

PEC1: Metadatos del dataset de metabolómica 2023-CIMCBTutorial

Sara Suárez Hernández

2025-04-02

Contents

1. Objecto summarizedExperiment	1
2. Metadatos	1
2.1 Metadatos muestras	1
2.2 Metadatos metabolitos	2

1. Objecto summarizedExperiment

```
## class: SummarizedExperiment
## dim: 149 140
## metadata(0):
## assays(1): metabolomics
## rownames(149): M1 M2 ... M148 M149
## rowData names(3): nombre_metabolito Perc_missing QC_RSD
## colnames(140): sample_1 sample_2 ... sample_139 sample_140
## colData names(2): SampleType Class
```

2. Metadatos

2.1 Metadatos muestras

Los metadatos disponibles para las muestras son escasos, limitándose a:

- **SampleType**: tipo de muestra.
 - QC: control.
 - Sample: muestra experimental.
- **Class**: clase.
 - HE: participante sano.
 - GC: participante con tumores malignos (de estómago).
 - BN: tumor benigno.

```
## DataFrame with 140 rows and 2 columns
##           SampleType      Class
##           <character> <character>
## sample_1           QC         QC
## sample_2           Sample      GC
## sample_3           Sample      BN
```

```
## sample_4      Sample      HE
## sample_5      Sample      GC
## ...           ...         ...
## sample_136    QC          QC
## sample_137    Sample      GC
## sample_138    Sample      BN
## sample_139    Sample      HE
## sample_140    QC          QC
```

Esta tabla muestra el número de tipo de muestras por clase:

SampleType	Class	n
QC	QC	17
Sample	BN	40
Sample	GC	43
Sample	HE	40

2.2 Metadatos metabolitos

También hay metadatos para los metabolitos, que incluyen:

- **nombre_metabolito**: nombre de la molécula.
- **Perc_missing**: porcentaje de muestras en la que se detectó el metabolito.
- **QC_RSD**: medida de la variación de la medida de cada uno de los metabolitos.

Estas dos últimas columnas determinan el criterio de inclusión de los metabolitos en el análisis de los datos (`perc_missing<10%` and `QC_RSD<20`).

```
## DataFrame with 149 rows and 3 columns
##      nombre_metabolito Perc_missing  QC_RSD
##      <character>      <numeric> <numeric>
## M1      1_3-Dimethylurate    11.428571  32.20800
## M2      1_6-Anhydro- -D-gluc..  0.714286  31.17803
## M3      1_7-Dimethylxanthine   5.000000  34.99060
## M4      1-Methylnicotinamide   8.571429  12.80420
## M5      2-Aminoadipate        1.428571   9.37266
## ...           ...         ...         ...
## M145      uarm1             23.57143   41.4070
## M146      uarm2              4.28571   34.4582
## M147      -Alanine           1.42857   27.6235
## M148      -Methylhistidine    1.42857   16.5619
## M149      -Methylhistidine    0.00000    8.3518
```

La siguiente tabla muestra los estadísticos de resumen de la distribución de *Perc_missing* y *QC_RSD*:

Perc_missing	QC_RSD
Min. : 0.0000	Min. : 2.432
1st Qu.: 0.7143	1st Qu.: 13.280
Median : 2.1429	Median : 26.605
Mean : 5.1246	Mean : 29.569
3rd Qu.: 6.4286	3rd Qu.: 40.175
Max. :31.4286	Max. :133.786