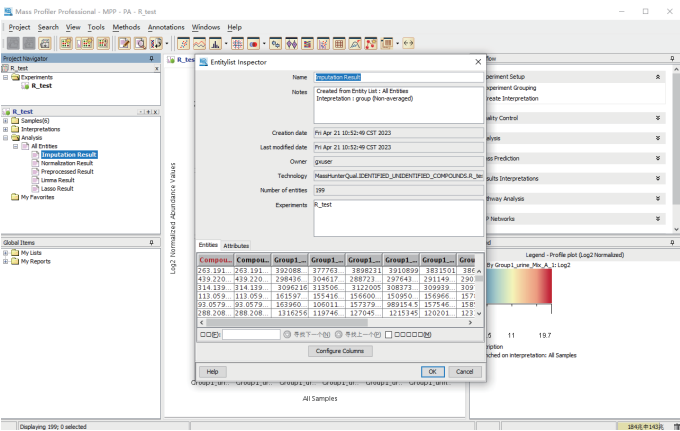


图 11. 新实体列表

结论

借助 MPP 的脚本编辑器，实现了在 MPP 中运行 POMA 的功能。一方面，该方法使研究人员能够应用 MPP 目前尚不支持的数据处理和分析方法分析现有数据；另一方面，处理后的数据依然可以通过 ID Browser，用安捷伦全面的 PCDL 数据库/谱库实现化合物定性分析。该方法使代谢组学和其他相关应用的数据处理和分析流程更为流畅和高效。

```
sessionInfo()

## R version 4.2.3 (2023-03-15 ucrt)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 19045)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=Chinese (Simplified)_China.utf8
## [2] LC_CTYPE=Chinese (Simplified)_China.utf8
## [3] LC_MONETARY=Chinese (Simplified)_China.utf8
## [4] LC_NUMERIC=C
## [5] LC_TIME=Chinese (Simplified)_China.utf8
##
## attached base packages:
## [1] stats4      stats      graphics  grDevices  utils      datasets  methods
## [8] base
##
## other attached packages:
## [1] POMA_1.6.0               QFeatures_1.7.1
## [3] MultiAssayExperiment_1.22.0 SummarizedExperiment_1.26.1
## [5] Biobase_2.56.0           GenomicRanges_1.48.0
## [7] GenomeInfoDb_1.32.4     IRanges_2.30.1
## [9] S4Vectors_0.34.0        BiocGenerics_0.42.0
## [11] MatrixGenerics_1.8.1    matrixStats_1.0.0
## [13] lubridate_1.9.2         forcats_1.0.0
## [15] stringr_1.5.0           dplyr_1.1.2
## [17] purrr_1.0.1             readr_2.1.4
## [19] tidyr_1.3.0            tibble_3.2.1
## [21] ggplot2_3.4.2          tidyverse_2.0.0
##
## loaded via a namespace (and not attached):
## [1] backports_1.4.1         circlize_0.4.15         plyr_1.8.8
##
## [4] igraph_1.4.3            lazyeval_0.2.2          splines_4.2.3
##
## [7] gmp_0.7-1              BiocParallel_1.30.4     listenv_0.9.0
##
## [10] digest_0.6.31          foreach_1.5.2          htmltools_0.5.5
##
## [13] fansi_1.0.4            magrittr_2.0.3         cluster_2.1.4
##
## [16] doParallel_1.0.17      tzdb_0.4.0             limma_3.52.4
##
## [19] recipes_1.0.6          ComplexHeatmap_2.12.1  globals_0.16.2
##
## [22] gower_1.0.1            rARPACK_0.11-0         vroom_1.6.3
##
## [25] hardhat_1.3.0          timechange_0.2.0       colorspace_2.1-0
```


## [28] ggrepel_0.9.3	xfun_0.39	crayon_1.5.2
## [31] RCurl_1.98-1.12	impute_1.70.0	survival_3.5-5
## [34] iterators_1.0.14	glue_1.6.2	gtable_0.3.3
## [37] ipred_0.9-14	zlibbioc_1.42.0	XVector_0.36.0
## [40] GetoptLong_1.0.5	DelayedArray_0.22.0	RankProd_3.22.0
## [43] future.apply_1.11.0	shape_1.4.6	Rmpfr_0.9-2
## [46] scales_1.2.1	Rcpp_1.0.10	viridisLite_0.4.2
## [49] clue_0.3-64	bit_4.0.5	proxy_0.4-27
## [52] MsCoreUtils_1.8.0	lava_1.7.2.1	prodlim_2023.03.31
## [55] glmnet_4.1-7	RColorBrewer_1.1-3	farver_2.1.1
## [58] pkgconfig_2.0.3	nnnet_7.3-19	utf8_1.2.3
## [61] caret_6.0-94	labeling_0.4.2	tidyselect_1.2.0
## [64] rlang_1.1.1	reshape2_1.4.4	munsell_0.5.0
## [67] tools_4.2.3	cli_3.6.1	generics_0.1.3
## [70] broom_1.0.4	evaluate_0.21	fastmap_1.1.1
## [73] yaml_2.3.7	ModelMetrics_1.2.2.2	knitr_1.43
## [76] bit64_4.0.5	randomForest_4.7-1.1	AnnotationFilter_1.2.0
## [79] glasso_1.11	future_1.32.0	nlme_3.1-162
## [82] compiler_4.2.3	rstudioapi_0.14	png_0.1-8
## [85] e1071_1.7-13	stringi_1.7.12	highr_0.10
## [88] RSpectra_0.16-1	lattice_0.21-8	ProtGenerics_1.28.0
## [91] Matrix_1.5-4.1	permute_0.9-7	vegan_2.6-4
## [94] vctrs_0.6.2	pillar_1.9.0	lifecycle_1.0.3
## [97] GlobalOptions_0.1.2	data.table_1.14.8	bitops_1.0-7
## [100] corpcor_1.6.10	R6_2.5.1	gridExtra_2.3
## [103] parallelly_1.36.0	writexl_1.4.2	codetools_0.2-19
## [106] MASS_7.3-60	rjson_0.2.21	withr_2.5.0
## [109] GenomeInfoDbData_1.2.8	mgcv_1.8-42	parallel_4.2.3
## [112] hms_1.1.3	mixOmics_6.20.0	grid_4.2.3
## [115] rpart_4.1.19	timeDate_4022.108	class_7.3-22
## [118] rmarkdown_2.22	pROC_1.18.2	ellipse_0.4.5

参考文献

1. R.A.C.-P. Castellano-Escuder Pol AND González-Domínguez, POMAShiny: A user-friendly web-based workflow for metabolomics and proteomics data analysis, PLOS Computational Biology. 17 (2021) 1–15. <https://doi.org/10.1371/journal.pcbi.1009148>
2. H. Wickham, M. Averick, J. Bryan, W. Chang, L.D. McGowan, R. François, G. Grolemond, A. Hayes, L. Henry, J. Hester, M. Kuhn, T.L. Pedersen, E. Miller, S.M. Bache, K. Müller, J. Ooms, D. Robinson, D.P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, H. Yutani, Welcome to the tidyverse, Journal of Open Source Software. 4 (2019) 1686. <https://doi.org/10.21105/joss.01686>
3. M. Morgan, V. Obenchain, J. Hester, H. Pages, SummarizedExperiment: SummarizedExperiment container, 2022. <https://bioconductor.org/packages/SummarizedExperiment>
4. L. Gatto, C. Vanderaa, QFeatures: Quantitative features for mass spectrometry data, 2022. <https://github.com/RforMassSpectrometry/QFeatures>
5. R. Berg, H.C. Hoefsloot, J.A. Westerhuis, A.K. Smilde, M. Werf, Centering, scaling, and transformations: Improving the biological information content of metabolomics data, BMC Genomics. 7 (2006) 142–142
6. M.E. Ritchie, P. Belinda, D. Wu, Y. Hu, C.W. Law, S. Wei, G.K. Smyth, Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Research. 43 (2015) e47
7. J.H. Friedman, T. Hastie, R. Tibshirani, Regularization paths for generalized linear models via coordinate descent, Journal of Statistical Software. 33 (2010)

附表

表 2. 新生成的组合数据（由于页面限制，已将表格转置以便查看）

化合物	263.191@6.162	439.2208@5.7410007	314.1397@5.781	113.0592@0.664	93.0579@0.82399994	288.2089@8.294
Group1_urine_Mix_A_1	21.902748	21.508993	21.562075	20.623976	20.644913	0
Group1_urine_Mix_A_2	21.849052	21.538568	21.580065	20.567707	20.015793	0
Group1_urine_Mix_A_3	21.894388	21.461256	21.574041	20.578653	20.585814	0
Group1_urine_Mix_A_4	21.8463727725498	21.4959668714137	21.5756270730498	20.5473819940025	20.1156702253799	0
Group1_urine_Mix_A_5	21.9014767232407	21.5982259623633	21.5778641934518	20.6011270318281	20.2863091480832	0
Group1_urine_Mix_A_6	21.8554404541851	21.502286795733	21.562449188124	20.5716890356257	20.3926333261921	0
Group1_urine_Mix_A_7	21.8553746950681	21.4662095913321	21.5720405282496	20.5949894840442	20.364492881382	0
Group1_urine_Mix_A_8	21.8812926748139	21.5141446486569	21.5802401771816	20.5640089381268	20.4083358581929	0
Group1_urine_Mix_A_9	21.8707518732287	21.4520627011545	21.5907347706265	20.6154200982367	19.9633612563534	0
Group1_urine_Mix_A_10	21.8396214128721	21.4745910580447	21.5678817617193	20.5751104427717	20.2310407044591	0
Group1_urine_Mix_A_11	21.9125600150167	21.5115412294641	21.5722656920994	20.5651230440106	20.3007989330638	0
Group1_urine_Mix_A_12	21.8488302140589	21.5187043853875	21.5800966340744	20.5450438916216	20.6006164131535	0
Group1_urine_Mix_A_13	21.8377473038845	21.4780507567406	21.5738072931151	20.5438403584327	20.2095394783369	0
Group1_urine_Mix_A_14	21.9380915456658	21.4984463954156	21.5603929717054	20.6104875835642	20.0903644990488	0
Group1_urine_Mix_A_15	21.8838139766387	21.4934787840867	21.5746719273725	20.6354035832134	20.2686480781405	0
Group1_urine_Mix_A_16	21.8791987737283	21.4816915700968	21.5842532631684	20.6141582864574	20.509914430027	0
Group1_urine_Mix_A_17	21.8878035249878	21.5329487358409	21.5463330158764	20.6263346104182	20.3476572748052	0
Group1_urine_Mix_A_18	21.9160958388278	21.4402056991927	21.5605946295174	20.6006197901612	20.1524639378738	0
Group1_urine_Mix_A_19	21.912496382637	21.4488305818965	21.575372490735	20.589614178908	20.5807228425876	0
Group1_urine_Mix_A_20	21.8834683065681	21.5411151972566	21.5658295926967	20.5351951106281	20.4825641218232	0
Group1_urine_Mix_B_1	21.899069	21.505156	21.556248	20.525642	19.915836	20.212934
Group1_urine_Mix_B_2	21.869478	21.473330	21.563557	20.582024	20.587343	20.197025
Group1_urine_Mix_B_3	21.883112	21.469435	21.562828	20.590416	20.596172	20.234135
Group1_urine_Mix_B_4	21.8899626032159	21.5024268233923	21.556527113117	20.5843913715121	21.1314332205299	20.2150972321196
Group1_urine_Mix_B_5	21.8790594986673	21.4583450382968	21.5652641222624	20.5666665751694	20.3967989184984	20.2551723097883
Group1_urine_Mix_B_6	21.8932757716064	21.4589483379203	21.5646589987082	20.6279358557639	21.0042523963234	20.2224822808584
Group1_urine_Mix_B_7	21.8954860372141	21.4560479937505	21.5598454934942	20.538041537517	19.2093757031625	20.240675815158
Group1_urine_Mix_B_8	21.8957373399131	21.4851998365106	21.5694949577068	20.541132100559	20.3210355991294	20.2211733721295
Group1_urine_Mix_B_9	21.8839017383608	21.5080523877315	21.5595137513033	20.6367992840178	20.0978542821457	20.2344532342627
Group1_urine_Mix_B_10	21.8962407654131	21.4756264556101	21.5648391005323	20.5920299628348	20.3374308306084	20.2051676355341
Group1_urine_Mix_B_11	21.8718827528707	21.4721894548459	21.5565800566223	20.5498359017668	20.7306315866167	20.2024748720726
Group1_urine_Mix_B_12	21.9001680548719	21.490347691638	21.5627114971664	20.5177429308716	20.0554680604006	20.2067326988035
Group1_urine_Mix_B_13	21.9032654060477	21.4755683464908	21.5639568688668	20.61004064364	19.5266318622143	20.222941182876
Group1_urine_Mix_B_14	21.8656595146731	21.4844704299372	21.5591879644127	20.6027280182418	20.4042019580737	20.2072028928793
Group1_urine_Mix_B_15	21.8806210244391	21.5215968232916	21.5600871889051	20.4878994432544	20.4487110138921	20.2439965690126
Group1_urine_Mix_B_16	21.9040323664132	21.4641013019919	21.5610916682548	20.521277420187	20.1314466258918	20.2254180754538
Group1_urine_Mix_B_17	21.901300476708	21.4747238835639	21.558240748729	20.5744451485395	20.0819378785682	20.2393912267688
Group1_urine_Mix_B_18	21.8685189947573	21.4812868602558	21.5557873404642	20.5833087327098	19.7387788053621	20.2213350541604
Group1_urine_Mix_B_19	21.8945531631717	21.4633242048608	21.5677310923358	20.5761537618766	20.7900582394032	20.2274622950154
Group1_urine_Mix_B_20	21.8898192593366	21.4910404474952	21.5641408412411	20.5359102158603	19.8666582802313	20.235816206677

表 3. 还原后的数据（由于页面限制，已将表格转置以便查看）

化合物	263.191@6.162	439.2208@5.7410007	314.1397@5.781	113.0592@0.664	93.0579@0.82399994	288.2089@8.294
Group1_urine_Mix_A_1	3920885.27886654	2984365.90358624	3096216.66436165	1615977.85816054	1639600.60869021	0
Group1_urine_Mix_A_2	3777635.30710391	3046176.27730762	3135067.33739193	1554163.65681905	1060117.68617779	0
Group1_urine_Mix_A_3	3898230.58613326	2887232.78539451	3122004.09724415	1566000.23668303	1573792.60087291	0
Group1_urine_Mix_A_4	3770626.37468451	2957541.27366663	3125438.25993072	1532421.7682374	1136109.31519219	0
Group1_urine_Mix_A_5	3917431.78752179	3174781.69640972	3130288.49187334	1590586.09338738	1278755.61847326	0
Group1_urine_Mix_A_6	3794400.29280873	2970525.60961	3097019.82631388	1558459.28660563	1376557.38841459	0
Group1_urine_Mix_A_7	3794227.3451522	2897163.33395661	3117678.05967817	1583833.75710634	1349967.16009482	0
Group1_urine_Mix_A_8	3863006.5010708	2995041.67824176	3135448.03157092	1550184.96877952	1391621.90084882	0
Group1_urine_Mix_A_9	3834884.97040602	2868892.94106499	3158339.37274413	1606422.66433063	1022281.61982045	0
Group1_urine_Mix_A_10	3753022.25052122	2914043.65153478	3108703.86551691	1562159.61914907	1230694.03472406	0
Group1_urine_Mix_A_11	3947642.77634428	2989641.84138512	3118164.67892725	1551382.54496387	1291663.5814807	0
Group1_urine_Mix_A_12	3777054.61473189	3004522.68790097	3135136.08098501	1529940.26168782	1590023.2306705	0
Group1_urine_Mix_A_13	3748150.11463247	2921040.14821815	3121498.39455792	1528664.47851852	1212488.36979262	0
Group1_urine_Mix_A_14	4018126.47813231	2962628.69668644	3092608.91003647	1600939.74995529	1116355.04157159	0
Group1_urine_Mix_A_15	3869763.5223593	2952445.06204768	3123369.72792947	1628828.84318032	1263196.87636448	0
Group1_urine_Mix_A_16	3857403.87114424	2928421.05151257	3144181.92127116	1605018.26713559	1493136.29349787	0
Group1_urine_Mix_A_17	3880479.56037002	3034334.56700448	3062615.91023964	1618621.92320119	1334305.20818434	0
Group1_urine_Mix_A_18	3957329.70779975	2845411.14779118	3093041.22061965	1590026.95254295	1165456.67033045	0
Group1_urine_Mix_A_19	3947468.66306503	2862472.85591384	3124886.78433514	1577943.56626797	1568248.5984027	0
Group1_urine_Mix_A_20	3868836.4371799	3051559.30381997	3104285.01717704	1519531.45677342	1465096.3580452	0
Group1_urine_Mix_B_1	3910899.41232924	2976439.21253219	3083736.36314879	1509502.7976401	989154.257744031	1215344.59635547
Group1_urine_Mix_B_2	3831500.3319043	2911497.60230498	3099398.8688999	1569663.62977201	1575461.42512841	1202016.27563666
Group1_urine_Mix_B_3	3867881.05743341	2903647.71836411	3097833.12496257	1578820.79954117	1585132.49066932	1233336.46276396
Group1_urine_Mix_B_4	3886291.27462947	2970813.94215655	3084333.02045238	1572241.46317782	2297181.18607522	1217168.29744547
Group1_urine_Mix_B_5	3857031.5026164	2881413.01723914	3103068.51793264	1553043.2484908	1380537.76057775	1251452.67971685
Group1_urine_Mix_B_6	3895226.4502885	2882618.20534789	3101767.24077205	1620419.4269032	2103342.55106836	1223414.8657581
Group1_urine_Mix_B_7	3901198.66424724	2876828.91074136	3091435.54024697	1522532.44084193	606175.36778045	1238940.79368509
Group1_urine_Mix_B_8	3901878.2722843	2935550.89014566	3112181.90718094	1525797.52877604	1309909.39884011	1222305.40592714
Group1_urine_Mix_B_9	3869998.93416231	2982420.78268314	3090724.75822635	1630405.37714611	1122165.69338342	1233608.54605274
Group1_urine_Mix_B_10	3903240.06235405	2916135.76148855	3102154.4804815	1580588.03647095	1324880.52042056	1208819.6907519
Group1_urine_Mix_B_11	3837892.18495646	2909196.78188235	3084446.21027657	1535030.51192158	1739970.70612792	1206565.55558666
Group1_urine_Mix_B_12	3913879.89717241	2946044.29552603	3097582.97385305	1501260.52621801	1089676.19457765	1210131.75309931
Group1_urine_Mix_B_13	3922291.7116364	2916018.30723571	3100258.04207842	1600443.86340357	755269.42669738	1223804.07958021
Group1_urine_Mix_B_14	3821372.63158225	2934067.09151645	3090026.89492287	1592352.17709152	1387640.05328676	1210526.21589085
Group1_urine_Mix_B_15	3861208.48705831	3010552.45337931	3091953.49335117	1470524.55418939	1431117.81453732	1241795.83645446
Group1_urine_Mix_B_16	3924377.42064389	2892932.64240106	3094107.0218574	1504943.00541382	1148601.26566323	1225906.97363029
Group1_urine_Mix_B_17	3916953.2445321	2914311.95299017	3087998.77309849	1561439.40028242	1109853.54230027	1237838.12086245
Group1_urine_Mix_B_18	3828954.25818992	2927599.67584597	3082751.86833127	1571062.05169568	874912.251781861	1222442.39669033
Group1_urine_Mix_B_19	3898676.8904215	2891374.80499032	3108379.22169124	1563289.74048229	1813139.17590355	1227645.24764694
Group1_urine_Mix_B_20	3885905.15810625	2947459.27198392	3100653.41198294	1520284.834455	956004.762858023	1234774.53668281

表 6. 经过插补后的数据（由于页面限制，已将表格转置以便查看）

化合物	263.191@6.162	439.2208@5.7410007	314.1397@5.781	113.0592@0.664	93.0579@0.82399994	288.2089@8.294
Group1_urine_Mix_A_1	3920885.27886654	2984365.90358624	3096216.66436165	1615977.85816054	1639600.6086902	1316255.92244626
Group1_urine_Mix_A_2	3777635.30710391	3046176.27730762	3135067.33739193	1554163.65681905	1060117.68617779	1197460.97409059
Group1_urine_Mix_A_3	3898230.58613326	2887232.78539451	3122004.09724415	1566000.23668303	1573792.60087291	1270453.58839057
Group1_urine_Mix_A_4	3770626.3746845	2957541.27366663	3125438.25993072	1532421.7682374	1136109.31519219	47232447.4714519
Group1_urine_Mix_A_5	3917431.7875218	3174781.69640972	3130288.49187333	1590586.09338738	1278755.61847326	87100381935920.8
Group1_urine_Mix_A_6	3794400.29280873	2970525.60961	3097019.82631388	1558459.28660563	1376557.38841458	1094735.87647697
Group1_urine_Mix_A_7	3794227.3451522	2897163.33395661	3117678.05967817	1583833.75710634	1349967.16009482	3070479.73224992
Group1_urine_Mix_A_8	3863006.5010708	2995041.67824176	3135448.03157092	1550184.96877952	1391621.90084882	1133422.24080514
Group1_urine_Mix_A_9	3834884.97040602	2868892.94106499	3158339.37274413	1606422.66433063	1022281.61982045	5405716.9446787
Group1_urine_Mix_A_10	3753022.25052122	2914043.65153478	3108703.86551691	1562159.61914907	1230694.03472406	1091203.52619051
Group1_urine_Mix_A_11	3947642.77634428	2989641.84138512	3118164.67892725	1551382.54496387	1291663.5814807	1231065.63496078
Group1_urine_Mix_A_12	3777054.61473189	3004522.68790097	3135136.08098501	1529940.26168782	1590023.2306705	1144204.10508998
Group1_urine_Mix_A_13	3748150.11463247	2921040.14821815	3121498.39455792	1528664.47851852	1212488.36979262	1105194.44021025
Group1_urine_Mix_A_14	4018126.47813231	2962628.69668644	3092608.91003647	1600939.74995529	1116355.04157159	1079836.74289692
Group1_urine_Mix_A_15	3869763.5223593	2952445.06204768	3123369.72792947	1628828.84318032	1263196.87636448	1120523.0998967
Group1_urine_Mix_A_16	3857403.87114424	2928421.05151257	3144181.92127116	1605018.26713559	1493136.29349787	1145553.69836958
Group1_urine_Mix_A_17	3880479.56037002	3034334.56700448	3062615.91023964	1618621.92320119	1334305.20818434	7989419.73543269
Group1_urine_Mix_A_18	3957329.70779975	2845411.14779118	3093041.22061965	1590026.95254294	1165456.67033045	1139164.908976
Group1_urine_Mix_A_19	3947468.66306503	2862472.85591384	3124886.78433514	1577943.56626797	1568248.5984027	1128569.53584566
Group1_urine_Mix_A_20	3868836.43717989	3051559.30381997	3104285.01717704	1519531.45677342	1465096.3580452	1147431.2356761
Group1_urine_Mix_B_1	3910899.41232923	2976439.21253218	3083736.36314879	1509502.7976401	989154.257744031	1215344.59635547
Group1_urine_Mix_B_2	3831500.3319043	2911497.60230498	3099398.8688999	1569663.62977201	1575461.42512841	1202016.27563666
Group1_urine_Mix_B_3	3867881.05743341	2903647.71836411	3097833.12496257	1578820.79954117	1585132.49066932	1233336.46276396
Group1_urine_Mix_B_4	3886291.27462947	2970813.94215655	3084333.02045238	1572241.46317782	2297181.18607522	1217168.29744546
Group1_urine_Mix_B_5	3857031.5026164	2881413.01723914	3103068.51793264	1553043.2484908	1380537.76057775	1251452.67971685
Group1_urine_Mix_B_6	3895226.4502885	2882618.20534789	3101767.24077205	1620419.4269032	2103342.55106836	1223414.8657581
Group1_urine_Mix_B_7	3901198.66424724	2876828.91074136	3091435.54024697	1522532.44084193	606175.367780453	1238940.79368509
Group1_urine_Mix_B_8	3901878.2722843	2935550.89014566	3112181.90718094	1525797.52877604	1309909.39884011	1222305.40592714
Group1_urine_Mix_B_9	3869998.93416231	2982420.78268314	3090724.75822635	1630405.37714611	1122165.69338342	1233608.54605274
Group1_urine_Mix_B_10	3903240.06235405	2916135.76148856	3102154.4804815	1580588.03647095	1324880.52042056	1208819.6907519
Group1_urine_Mix_B_11	3837892.18495646	2909196.78188235	3084446.21027657	1535030.51192158	1739970.70612792	1206565.55558666
Group1_urine_Mix_B_12	3913879.89717241	2946044.29552603	3097582.97385305	1501260.52621801	1089676.19457765	1210131.75309931
Group1_urine_Mix_B_13	3922291.7116364	2916018.30723572	3100258.04207842	1600443.86340357	755269.426697382	1223804.07958021
Group1_urine_Mix_B_14	3821372.63158225	2934067.09151645	3090026.89492287	1592352.17709152	1387640.05328676	1210526.21589085
Group1_urine_Mix_B_15	3861208.48705831	3010552.45337931	3091953.49335117	1470524.55418939	1431117.81453732	1241795.83645445
Group1_urine_Mix_B_16	3924377.42064389	2892932.64240106	3094107.0218574	1504943.00541382	1148601.26566323	1225906.97363029
Group1_urine_Mix_B_17	3916953.2445321	2914311.95299017	3087998.77309849	1561439.40028242	1109853.54230027	1237838.12086245
Group1_urine_Mix_B_18	3828954.25818992	2927599.67584597	3082751.86833127	1571062.05169569	874912.251781861	1222442.39669033
Group1_urine_Mix_B_19	3898676.8904215	2891374.80499032	3108379.22169124	1563289.74048229	1813139.17590355	1227645.24764693
Group1_urine_Mix_B_20	3885905.15810625	2947459.27198392	3100653.41198294	1520284.834455	956004.762858023	1234774.53668281

表 7. 经过变换后的数据（由于页面限制，已将表格转置以便查看）

化合物	263.191@6.162	439.2208@5.7410007	314.1397@5.781	113.0592@0.664	93.0579@0.82399994	288.2089@8.294
Group1_urine_Mix_A_1	6.593	6.475	6.491	6.208	6.215	0
Group1_urine_Mix_A_2	6.577	6.484	6.496	6.191	6.025	0
Group1_urine_Mix_A_3	6.591	6.46	6.494	6.195	6.197	0
Group1_urine_Mix_A_4	6.576	6.471	6.495	6.185	6.055	0
Group1_urine_Mix_A_5	6.593	6.502	6.496	6.202	6.107	0
Group1_urine_Mix_A_6	6.579	6.473	6.491	6.193	6.139	0
Group1_urine_Mix_A_7	6.579	6.462	6.494	6.2	6.13	0
Group1_urine_Mix_A_8	6.587	6.476	6.496	6.19	6.144	0
Group1_urine_Mix_A_9	6.584	6.458	6.499	6.206	6.01	0
Group1_urine_Mix_A_10	6.574	6.464	6.493	6.194	6.09	0
Group1_urine_Mix_A_11	6.596	6.476	6.494	6.191	6.111	0
Group1_urine_Mix_A_12	6.577	6.478	6.496	6.185	6.201	0
Group1_urine_Mix_A_13	6.574	6.466	6.494	6.184	6.084	0
Group1_urine_Mix_A_14	6.604	6.472	6.49	6.204	6.048	0
Group1_urine_Mix_A_15	6.588	6.47	6.495	6.212	6.101	0
Group1_urine_Mix_A_16	6.586	6.467	6.498	6.205	6.174	0
Group1_urine_Mix_A_17	6.589	6.482	6.486	6.209	6.125	0
Group1_urine_Mix_A_18	6.597	6.454	6.49	6.201	6.066	0
Group1_urine_Mix_A_19	6.596	6.457	6.495	6.198	6.195	0
Group1_urine_Mix_A_20	6.588	6.485	6.492	6.182	6.166	0
Group1_urine_Mix_B_1	6.592	6.474	6.489	6.179	5.995	6.085
Group1_urine_Mix_B_2	6.583	6.464	6.491	6.196	6.197	6.08
Group1_urine_Mix_B_3	6.587	6.463	6.491	6.198	6.2	6.091
Group1_urine_Mix_B_4	6.59	6.473	6.489	6.197	6.361	6.085
Group1_urine_Mix_B_5	6.586	6.46	6.492	6.191	6.14	6.097
Group1_urine_Mix_B_6	6.591	6.46	6.492	6.21	6.323	6.088
Group1_urine_Mix_B_7	6.591	6.459	6.49	6.183	5.783	6.093
Group1_urine_Mix_B_8	6.591	6.468	6.493	6.183	6.117	6.087
Group1_urine_Mix_B_9	6.588	6.475	6.49	6.212	6.05	6.091
Group1_urine_Mix_B_10	6.591	6.465	6.492	6.199	6.122	6.082
Group1_urine_Mix_B_11	6.584	6.464	6.489	6.186	6.241	6.082
Group1_urine_Mix_B_12	6.593	6.469	6.491	6.176	6.037	6.083
Group1_urine_Mix_B_13	6.594	6.465	6.491	6.204	5.878	6.088
Group1_urine_Mix_B_14	6.582	6.467	6.49	6.202	6.142	6.083
Group1_urine_Mix_B_15	6.587	6.479	6.49	6.167	6.156	6.094
Group1_urine_Mix_B_16	6.594	6.461	6.491	6.178	6.06	6.088
Group1_urine_Mix_B_17	6.593	6.465	6.49	6.194	6.045	6.093
Group1_urine_Mix_B_18	6.583	6.467	6.489	6.196	5.942	6.087
Group1_urine_Mix_B_19	6.591	6.461	6.493	6.194	6.258	6.089
Group1_urine_Mix_B_20	6.589	6.469	6.491	6.182	5.98	6.092

表 8. 经过完整预处理后的数据（由于页面限制，已将表格转置以便查看）

化合物	263.191@6.162	439.2208@5.7410007	314.1397@5.781	113.0592@0.664	93.0579@0.82399994	288.2089@8.294
Group1_urine_Mix_A_1	6.593	6.475	6.491	6.208	6.215	0
Group1_urine_Mix_A_2	6.577	6.484	6.496	6.191	6.025	0
Group1_urine_Mix_A_3	6.591	6.46	6.494	6.195	6.197	0
Group1_urine_Mix_A_4	6.576	6.471	6.495	6.185	6.055	0
Group1_urine_Mix_A_6	6.579	6.473	6.491	6.193	6.139	0
Group1_urine_Mix_A_7	6.579	6.462	6.494	6.2	6.13	0
Group1_urine_Mix_A_8	6.587	6.476	6.496	6.19	6.144	0
Group1_urine_Mix_A_9	6.584	6.458	6.499	6.206	6.01	0
Group1_urine_Mix_A_10	6.574	6.464	6.493	6.194	6.09	0
Group1_urine_Mix_A_11	6.596	6.476	6.494	6.191	6.111	0
Group1_urine_Mix_A_12	6.577	6.478	6.496	6.185	6.201	0
Group1_urine_Mix_A_13	6.574	6.466	6.494	6.184	6.084	0
Group1_urine_Mix_A_14	6.604	6.472	6.49	6.204	6.048	0
Group1_urine_Mix_A_15	6.588	6.47	6.495	6.212	6.101	0
Group1_urine_Mix_A_16	6.586	6.467	6.498	6.205	6.174	0
Group1_urine_Mix_A_17	6.589	6.482	6.486	6.209	6.125	0
Group1_urine_Mix_A_18	6.597	6.454	6.49	6.201	6.066	0
Group1_urine_Mix_A_19	6.596	6.457	6.495	6.198	6.195	0
Group1_urine_Mix_A_20	6.588	6.485	6.492	6.182	6.166	0
Group1_urine_Mix_B_1	6.592	6.474	6.489	6.179	5.995	6.085
Group1_urine_Mix_B_2	6.583	6.464	6.491	6.196	6.197	6.08
Group1_urine_Mix_B_3	6.587	6.463	6.491	6.198	6.2	6.091
Group1_urine_Mix_B_4	6.59	6.473	6.489	6.197	6.361	6.085
Group1_urine_Mix_B_5	6.586	6.46	6.492	6.191	6.14	6.097
Group1_urine_Mix_B_6	6.591	6.46	6.492	6.21	6.323	6.088
Group1_urine_Mix_B_7	6.591	6.459	6.49	6.183	5.783	6.093
Group1_urine_Mix_B_8	6.591	6.468	6.493	6.183	6.117	6.087
Group1_urine_Mix_B_9	6.588	6.475	6.49	6.212	6.05	6.091
Group1_urine_Mix_B_10	6.591	6.465	6.492	6.199	6.122	6.082
Group1_urine_Mix_B_11	6.584	6.464	6.489	6.186	6.241	6.082
Group1_urine_Mix_B_12	6.593	6.469	6.491	6.176	6.037	6.083
Group1_urine_Mix_B_13	6.594	6.465	6.491	6.204	5.878	6.088
Group1_urine_Mix_B_14	6.582	6.467	6.49	6.202	6.142	6.083
Group1_urine_Mix_B_15	6.587	6.479	6.49	6.167	6.156	6.094
Group1_urine_Mix_B_16	6.594	6.461	6.491	6.178	6.06	6.088
Group1_urine_Mix_B_17	6.593	6.465	6.49	6.194	6.045	6.093
Group1_urine_Mix_B_18	6.583	6.467	6.489	6.196	5.942	6.087
Group1_urine_Mix_B_19	6.591	6.461	6.493	6.194	6.258	6.089
Group1_urine_Mix_B_20	6.589	6.469	6.491	6.182	5.98	6.092

查找当地的安捷伦客户中心：

www.agilent.com/chem/contactus-cn

免费专线：

800-820-3278, 400-820-3278 (手机用户)

联系我们：

LSCA-China_800@agilent.com

在线询价：

www.agilent.com/chem/erfq-cn



微信搜一搜

Q 安捷伦视界

www.agilent.com

DE42748598

安捷伦对本资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、说明和技术指标如有变更，恕不另行通知。

© 安捷伦科技（中国）有限公司，2023

2023 年 7 月 10 日，中国出版

5994-6513ZHCN



Agilent

Trusted Answers