

# Package ‘SuperCurveSampleData’

August 4, 2017

**Version** 1.0.2

**Date** 2015-07-21

**Title** 'SuperCurve' RPPA Sample Data

**Author** P. Roebuck

**Maintainer** ``P. Roebuck" <proebuck@mdanderson.org>

**Description** Reverse phase protein lysate array sample data for 'SuperCurve'.

**Depends** R (>= 2.15)

**NeedsCompilation** no

**URL** <http://r-forge.r-project.org/projects/supercurve/>

**License** Artistic-2.0

**Copyright** file COPYRIGHTS

**LazyData** no

**LazyLoad** no

**biocViews** ExperimentData, Microarray

**Repository** R-Forge

**Repository/R-Forge/Project** supercurve

## R topics documented:

rppaCell-extdata . . . . .	2
rppaSingleSubgrid-extdata . . . . .	3
rppaTriple-extdata . . . . .	5
rppaTumor-extdata . . . . .	7
<b>Index</b>	<b>9</b>

---

rppaCell-extdata*AKT, ERK2, and CTNNB1 expression in cell lines*

---

**Description**

This data set contains the expression levels of three proteins: AKT, ERK2, and beta catenin (CTNNB1) in 40 cell lines, measured in duplicate dilution series using reverse-phase protein arrays.

The data set also contains a description of the design of the reverse-phase protein array used in a set of experiments to measure protein levels in 40 different cell lines. Cell lysates were spotted on the array in duplicate in eight-step dilution series. The layout of the array consisted of a grid of 4x4 subgrids. Each subgrid contained the duplicate dilution series for a single sample. Each of the identical top two rows of a subgrid contained the four most concentrated dilution steps (in decreasing concentrations from left to right), and the identical bottom two rows contained the four least concentrated dilution steps.

**Format**

The MicroVigene TXT datafiles contain the following columns:

- Main Row
- Main Col
- Sub Row
- Sub Col
- GeneID
- mean\_net
- mean\_total
- median\_net
- vol\_bkg
- vol\_dust

The file 'proteinAssay.tsv' contains the following columns:

- Antibody
- Filename

The file 'cellLineInfo.tsv' contains the following columns:

- CellLine
- ProteinConcentration
- LoadingControl

**Author(s)**

P. Roebuck <proebuck@mdanderson.org>

**Source**

Bryan Hennessey and Gordon Mills

## Examples

```
## Path to raw data directory
extdata.dir <- system.file("extdata", package="SuperCurveSampleData")
rawdata.dir <- file.path(extdata.dir, "rppaCellData")
rawdata.dir

## Directory Listing
files <- list.files(rawdata.dir, full.names=TRUE)
files
```

---

```
rppaSingleSubgrid-extdata
```

*FOO, BAR, PLUGH, and WALDO expression*

---

## Description

This data set contains the expression levels of four imaginary proteins: FOO, BAR, PLUGH, and WALDO. Each sample on the RPPA slide is printed as a 5-step dilution series. There are 4x12 subgrids on each slide, corresponding to the 48 pins on the printer. Each subgrid contains 11 subrows and 11 subcolumns. Each subrow is composed of two 5-step dilution series from left to right, each series itself from most concentrated to most diluted, and a positive control spot. The eleventh subcolumn is composed of two mirrored 5-step positive control series, with most concentrated on top/bottom and most diluted near the middle; the middle is a negative control (typically blank).

When these slide images were quantified using MicroVigene, the whole slide was treated as a single subgrid (1x1x44x132) to save time (by not having to perform alignment on 48 individual subgrids).

The data set also contains a description of the design of the reverse-phase protein array used in a set of experiments to measure protein levels. The associated MicroVigene datafiles were converted back into their logical format (4x12x11x11) when read to facilitate downstream processing.

## Format

The MicroVigene TXT datafiles contain the following columns:

- Main Row
- Main Col
- Sub Row
- Sub Col
- GeneID
- mean\_net
- mean\_total
- mean\_bkg
- bkg\_used
- mean\_dust
- median\_net
- median\_total
- median\_bkg

- vol\_net
- vol\_total
- vol\_bkg
- vol\_dust
- cv\_spot
- cv\_bkg
- xcenter
- ycenter
- area\_signal
- area\_spot
- area\_bkg
- solidity
- circularity
- roundness
- aspect
- dustiness
- rank
- type

The file 'proteinAssay.tsv' contains the following columns:

- Antibody
- Filename

The file 'slidedesign.tsv' contains the following columns:

- Main.Row
- Main.Col
- Sub.Row
- Sub.Col
- SubgridAlias
- SpotType
- Dilution
- GridAlias

The file 'slidedesign-singlesubgrid.tsv' contains the following columns:

- Main.Row
- Main.Col
- Sub.Row
- Sub.Col
- SubgridAlias
- SpotType
- Dilution
- GridAlias
- Main.Row.Real
- Main.Col.Real
- Sub.Row.Real
- Sub.Col.Real

**Author(s)**

P. Roebuck <proebuck@mdanderson.org>

**Examples**

```
## Path to raw data directory
extdata.dir <- system.file("extdata", package="SuperCurveSampleData")
rawdata.dir <- file.path(extdata.dir, "rppaSingleSubgridData")
rawdata.dir

## Directory Listing
files <- list.files(rawdata.dir, full.names=TRUE)
files
```

---

rppaTriple-extdata	<i>ACTB, CAS3, FAK, and ODC1 expression in 14 fed/starved cell lines</i>
--------------------	--

---

**Description**

This data set contains the expression levels of four proteins: beta-Actin (ACTB), Caspase 3 (CAS3), Focal adhesion kinase (FAK), and Ornithine decarboxylase (ODC1) from a study that was done to compare protein levels in 14 cell lines from both a “fed” and a “starved” state. There are two files included for beta-Actin, one that was scanned in color (*actb*) and the other in 16-bit grayscale (*actb.gray*); all other proteins were scanned in color.

This data set also contains a description of the design used for the reverse-phase protein arrays from this study. Cell lysates were spotted on the array in six-step dilution series. The layout of the array consists of a grid of 6x6 subgrids. The first three rows of a subgrid contain 3 replicates of a cell line. The last three rows contain 3 replicates of another cell line. Each subgrid is replicated on the array 3 times, so that there are a total of 9 replicates per cell line per state. The top part of the array contains the fed cell lines and the bottom part of the array contains the starved cell lines. There is one subgrid on the array that contains only buffer material and another subgrid that did not have anything printed (blank). There are a total of 18 spots each of buffer and blank material.

**Format**

The MicroVigene TXT datafiles contain the following columns:

- Main Row
- Main Col
- Sub Row
- Sub Col
- GeneID
- mean\_net
- mean\_total
- mean\_bkg
- bkg\_used
- mean\_dust
- median\_net

- median\_total
- median\_bkg
- vol\_net
- vol\_total
- vol\_bkg
- vol\_dust
- cv\_spot
- cv\_bkg
- xcenter
- ycenter
- area\_signal
- area\_spot
- area\_bkg
- solidity
- circularity
- roundness
- aspect
- dustiness
- rank
- type

The file 'proteinAssay.tsv' contains the following columns:

- Antibody
- Filename
- Alias

The file 'layoutInfo.tsv' contains the following columns:

- Alias
- Sample

### Author(s)

P. Roebuck <proebuck@mdanderson.org>

### Source

Victor Levin

### Examples

```
## Path to raw data directory
extdata.dir <- system.file("extdata", package="SuperCurveSampleData")
rawdata.dir <- file.path(extdata.dir, "rppaTripleData")
rawdata.dir

## Directory Listing
files <- list.files(rawdata.dir, full.names=TRUE)
files
```

---

rppaTumor-extdata	<i>ERK2, GSK3, and JNK expression in tumor samples</i>
-------------------	--

---

## Description

This data set contains the expression levels of three proteins: ERK2, GSK3, and JNK in 96 breast tumor samples and controls, measured in dilution series using reverse-phase protein arrays.

This data set also contains a description of the design of the reverse-phase protein array used in a set of experiments to measure protein levels. Cell lysates were spotted on the array in seven-step dilution series with either a positive or negative control at the end of the series. The layout of the array consisted of a grid of 4x4 subgrids. The first two rows of a subgrid contained a single dilution series and a negative control spot. The last two rows of the subgrid contained another dilution series and a positive control spot.

## Format

The MicroVigene TXT datafiles contain the following columns:

- Main Row
- Main Col
- Sub Row
- Sub Col
- GeneID
- mean\_net
- mean\_total
- median\_net
- vol\_net
- vol\_bkg
- vol\_dust

The file 'proteinAssay.tsv' contains the following columns:

- Antibody
- Filename

The file 'layoutInfo.tsv' contains the following columns:

- Alias
- Sample

The file 'slidedesign.tsv' contains the following columns:

- Main.Row
- Main.Col
- Sub.Row
- Sub.Col
- SubgridAlias

- SpotType
- Dilution

The file 'supercurve.csv' contains the following columns:

- CA19-9
- ERK2
- GSK3
- JNK
- p38

**Author(s)**

P. Roebuck <proebuck@mdanderson.org>

**Source**

Doris Swank and Gordon Mills

**Examples**

```
## Path to raw data directory
extdata.dir <- system.file("extdata", package="SuperCurveSampleData")
rawdata.dir <- file.path(extdata.dir, "rppaTumorData")
rawdata.dir

## Directory Listing
files <- list.files(rawdata.dir, full.names=TRUE)
files
```



# Index

## \*Topic **datasets**

- rppaCell-extdata, [2](#)
- rppaSingleSubgrid-extdata, [3](#)
- rppaTriple-extdata, [5](#)
- rppaTumor-extdata, [7](#)

- rppaCell-extdata, [2](#)
- rppaSingleSubgrid-extdata, [3](#)
- rppaTriple-extdata, [5](#)
- rppaTumor-extdata, [7](#)