

# Package ‘HMP2Data’

September 11, 2025

**Title** 16s rRNA sequencing data from the Human Microbiome Project 2

**Version** 1.23.0

**Author** John Stansfield <stansfieldjc@vcu.edu>,  
Ekaterina Smirnova <ekaterina.smirnova@vcuhealth.org>,  
Ni Zhao <nzhao10@jhu.edu>,  
Jennifer Fettweis <jennifer.fettweis@vcuhealth.org>,  
Levi Waldron <levi.waldron@sph.cuny.edu>,  
Mikhail Dozmorov <mikhail.dozmorov@vcuhealth.org>

**Maintainer** John Stansfield <stansfieldjc@vcu.edu>, Ekaterina Smirnova <ekaterina.smirnova@vcuhealth.org>

**Description** HMP2Data is a Bioconductor package of the Human Microbiome Project 2 (HMP2) 16S rRNA sequencing data. Processed data is provided as phyloseq, SummarizedExperiment, and MultiAssayExperiment class objects. Individual matrices and data.frames used for building these S4 class objects are also provided in the package.

**Depends** R (>= 3.6.0)

**License** Artistic-2.0

**Encoding** UTF-8

**LazyData** false

**BugReports** <https://github.com/jstansfield0/HMP2Data/issues>

**URL** <https://github.com/jstansfield0/HMP2Data>

**Imports** AnnotationHub, assertthat, dplyr, ExperimentHub, kableExtra, knitr, magrittr, methods, readr, S4Vectors, SummarizedExperiment, phyloseq, MultiAssayExperiment, data.table

**Suggests** Biobase, BiocCheck, BiocManager, BiocStyle, cowplot, devtools, ggplot2, gridExtra, haven, rmarkdown, roxygen2, stats, testthat, tibble, UpSetR

**biocViews** ExperimentData, Homo\_sapiens\_Data, MicrobiomeData, ReproducibleResearch, SequencingData

**RoxygenNote** 7.1.1

**VignetteBuilder** knitr

**git\_url** <https://git.bioconductor.org/packages/HMP2Data>

**git\_branch** devel

**git\_last\_commit** 982e03c  
**git\_last\_commit\_date** 2025-04-15  
**Repository** Bioconductor 3.22  
**Date/Publication** 2025-09-11

**Contents**

|                            |           |
|----------------------------|-----------|
| IBD16S . . . . .           | 2         |
| IBD16S_mtx . . . . .       | 3         |
| IBD16S_samp . . . . .      | 3         |
| IBD16S_tax . . . . .       | 4         |
| momspi16S . . . . .        | 4         |
| momspi16S_mtx . . . . .    | 5         |
| momspi16S_samp . . . . .   | 5         |
| momspi16S_tax . . . . .    | 5         |
| momspiCytokines . . . . .  | 6         |
| momspiCyto_mtx . . . . .   | 7         |
| momspiCyto_samp . . . . .  | 7         |
| momspiMultiAssay . . . . . | 7         |
| patient_table . . . . .    | 8         |
| T2D16S . . . . .           | 9         |
| T2D16S_mtx . . . . .       | 10        |
| T2D16S_samp . . . . .      | 10        |
| T2D16S_tax . . . . .       | 10        |
| table_two . . . . .        | 11        |
| visit_table . . . . .      | 11        |
| <b>Index</b>               | <b>12</b> |

---

|        |                                                                                             |
|--------|---------------------------------------------------------------------------------------------|
| IBD16S | <i>Create HMP2 16S sequencing data for Inflammatory Bowel Disease (IBD) phyloseq object</i> |
|--------|---------------------------------------------------------------------------------------------|

---

**Description**

The HMP2 study gathered data on the microbiome for patients with different disease. This dataset contains 16S rRNA data for 196 patients with IBD over multiple visits.

**Usage**

IBD16S()

**Format**

A phyloseq object with 982 taxa and 178 samples.

**OTU Table:** A counts matrix for the rRNA-seq results.

**Sample Data:**

**Project** Name of the project  
**sample\_id** Sample identifier

**subject\_id** Subject identifier See `colnames(sample_data(IBD16S()))` for remaining values.

**Taxonomy Table:** Taxonomy table for the taxa identified.

### Value

a phyloseq object

### Examples

```
IBD <- IBD16S()
```

---

|            |                   |
|------------|-------------------|
| IBD16S_mtx | <i>IBD16S_mtx</i> |
|------------|-------------------|

---

### Description

Matrix of OTU values for 16S rRNA data for IBD

### Usage

```
IBD16S_mtx
```

### Format

An object of class `matrix` (inherits from `array`) with 982 rows and 178 columns.

---

|             |                    |
|-------------|--------------------|
| IBD16S_samp | <i>IBD16S_samp</i> |
|-------------|--------------------|

---

### Description

Sample data for IBD16S data

### Usage

```
IBD16S_samp
```

### Format

An object of class `data.frame` with 178 rows and 490 columns.

---

IBD16S\_tax

*IBD16S\_tax*


---

### Description

Taxonomy data for IBD 16S data

### Usage

IBD16S\_tax

### Format

An object of class `matrix` (inherits from `array`) with 982 rows and 6 columns.

---

momspi16S

*Construct MOMS-PI 16S rRNA phyloseq object*


---

### Description

The Multi-Omic Microbiome Study-Pregnancy Initiative (MOMS-PI) was funded by the NIH Roadmap HUMAN Microbiome Project to understand the impact of the vaginal microbiome on pregnancy and the fetal microbiome. This longitudinal study contains samples from various body sites of 596 women. This `phyloseq` object contains the 16S rRNA sequencing data that was collected. Also contained are the taxa data and participant metadata.

### Usage

`momspi16S()`

### Format

A `phyloseq` object with 7,665 taxa and 9,107 samples.

**OTU Table:** A counts matrix for the rRNA-seq results.

#### Sample Data:

**file\_id** File identifier

**md5** md5 hash for the file

**size** file size

**urls** URL for the file

**sample\_id** Sample identifier

**file\_name** Filename which the sample was taken from

**subject\_id** Participant identifier

**sample\_body\_site** Body site of the sample

**visit\_number** Visit number

**subject\_gender** Participant gender

**subject\_race** Participant race

**study\_full\_name** Name of the study

**project\_name** Name of project

**Taxonomy Table:** Taxonomy table for the taxa identified.

**Value**

a phyloseq object

**Examples**

```
momspi <- momspi16S()
```

---

|               |                      |
|---------------|----------------------|
| momspi16S_mtx | <i>momspi16S_mtx</i> |
|---------------|----------------------|

---

**Description**

16S rRNA OTU matrix for MOMS-PI data

**Usage**

```
momspi16S_mtx
```

**Format**

An object of class `matrix` (inherits from `array`) with 7665 rows and 9107 columns.

---

|                |                       |
|----------------|-----------------------|
| momspi16S_samp | <i>momspi16S_samp</i> |
|----------------|-----------------------|

---

**Description**

Sample data for MOMS-PI 16S data

**Usage**

```
momspi16S_samp
```

**Format**

An object of class `data.frame` with 9107 rows and 13 columns.

---

|               |                      |
|---------------|----------------------|
| momspi16S_tax | <i>momspi16S_tax</i> |
|---------------|----------------------|

---

**Description**

Taxonomy data for MOMS-PI 16S data

**Usage**

```
momspi16S_tax
```

**Format**

An object of class `matrix` (inherits from `array`) with 7665 rows and 7 columns.

---

momspiCytokines

---

Create MOMS-PI cytokines SummarizedExperiment object

---

## Description

The Multi-Omic Microbiome Study-Pregnancy Initiative (MOMS-PI) was funded by the NIH Roadmap HUmAn Microbiome Project to understand the impact of the vaginal microbiome on pregnancy and the fetal microbiome. This longitudinal study contains samples from three body sites of 116 women. This summarizedExperiment object contains the cytokine data that was collected and the participant metadata.

## Usage

```
momspiCytokines()
```

## Format

A SummarizedExperiment object with 29 features and 1396 samples.

### colData:

**file\_id** File identifier

**md5** md5 hash for the file

**size** file size

**urls** URL for the file

**sample\_id** Sample identifier

**file\_name** Filename which the sample was taken from

**subject\_id** Participant identifier

**sample\_body\_site** Body site of the sample

**visit\_number** Visit number

**subject\_gender** Participant gender

**subject\_race** Participant race

**study\_full\_name** Name of the study

**project\_name** Name of project

## Value

A SummarizedExperiment object

## Examples

```
momspiCyto <- momspiCytokines()
```

---

|                |                       |
|----------------|-----------------------|
| momspiCyto_mtx | <i>momspiCyto_mtx</i> |
|----------------|-----------------------|

---

**Description**

Cytokines matrix for MOMS-PI data

**Usage**

```
momspiCyto_mtx
```

**Format**

An object of class `matrix` (inherits from `array`) with 29 rows and 1396 columns.

---

|                 |                        |
|-----------------|------------------------|
| momspiCyto_samp | <i>momspiCyto_samp</i> |
|-----------------|------------------------|

---

**Description**

Sample data for MOMS-PI cytokines data

**Usage**

```
momspiCyto_samp
```

**Format**

An object of class `data.frame` with 1396 rows and 13 columns.

---

|                  |                                          |
|------------------|------------------------------------------|
| momspiMultiAssay | <i>Make MOMS-PI MultiAssayExperiment</i> |
|------------------|------------------------------------------|

---

**Description**

Construct `MultiAssayExperiment` for MOMS-PI 16S rRNA and cytokines data.

**Usage**

```
momspiMultiAssay()
```

**Format**

A MultiAssayExperiment object with a 16S rRNA matrix and Cytokine matrix

**16S:** A counts matrix for the 16S rRNA-seq results.

**cytokines:** A counts matrix for the cytokines results.

**colData:**

**file\_id** File identifier

**md5** md5 hash for the file

**size** file size

**urls** URL for the file

**sample\_id** Sample identifier

**file\_name** Filename which the sample was taken from

**subject\_id** Participant identifier

**sample\_body\_site** Body site of the sample

**visit\_number** Visit number

**subject\_gender** Participant gender

**subject\_race** Participant race

**study\_full\_name** Name of the study

**project\_name** Name of project

**Value**

A multiAssay Experiment object

**Examples**

```
momspiMA <- momspiMultiAssay()
```

---

patient\_table

*Generate patient summary table for HMP2 sample data*

---

**Description**

This function allows you to produce a summary table for the HMP2Data data sets.

**Usage**

```
patient_table(x)
```

**Arguments**

**x** A named list of phyloseq or SummarizedExperiment objects.

**Value**

A knitr::kable table.



**Examples**

```
patient_table(list(momspi16S = momspi16S(),
  momspiCytokines = momspiCytokines(),
  IBD16S = IBD16S(), T2D16S = T2D16S()))
```

T2D16S

---

*Create HMP2 16S sequencing data for Type 2 Diabetes (T2D) phyloseq object*

---

**Description**

The HMP2 study gathered data on the microbiome for patients with different disease. This dataset contains 16S rRNA data for 79 patients with T2D over multiple visits.

**Usage**

```
T2D16S()
```

**Format**

A phyloseq object with 12,062 taxa and 2,208 samples.

**OTU Table:** A counts matrix for the rRNA-seq results.

**Sample Data:**

**file\_id** File identifier

**md5** md5 hash for the file

**size** file size

**urls** URL for the file

**sample\_id** Sample identifier

**file\_name** Filename which the sample was taken from

**subject\_id** Participant identifier

**sample\_body\_site** Body site of the sample

**visit\_number** Visit number

**subject\_gender** Participant gender

**subject\_race** Participant race

**study\_full\_name** Name of the study

**project\_name** Name of project

**Taxonomy Table:** Taxonomy table for the taxa identified.

**Value**

a phyloseq object

**Examples**

```
T2D <- T2D16S()
```

---

|            |                   |
|------------|-------------------|
| T2D16S_mtx | <i>T2D16S_mtx</i> |
|------------|-------------------|

---

**Description**

T2D 16S rRNA matrix

**Usage**

T2D16S\_mtx

**Format**

An object of class `matrix` (inherits from `array`) with 12062 rows and 2208 columns.

---

|             |                    |
|-------------|--------------------|
| T2D16S_samp | <i>T2D16S_samp</i> |
|-------------|--------------------|

---

**Description**

T2D 16S sample data

**Usage**

T2D16S\_samp

**Format**

An object of class `data.frame` with 2208 rows and 13 columns.

---

|            |                   |
|------------|-------------------|
| T2D16S_tax | <i>T2D16S_tax</i> |
|------------|-------------------|

---

**Description**

Taxonomy data for T2D 16S data

**Usage**

T2D16S\_tax

**Format**

An object of class `matrix` (inherits from `array`) with 12062 rows and 7 columns.

---

table\_two

*Generate demographic summary table for HMP2 sample data*


---

**Description**

This function allows you to produce a summary table for the HMP2Data data sets.

**Usage**

```
table_two(x)
```

**Arguments**

x                      A named list of phyloseq or SummarizedExperiment objects.

**Value**

A knitr::kable table.

**Examples**

```
table_two(list(momspi16S = momspi16S(),
  momspiCytokines = momspiCytokines(),
  IBD16S = IBD16S(), T2D16S = T2D16S()))
```

---

visit\_table

*Make table of visits*


---

**Description**

This function allows you to produce a summary table for the HMP2Data visits.

**Usage**

```
visit_table(x)
```

**Arguments**

x                      A named list of phyloseq or SummarizedExperiment objects.

**Value**

A knitr::kable table.

**Examples**

```
visit_table(list(momspi16S = momspi16S(),
  momspiCytokines = momspiCytokines(),
  IBD16S = IBD16S(), T2D16S = T2D16S()))
```

# Index

## \* datasets

- IBD16S\_mtx, [3](#)
- IBD16S\_samp, [3](#)
- IBD16S\_tax, [4](#)
- momspi16S\_mtx, [5](#)
- momspi16S\_samp, [5](#)
- momspi16S\_tax, [5](#)
- momspiCyto\_mtx, [7](#)
- momspiCyto\_samp, [7](#)
- T2D16S\_mtx, [10](#)
- T2D16S\_samp, [10](#)
- T2D16S\_tax, [10](#)

- IBD16S, [2](#)
- IBD16S\_mtx, [3](#)
- IBD16S\_samp, [3](#)
- IBD16S\_tax, [4](#)

- momspi16S, [4](#)
- momspi16S\_mtx, [5](#)
- momspi16S\_samp, [5](#)
- momspi16S\_tax, [5](#)
- momspiCyto\_mtx, [7](#)
- momspiCyto\_samp, [7](#)
- momspiCytokines, [6](#)
- momspiMultiAssay, [7](#)

- patient\_table, [8](#)

- T2D16S, [9](#)
- T2D16S\_mtx, [10](#)
- T2D16S\_samp, [10](#)
- T2D16S\_tax, [10](#)
- table\_two, [11](#)

- visit\_table, [11](#)