

Package ‘OSTA.data’

April 24, 2025

Title OSTA book data

Depends R (>= 4.5)

Version 1.0.0

Description 'OSTA.data' is a companion package for the ``Orchestrating Spatial Transcriptomics Analysis" (OSTA) with Bioconductor online book. Throughout OSTA, we rely on a set of publicly available datasets that cover different sequencing- and imaging-based platforms, such as Visium, Visium HD, Xenium (10x Genomics) and CosMx (NanoString). In addition, we rely on scRNA-seq (Chromium) data for tasks, e.g., spot deconvolution and label transfer (i.e., supervised clustering). These data been deposited in an Open Storage Framework (OSF) repository, and can be queried and downloaded using functions from the 'osfr' package. For convenience, we have implemented 'OSTA.data' to query and retrieve data from our OSF node, and cache retrieved Zip archives using 'BiocFileCache'.

Imports osfr, utils, BiocFileCache

Suggests BiocStyle, DropletUtils, knitr, VisiumIO, SpatialExperimentIO

biocViews DataImport, DataRepresentation, ExperimentHubSoftware, Infrastructure, ImmunoOncology, GeneExpression, Transcriptomics, SingleCell, Spatial

Encoding UTF-8

RoxygenNote 7.3.2

License Artistic-2.0

VignetteBuilder knitr

URL <https://github.com/estellad/OSTA.data>

BugReports <https://github.com/estellad/OSTA.data>

git_url <https://git.bioconductor.org/packages/OSTA.data>

git_branch RELEASE_3_21

git_last_commit 6d32e00

git_last_commit_date 2025-04-15

Repository Bioconductor 3.21

Date/Publication 2025-04-23

Author Yixing E. Dong [aut, cre] (ORCID:
<https://orcid.org/0009-0003-5115-5686>),
 Helena L. Crowell [aut] (ORCID:
<https://orcid.org/0000-0002-4801-1767>),
 Vince Carey [aut] (ORCID: <https://orcid.org/0000-0003-4046-0063>)

Maintainer Yixing E. Dong <estelladong729@gmail.com>

Contents

OSTA.data	2
Index	4

OSTA.data	<i>Datasets used in OSTA</i>
-----------	------------------------------

Description

Datasets used in OSTA

Usage

```
OSTA.data_list(url = "https://osf.io/5n4q3")
```

```
OSTA.data_load(
  id,
  bfc = BiocFileCache(),
  url = "https://osf.io/5n4q3",
  pol = TRUE,
  mol = TRUE
)
```

Arguments

url	character string; an OSF identifier (passed to osf_retrieve_node)
id	character string; dataset identifier (see OSTA.data_list for valid options)
bfc	BiocFileCache instance giving the location of files stored on disk
pol, mol	logical scalar; specifies whether or not to retrieve boundaries and molecules, respectively

Details

The following datasets are currently available:

- Janesick *et al.*: Visium, Visium HD, Xenium and Chromium data on human colorectal cancer; source: [10x Genomics](#)
- Oliveira *et al.*: Visium, Xenium and Chromium data on human breast cancer; source: [GSE243280](#)
- 1k-plex CosMx data on 2 mouse brain sections (coronal hippocampus and cortex, coronal hemisphere) source: [NanoString](#)
- 6k-plex CosMx data on human brain (frontal cortex); source: [NanoString](#)

Value

- `OSTA.data_list` returns a character vector of currently available datasets
- `OSTA.data_load` return a character string; the path to a Zip archive

References

Oliveira *et al.* Characterization of immune cell populations in the tumor microenvironment of colorectal cancer using high definition spatial profiling. *bioRxiv* 2024.06.04.597233 (2024).

Janesick *et al.* High resolution mapping of the tumor microenvironment using integrated single-cell, spatial and in situ analysis. *Nature Communications* 14:8353 (2023).

Examples

```
# view available datasets
OSTA.data_list()

# retrieve 10x Xenium data from Oliveira et al.
id <- "Xenium_HumanColon_Oliveira"
pa <- OSTA.data_load(id)

# unpack & view contents
td <- tempfile()
dir.create(td)
unzip(pa, exdir=td)
list.files(td)

# read into 'SpatialExperiment'
library(SpatialExperimentIO)
(spe <- readXeniumSXE(td))
```

Index

`BiocFileCache`, [2](#)

`osf_retrieve_node`, [2](#)

`OSTA.data`, [2](#)

`OSTA.data_list`, [2](#)

`OSTA.data_list (OSTA.data)`, [2](#)

`OSTA.data_load (OSTA.data)`, [2](#)