Package ‘timecoursedata’

August 8, 2024

Title  A data package for timecourse RNA-seq and microarray gene expression data sets

Version  1.14.0

Description  This data package contains timecourse gene expression data sets. The first dataset, from Shoemaker et al, consists of microarray samples from lung tissue of mice exposed to different influenza strains from 14 timepoints. The two other datasets are leaf and root samples from sorghum crops exposed to pre- and post-flowering drought stress and a control condition, sampled across the plants lifetime.

Depends  R (>= 4.0), SummarizedExperiment

Suggests  testthat (>= 1.0.0), knitr, rmarkdown, markdown, covr

VignetteBuilder  knitr

License  BSD 3-clause License + file LICENSE

Encoding  UTF-8

LazyData  false

RoxygenNote  7.1.1

biocViews  ExpressionData, MicroarrayData, RNASeqData

git_url  https://git.bioconductor.org/packages/timecoursedata

git_branch  RELEASE_3_19

git_last_commit  3e95200

Date/Publication  2024-08-08

Author  Elizabeth Purdom [aut] (<https://orcid.org/0000-0001-9455-7990>),
Nelle Varoquaux [aut, cre] (<https://orcid.org/0000-0002-8748-6546>)

Maintainer  Nelle Varoquaux <nelle.varoquaux@gmail.com>
Contents

load_shoemaker2015 ......................................................... 2
load_varoquaux2019 ....................................................... 2
shoemaker2015 ............................................................. 3
varoquaux2019leaf .......................................................... 4

Index 7

load_shoemaker2015  Loading the shoemaker2015 dataset as SummarizedExperiment

Description

See data(shoemaker2015) for a full description of the dataset and the metadata.

Usage

load_shoemaker2015()

Value

a SummarizedExperiment of the Shoemaker 2015 data

Examples

mice_timecourse_data = load_shoemaker2015()
head(mice_timecourse_data)

load_varoquaux2019  Loading the varoquaux2019 leaf or root dataset as SummarizedExperiment

Description

See data(varoquaux2019leaf) or data(varoquaux2019root) for a full description of the dataset and the metadata.

Usage

load_varoquaux2019(sample_type = "leaf")

Arguments

sample_type  "leaf", "root", optional, default: leaf Sets which sample to load.
Value

a SummarizedExperiment of the leaf or root samples of Varoquaux et al.

Examples

```r
sorghum_timecourse_data = load_varoquaux2019(sample_type="leaf")
head(sorghum_timecourse_data)
```

Description

Time-course microarray data from "An Ultrasensitive Mechanism Regulates Influenza Virus-Induced Inflammation".

Usage

data(shoemaker2015)

Format

A list with two related datasets: a gene expression data set (element `data`) and corresponding meta data (element `meta`). The column names of `data` correspond to the rownames of `meta`, linking the samples together in the two datasets. See examples for accessing these two data.frames.

Details

This is data from a micro-array time-course experiment, exposing mice to three different strains of influenza, and collecting lung tissue during 14 time-points after infection (0, 3, 6, 9, 12, 18, 24, 30, 36, 48, 60 hours, then 3, 5, and 7 days later) [8]. The three strains of influenza used in the study are (1) a low pathogenicity seasonal H1N1 influenza virus (A/Kawasaki/UTK4/2009 [H1N1]), a mildly pathogenic virus from the 2009 pandemic season (A/California/04/2009 [H1N1]), and a highly pathogenic H5N1 avian influenza virus (A/Vietnam/1203/2004 [H5N1]). Mice were injected with 105 PFU of each virus. An additional 42 mice were injected with a lower dose of the Vietnam avian influenza virus (103 PFU).

data is a data frame with 39544 rows corresponding to genes and 209 corresponding to samples. The rownames give the RefSeq name of the gene.

meta is a data frame with 209 rows corresponding to samples and 3 named columns (the first column is just an index from 0-208):


- **Replicate** Identifies the replicate – each combination of treatment and timepoint was replicated three times (except for VH at timepoint 3, which has only 2 replicates).

- **Timepoint** Identifies the time passed (in days) since infection of the sample
References


Examples

data(shoemaker2015)
data = attr(shoemaker2015, "data")
meta = attr(shoemaker2015, "meta")

---

varoquaux2019leaf  Time course transcriptomic data from leaf and root tissues exposed to drought stress

Description

Time-course RNA-seq leaf and root data from "Transcriptomic analysis of field-droughted sorghum from seedling to maturity reveals biotic and metabolic responses"

Usage

data(varoquaux2019leaf)
data(varoquaux2019root)

Format

A list with two related datasets: a gene expression data set (element data) and corresponding meta data (element meta). The column names of data correspond to the rownames of meta, linking the samples together in the two datasets. See examples for accessing these two data.frames.

Details

This are two datasets from a mRNA-Seq time-course experiment, measuring gene expression across time in the plant sorghum. One dataset corresponds to the leaf of the plant (varoquaux2019leaf), the other to the root from the same sets of plants (varoquaux2019leaf). Each sample sequenced is actually a pool of at least 10 plants. Note that leaf and root samples were processed and sequenced as two separate batches. Any comparison between the two datasets should be traded with care.

data is a data frame with 34,211 rows corresponding to sorghum genes with (Phytozome v3.0 gene ids). Root has 198 columns and leaf has 197 columns, corresponding to samples. The root and leaf samples are from the same plants, and have ids in the form of "0622162L05" or "0622162R05", with "L" corresponding to the leaf and "R" corresponding to the root sample. Some samples were removed for quality control reasons, resulting in slightly different numbers of samples between the leaf and root.

The data provided is the raw data. Note that leafs and roots were sequenced as two different batches.

meta is a data frame with rows corresponding to samples and and columns with meta data. These columns are the same across leaf and root, allowing for easy reuse of code across the datasets:

Barcode  Barcode of the sample, same as row name
libraryName  Internal Barcode given to the sample by the sequencing facility (JGI)
Block  Plot from which the sample was taken
Week  Numeric value corresponding to the week sample was taken
Replicate  Each condition combination was replicated three times, but not all replicates were se-
  quenced. Each replicate came from a different plot. The replicate id tracks this plot for each
  condition. However, replicate 1 in one condition is a completely different plot from replicate
  1 in another condition, though the replicates were physically close to each other in the field.
Genotype  Genotype (i.e. variety) of the sample, one of BT642 or RT430
Condition  The drought condition, one of "Control" (regular watering), "Postflowering" (drought
  applied post flowering of the plant), or "Preflowering" (drought applied pre-flowering of the
  plant, and then watering resumed after flowering)
Sample.type  Either "Leaves" or "Root", only relevant if merging the two datasets
Day  The day sample collected, in form of "Day14" (corresponding to Week 2)
Row  Row plant was sampled from within the block
isFloweringWeek  logical, indicating whether the week sampled corresponding to the plant flow-
  ering
Date.Harvested  The date sample was collected
Collection.Time  The time of day plant was collected
No.plants.pooled  The number of plants pooled for the sample
FloweringJulianDate  The Julian date of flowering for the block the sample was collected from
FloweringDate  The date of flowering for the block the sample was collected from
FloweringWeek  The week of flowering for the block the sample was collected from
FloweringDay  The day of flowering for the block the sample was collected from
Unreplicated  Identifies timepoints with no replicates
CompleteSampleTreatment  Full set of factors (Genotype x Treatment x Week)

The following variables are features measured at the end of planting when the remaining plants
from the plot were harvested:

LbsGrainHarvestPerPlot
TonsOfGrain.ac
PctGrainMoisture
TonsOfGrain.acAt13PctMoisture
X1000SeedWt_g
ApproximateEndOfSeasonStandCount_aveOfGrainRows
TotalFreshForageWtPerPlot_kg
TotalFreshForageWtPerPlot_lbs
TonsOfForage.ac
PctMoistureForage
TonsOfForage.acAt65PctMoisture
ApproximateEndOfSeasonStandCount_aveOfForageRows
DryBiomass
FreshBiomass
AveHeight

The following variables are features regarding the quality of the mapping of the sequences from the sample:

rawReads
filteredReads
sequencerType
runType
totalFragments
mappedFragments
assignedFragments
unassignedAmbiguous
unassignedNoFeatures
unassignedSecondaryHits
ratioStrandedness
MappedRatio
AssignedRatio
Plate
Location

References
Varoquaux et al, 2019 (PNAS)

Examples

data(varoquaux2019leaf)
dataLeaf = attr(varoquaux2019leaf, "data")
metaLeaf = attr(varoquaux2019leaf, "meta")
data(varoquaux2019root)
dataRoot = attr(varoquaux2019root, "data")
metaRoot = attr(varoquaux2019root, "meta")
Index

* datasets
  shoemaker2015, 3
  varoquaux2019leaf, 4

load_shoemaker2015, 2
load_varoquaux2019, 2

shoemaker2015, 3
varoquaux2019leaf, 4
varoquaux2019root (varoquaux2019leaf), 4