Package 'msDilution'

April 23, 2008

Type Package

Title Protein Mass Spectra Dataset from a Dilution Experiment

Version 1.0.1

Date Fri Apr 18 14:54:53 PDT 2008

Author Lixin Gong, William Constantine, and Yu Alex Chen

Maintainer Lixin Gong <lgong@insightful.com>

Depends R (>= 2.5.0)

DependsSplus

Suggests msProcess

Description This package provides a dataset of 280 MALDI-TOF mass spectra generated from a dilution experiment aimed at elucidating which features in MALDI-TOF mass spectrometry data are informative for quantifying peptide content.

License GNU General Public License Version 2

Copyright 2007-2008 Insightful Corporation. All rights reserved.

URL http://www.insightful.com/services/research/proteome/default.asp

R topics documented:

1

`Dilution2005Raw`   *Dilution 2005 Raw*

Description

A data object of class `msSet` consisting of 280 MALDI-TOF mass spectra. These data come from a dilution experiment aimed at elucidating which features in MALDI-TOF mass spectrometry data are informative for quantifying peptide content. The details of the experiment are described in [2].

The primary dataset consists of 250 spectra collected from 5 different serum sources (5 people from a health study), each subject to 10 different concentrations of a peptide mixture that contained several known peptides. Each of the 50 prepared samples were spotted, randomly, 5 times each on a single plate producing a total of 5x10x5 = 250 spectra.

An additional 30 spectra arise from: 2 replicate spectra from each of the 10 concentrations of the peptide mixture, plus 2 replicates of serum-only spectra from each of the 5 serum samples.

NOTE: An error was made during the process of randomly spotting samples to the plate: one of the replicates from concentration 6 was spotted on top of a serum-only sample. The result is that two spectra from this design are missing: serum 1, concentration 6, replicate 4 (sample 36), and serum 4, concentration 0, replicate 1. This left one empty spot on the plate (sample 280) to which a sample containing only cytochrome c was spotted. The latter was not used in any subsequent analysis in [2], and no adjustment was made for former in the analysis in [2] (sample 36 was used as is).

Dilution2005Raw$coding is a list with 6 elements: "dir.org", "cyt.ind", "pep.ind", "ser.ind", "mix.ind", "err.ind". The "dir.org" is a matrix describing the source of the 280 MALDI spectra contained in this dataset. The others are vectors of indices for various types of samples: "cyt.ind" for cytochrome c, "pep.ind" for peptide mixture, "ser.ind" for serum-only, "mix.ind" for serum + peptide mixture, and "err.ind" for the bad spot.

Samples were randomly assigned to spots on the plate and this matrix decodes the positions of the various samples/concentrations/replicates. The column headings of "dir.org" are:

Index this just enumerates the 280 spectra.

Serum The numbers 1-5 indicate the different serum samples; 0 indicates peptide mix; 6 is cytochrome c (see NOTE above).

Conc The numbers 0-10 indicate the concentration of peptide mixture added to each serum sample (see [2] for details): 1=highest concentration, 2=(1/2 concentration of 1), 3=(1/2 the concentration of 2), ..., 10=(1/2 the concentration of 9), 0=serum only (no peptide mix added).

Replicate The numbers 1-5 enumerate the replicate spots (spectra) for each concentration and serum sample. Note there are 5 replicates for each of 10 concentration levels of peptide mix added to each serum sample, but only 2 replicates of each serum-only sample.

Manual The numbers 0 or 1 in this column indicate: 0=machine laser positioning; 1=manual laser positioning. Explanation: 14 of the spectra were deemed overly noisy (by the operator) as they were being collected during the machine's automatic positioning of the laser. These spots were re-shot after manually repositioning the laser on the spot. This information was not used in the analysis in [2].

Usage

```
Dilution2005Raw
```

Format

An object of msSet. See [msSet](#) in package [msProcess](#) for details.

References

Randolph TW and Yasui Y, "Multiscale processing of mass spectrometry data," Biometrics, 62:589–597, 2006.

Randolph TW, Mitchell BL, McLerran DF, Lampe PD, and Feng Z, "Quantifying peptide signal in MALDI-TOF mass spectrometry data," Molecular and Cellular Proteomics, 4(12):1990–1999, 2005.

Randolph TW, "Scale-based normalization of spectral data," Cancer Biomarkers, 2(3–4):135-144, 2006.

Examples

```
data(Dilution2005Raw, package="msDilution")

## illustration of how to get the indices for
## various types of samples
dir.org <- Dilution2005Raw$coding$dir.org

## order the indices based on "Serum", "Conc",
## "Replicate", and "Manual"
dir.ind <- order(dir.org[, "Serum"], dir.org[, "Conc"], dir.org[, "Replicate"], dir.org[,

## get the ordered indices
dir.ord <- dir.org[dir.ind,]

## index for cytochrome c
cyt.ind <- dir.ord[dir.ord[,"Serum"]==6, "Index"]

## indices for peptide mixture
pep.ind <- dir.ord[dir.ord[,"Serum"]==0, "Index"]

## indices for serum-only
ser.ind <- dir.ord[dir.ord[,"Serum"]!=6 & dir.ord[,"Conc"]==0, "Index"]

## indices for serum + peptide mixture
mix.ind <- dir.ord[dir.ord[,"Serum"] %in% 1:5 & dir.ord[,"Conc"]!=0, "Index"]

## index for the bad spot (serum 1, concentration
## 6, replicate 4 + serum 4, concentration 0,
## replicate 1)
err.ind <- dir.ord[dir.ord[,"Serum"]==1 & dir.ord[,"Conc"]==6 & dir.ord[,"Replicate"]==4,
```

Index