Cancer

Lieven Clement

statOmics, Ghent University (https://statomics.github.io)

Contents

This is part of the online course Proteomics Data Analysis (PDA)

1 Background

Twelve Estrogen Receptor Positive Breast cancer tissues from from patients treated with tamoxifen upon recurrence have been assessed in a proteomics study. Six patients had a good outcome (OR) and the other Six had a poor outcome (PD). The proteomes have been assessed using an LTQ-Orbitrap and the thermo output .RAW files were searched with MaxQuant (version 1.4.1.2) against the human proteome database (FASTA version 2012-09, human canonical proteome).

2 Data

We first import the data from peptide.txt file. This is the file containing your peptide-level intensities. For a MaxQuant search [6], this peptide.txt file can be found by default in the "path_to_raw_files/combined/txt/" folder from the MaxQuant output, with "path_to_raw_files" the folder where the raw files were saved.

We generate the object peptideFile with the path to the peptide.txt file. Using the `grepEcols` function, we find the columns that contain the expression data of the peptide in the peptide.txt file.

```
library(tidyverse)
library(limma)
library(QFeatures)
```

```
library(msqrob2)
library(plotly)

peptidesFile <- "https://raw.githubusercontent.com/statOmics/PDA22GTPB/data/quantification/cancer/peptic

ecols <- grep(
  "Intensity\\.",
  names(read.delim(peptidesFile))
  )
```

Next, we read the data and store it in QFeatures object

```
pe <- readQFeatures(
  assayData = read.delim(peptidesFile),
  fnames = 1,
  quantCols = ecols,
  name = "peptideRaw")
```

```
## Checking arguments.

## Loading data as a 'SummarizedExperiment' object.

## Formatting sample annotations (colData).

## Formatting data as a 'QFeatures' object.
```

The QFeatures object pe currently contains a single assay, named peptideRaw.

We extract the column names from the peptideRaw assay and see that this contains information about the prognosis.

```
colnames(pe[["peptideRaw"]])
```

```
##  [1] "Intensity.OR.01" "Intensity.OR.04" "Intensity.OR.07" "Intensity.OR.09"
##  [5] "Intensity.OR.10" "Intensity.OR.13" "Intensity.PD.02" "Intensity.PD.03"
##  [9] "Intensity.PD.04" "Intensity.PD.06" "Intensity.PD.07" "Intensity.PD.08"
```

We rename the colnames by dropping the "Intensity." from the name.

```
(newNames <-  sub(
  pattern = "Intensity\\.",
  replacement = "",
  colnames(pe[["peptideRaw"]]))
 )
```

```
##  [1] "OR.01" "OR.04" "OR.07" "OR.09" "OR.10" "OR.13" "PD.02" "PD.03" "PD.04"
## [10] "PD.06" "PD.07" "PD.08"
```

```
pe <- renameColname(pe,
                    i = "peptideRaw",
                    newNames)
pe <- renamePrimary(pe, newNames)
colnames(pe[["peptideRaw"]])
```

```
##  [1] "OR.01" "OR.04" "OR.07" "OR.09" "OR.10" "OR.13" "PD.02" "PD.03" "PD.04"
## [10] "PD.06" "PD.07" "PD.08"
```

In the following code chunk, we add the prognosis of the patients that we can read in the raw file name to the colData.

```
colData(pe)$prognosis <-
  colnames(pe[["peptideRaw"]]) %>%
  substr(start = 1, stop = 2) %>%
  as.factor
colData(pe)$prognosis
```

```
##  [1] OR OR OR OR OR OR PD PD PD PD PD PD
## Levels: OR PD
```

We calculate how many non zero intensities we have per peptide and this will be useful for filtering.

```
rowData(pe[["peptideRaw"]])$nNonZero <- rowSums(assay(pe[["peptideRaw"]]) > 0)
```

Peptides with zero intensities are missing peptides and should be represent with a `NA` value rather than `0`.

```
pe <- zeroIsNA(pe, "peptideRaw") # convert 0 to NA
```

Look at the column names of the data to know the variables that you can use for filtering.

```
pe[["peptideRaw"]] %>% rowData %>% names
```

```
##  [1] "Sequence"               "Proteins"               "Leading.razor.protein"
##  [4] "Gene.names"             "Protein.names"          "Unique..Groups."
##  [7] "Unique..Proteins."      "Charges"                "PEP"
## [10] "Score"                  "Slice.Average"          "Slice.Std..Dev."
## [13] "Slice.1"                "Unique.Slice.Average"   "Unique.Slice.Std..Dev."
## [16] "Unique.Slice.1"         "Experiment.OR.01"       "Experiment.OR.04"
## [19] "Experiment.OR.07"       "Experiment.OR.09"       "Experiment.OR.10"
## [22] "Experiment.OR.13"       "Experiment.PD.02"       "Experiment.PD.03"
## [25] "Experiment.PD.04"       "Experiment.PD.06"       "Experiment.PD.07"
## [28] "Experiment.PD.08"       "Intensity"              "Reverse"
## [31] "Contaminant"            "id"                     "Protein.group.IDs"
## [34] "Mod..peptide.IDs"       "Evidence.IDs"           "MS.MS.IDs"
## [37] "Best.MS.MS"             "Oxidation..M..site.IDs" "nNonZero"
```

So we will filter on the "Reverse", "Contaminant" and "nNonZero" column.

2.1 Data exploration

46% of all peptide intensities are missing and for some peptides we do not even measure a signal in any sample.

3 Preprocessing

This section preforms preprocessing for the peptide data. This include

- log transformation,
- filtering and
- summarisation of the data.

3.1 Log transform the data

```
pe <- logTransform(pe, base = 2, i = "peptideRaw", name = "peptideLog")
```

3.2 Filtering

1. Handling overlapping protein groups

In our approach a peptide can map to multiple proteins, as long as there is none of these proteins present in a smaller subgroup.

```
pe <- filterFeatures(pe, ~ Proteins %in% smallestUniqueGroups(rowData(pe[["peptideLog"]])$Proteins))
```

```
## 'Proteins' found in 2 out of 2 assay(s)
```

2. Remove reverse sequences (decoys) and contaminants

We now remove the contaminants and peptides that map to decoy sequences.

```
pe <- filterFeatures(pe,~Reverse != "+")
```

```
## 'Reverse' found in 2 out of 2 assay(s)
```

```
pe <- filterFeatures(pe,~ Contaminant != "+")
```

```
## 'Contaminant' found in 2 out of 2 assay(s)
```

3. Drop peptides that were only identified in one sample

We keep peptides that were observed at last twice.

```
pe <- filterFeatures(pe,~ nNonZero >=2)
```

```
## 'nNonZero' found in 2 out of 2 assay(s)
```

```
nrow(pe[["peptideLog"]])
```

```
## [1] 25452
```

We keep 25452 peptides upon filtering.

3.3 Normalize the data using median centering

We normalize the data by substracting the sample median from every intensity for peptide p in a sample i:

$$y_{ip}^{\text{norm}} = y_{ip} - \hat{\mu}_i$$

with $\hat{\mu}_i$ the median intensity over all observed peptides in sample i.

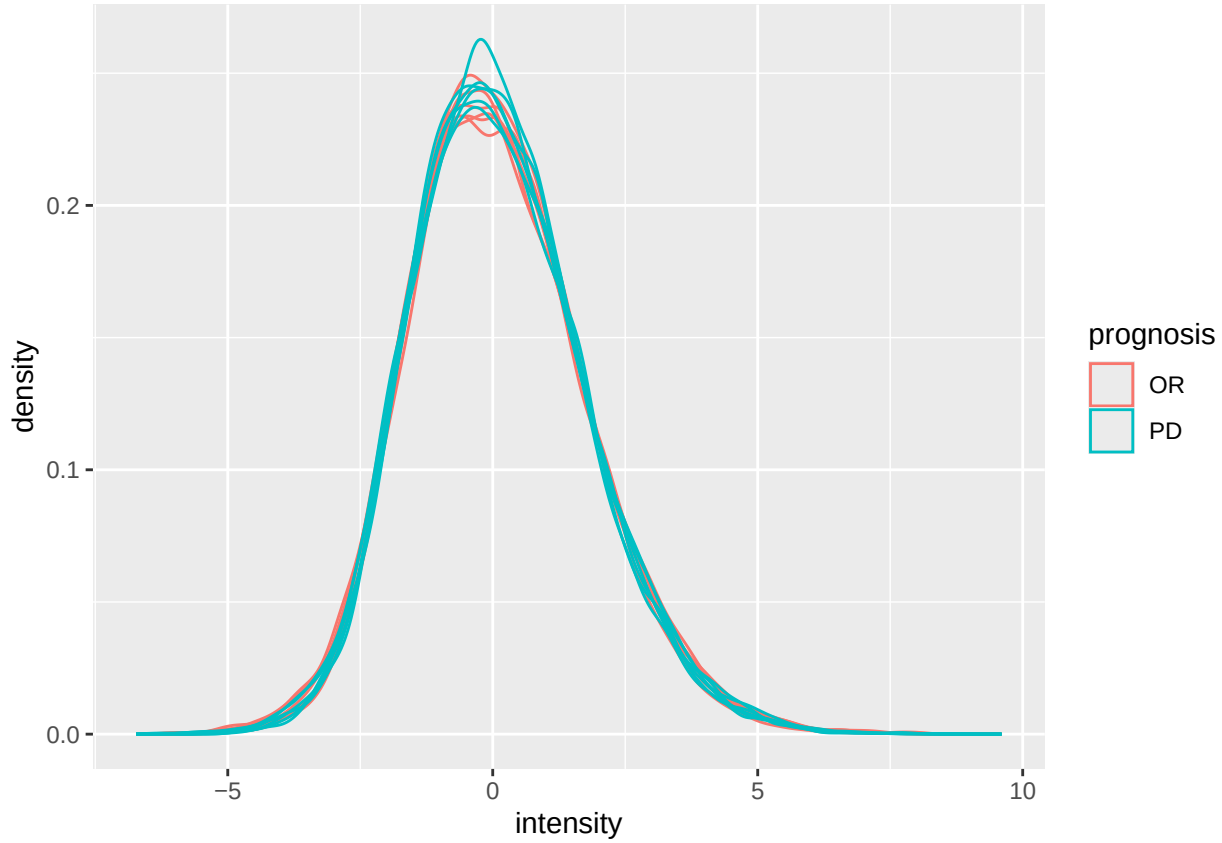
```
pe <- normalize(pe,
                i = "peptideLog",
                name = "peptideNorm",
                method = "center.median")
```

3.4 Explore normalized data

Upon the normalisation the density curves are nicely registered

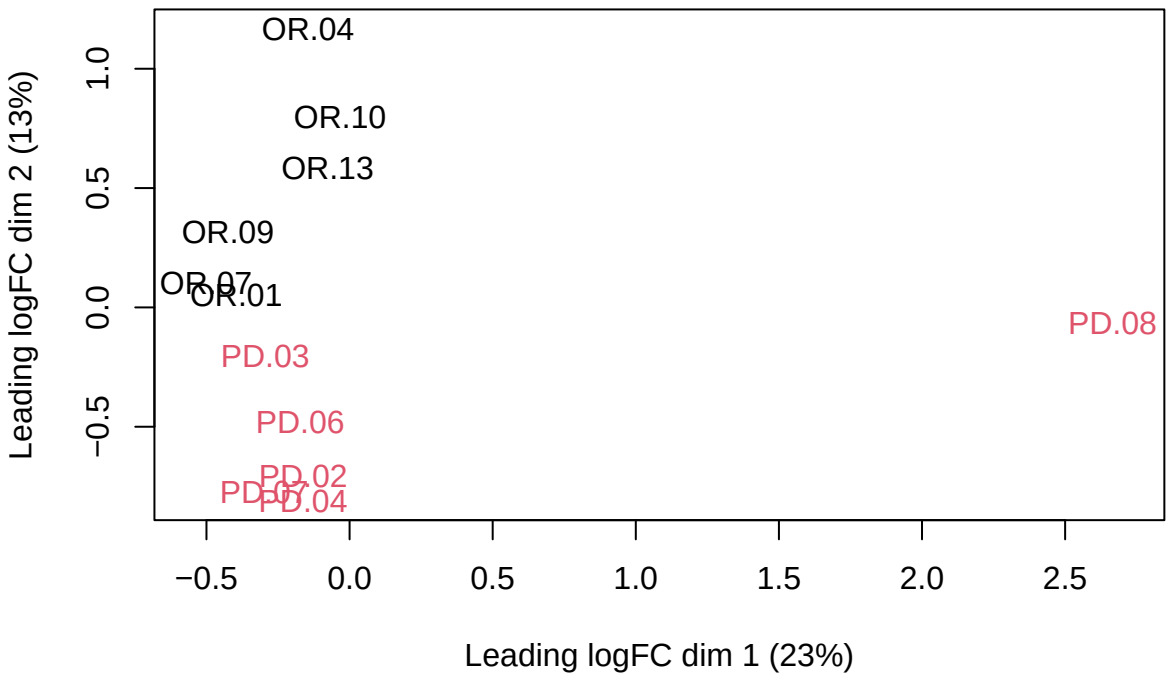
```
pe[["peptideNorm"]] %>%
  assay %>%
  as.data.frame() %>%
  gather(sample, intensity) %>%
  mutate(prognosis = colData(pe)[sample,"prognosis"]) %>%
  ggplot(aes(x = intensity,group = sample,color = prognosis)) +
    geom_density()
```

```
## Warning: Removed 108107 rows containing non-finite outside the scale range
## (`stat_density()`).
```



We can visualize our data using a Multi Dimensional Scaling plot, eg. as provided by the `limma` package.

```
pe[["peptideNorm"]] %>%
  assay %>%
  limma::plotMDS(col = as.numeric(colData(pe)$prognosis))
```



The first axis in the plot is showing the leading log fold changes (differences on the log scale) between the samples. We observe one outlying sample. In the second dimension we observe a separation according to prognosis.

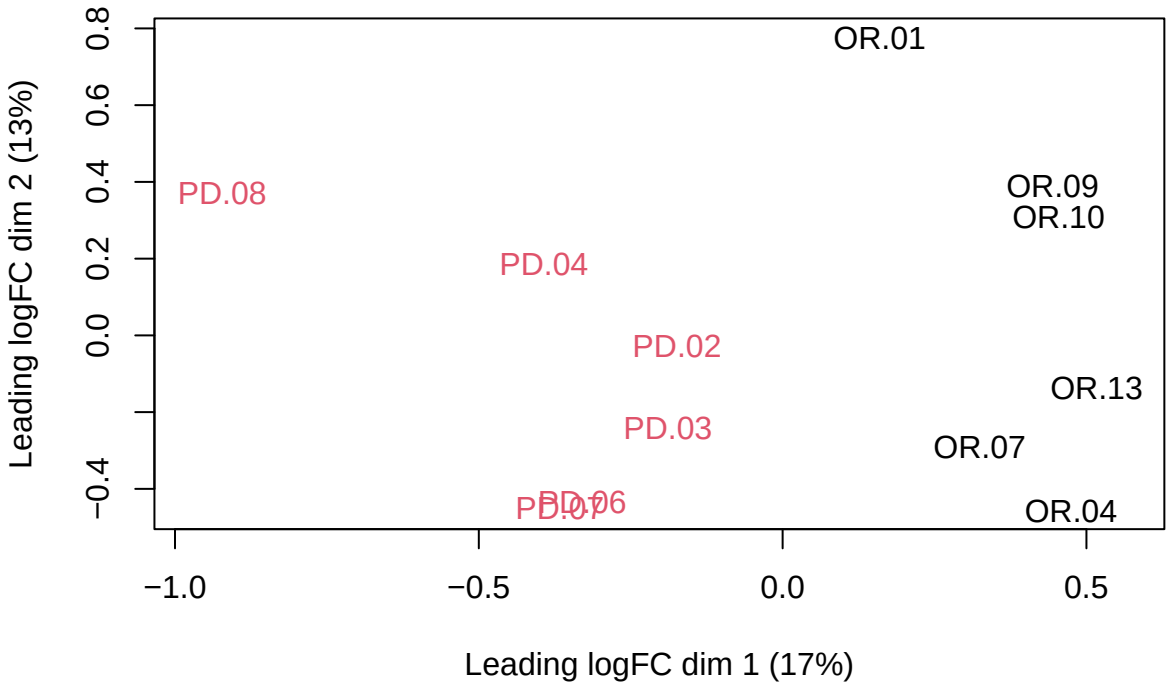
3.5 Summarization to protein level

- By default robust summarization is used: `fun = MsCoreUtils::robustSummary()`

```
pe <- aggregateFeatures(pe,
  i = "peptideNorm",
  fcol = "Proteins",
  na.rm = TRUE,
  name = "protein")
```

```
## Your quantitative and row data contain missing values. Please read the
## relevant section(s) in the aggregateFeatures manual page regarding the
## effects of missing values on data aggregation.
```

```
plotMDS(assay(pe[["protein"]]), col = as.numeric(colData(pe)$prognosis))
```



Note that the samples upon robust summarisation show a separation according to the prognosis.

4 Data Analysis

4.1 Estimation

We model the protein level expression values using `msqrob`. By default `msqrob2` estimates the model parameters using robust regression.

We will model the data with a different group mean. The group is incoded in the variable `prognosis` of the colData. We can specify this model by using a formula with the factor condition as its predictor: `formula = ~prognosis`.

Note, that a formula always starts with a symbol '~'.

```
pe <- msqrob(object = pe, i = "protein", formula = ~prognosis)
```

4.2 Inference

First, we extract the parameter names of the model by looking at the first model. The models are stored in the row data of the assay under the default name msqrobModels.

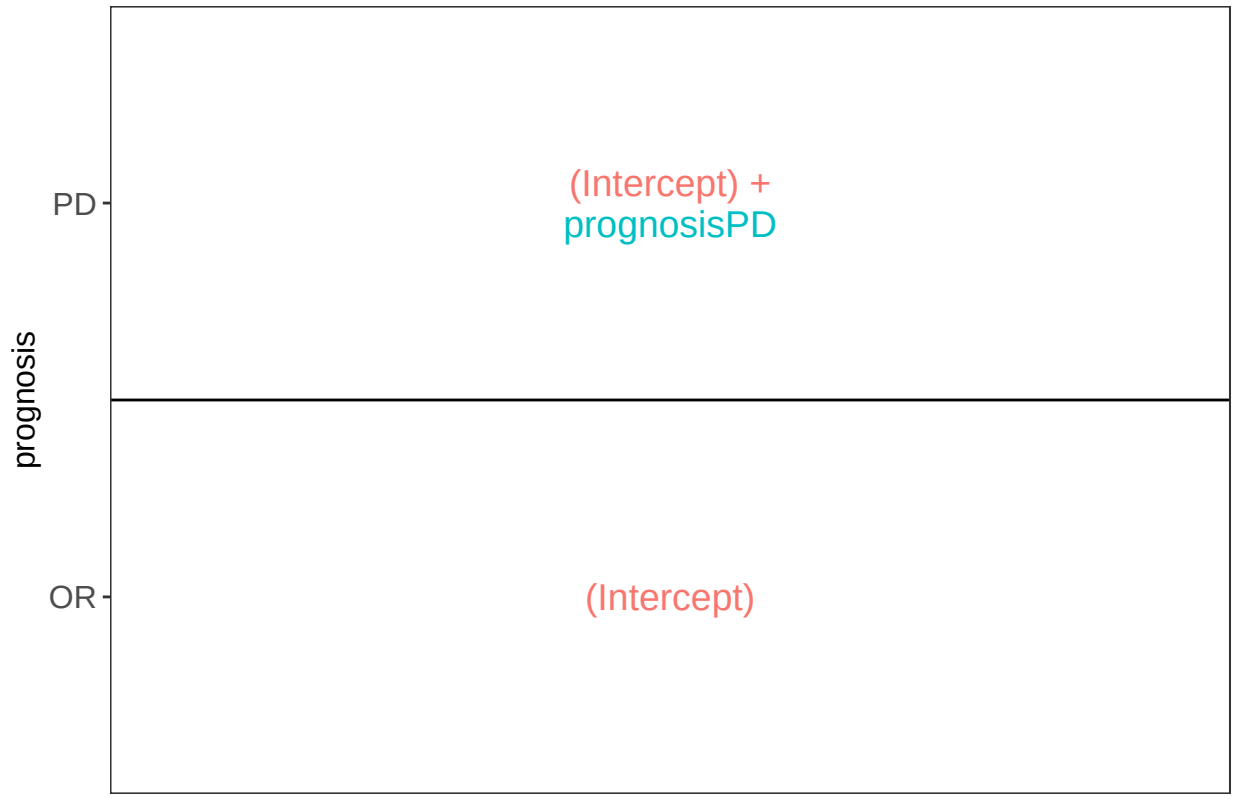
```
getCoef(rowData(pe[["protein"]])$msqrobModels[[1]])
```

```
## (Intercept) prognosisPD
##  -1.1067176   0.5721013
```

We can also explore the design of the model that we specified using the the package `ExploreModelMatrix`

```
library(ExploreModelMatrix)
VisualizeDesign(colData(pe),~prognosis)$plotlist
```

```
## [[1]]
```



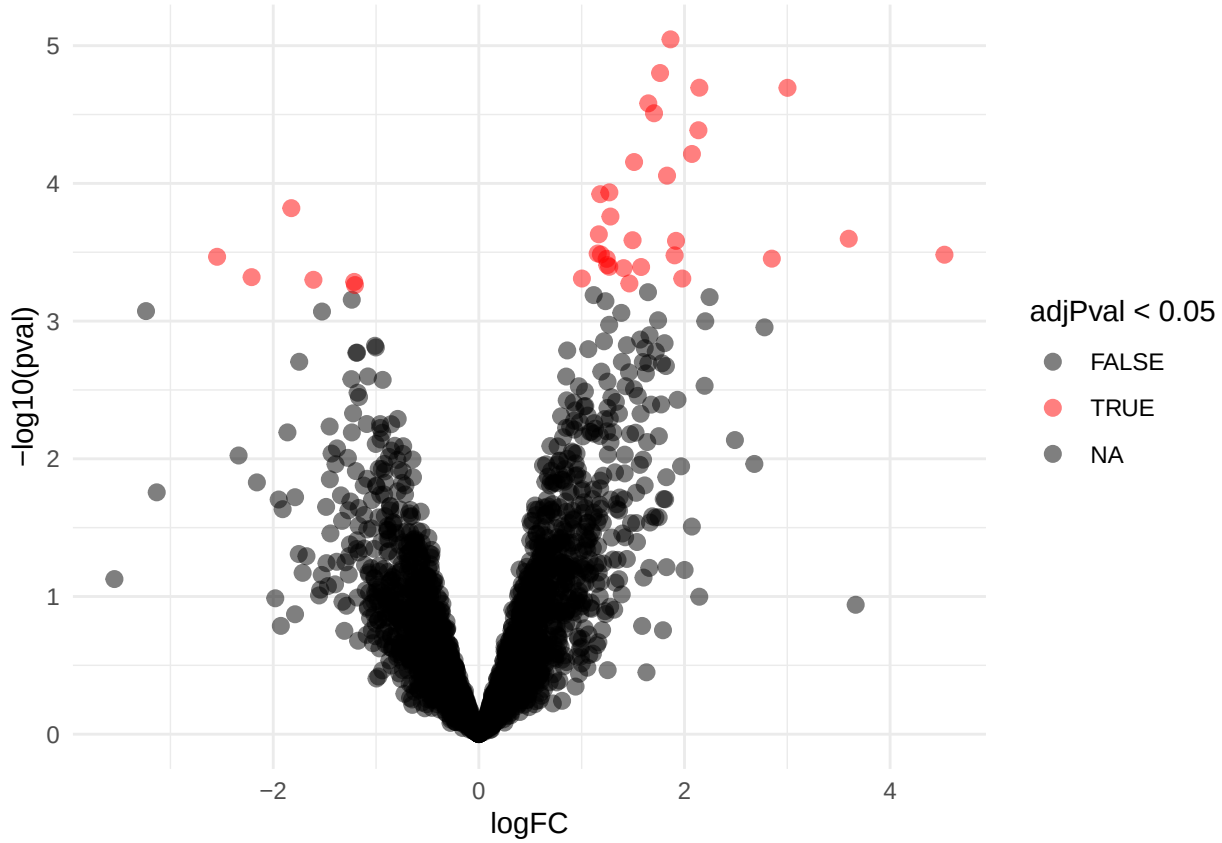
Spike-in condition `A` is the reference class. So the mean log2 expression for samples from good prognosis (OR) is '(Intercept). The mean log2 expression for samples from poor prognosis (PD) is'(Intercept)+prognosisPD'. Hence, the average log2 fold change between prognosis PD and prognosis OR is modelled using the parameter 'conditionPD'. Thus, we assess the contrast 'conditionPD = 0' with our statistical test.

```
L <- makeContrast("prognosisPD=0", parameterNames = c("prognosisPD"))
pe <- hypothesisTest(object = pe, i = "protein", contrast = L)
```

4.3 Plots

4.3.1 Volcano-plot

```
volcano <- ggplot(rowData(pe[["protein"]])$prognosisPD,
                  aes(x = logFC, y = -log10(pval), color = adjPval < 0.05)) +
  geom_point(cex = 2.5) +
  scale_color_manual(values = alpha(c("black", "red"), 0.5)) + theme_minimal()
volcano
```

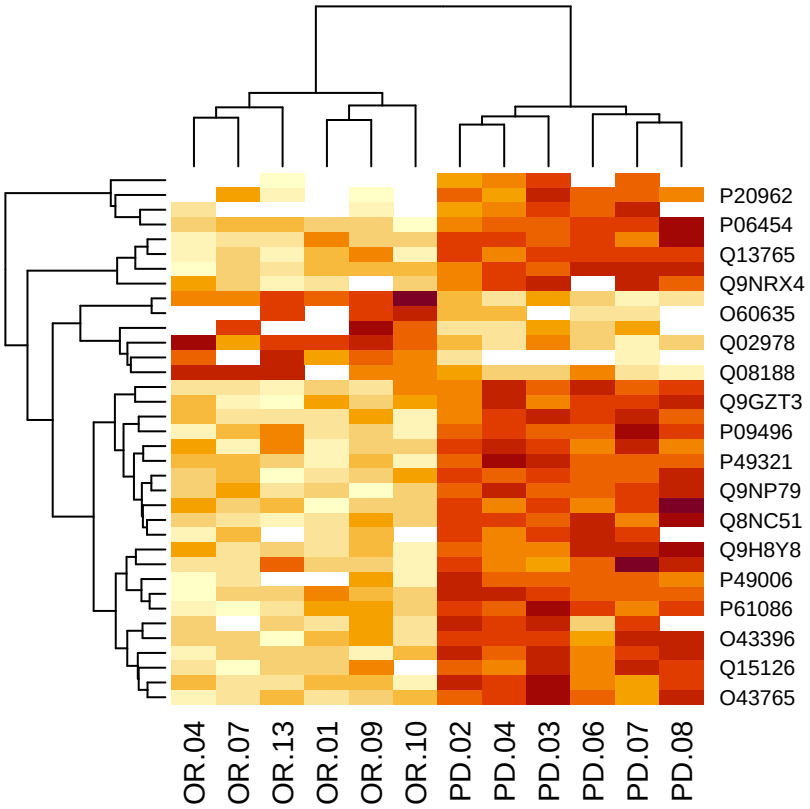


Note, that 36 proteins are found to be differentially abundant.

4.3.2 Heatmap

Note, that we also order the sigNames according to statistical significance.

```
sigNames <- rowData(pe[["protein"]])$prognosisPD %>%
  rownames_to_column("protein") %>%
  arrange(pval) %>%
  filter(adjPval<0.05) %>%
  pull(protein)
heatmap(assay(pe[["protein"]])[sigNames, ])
```

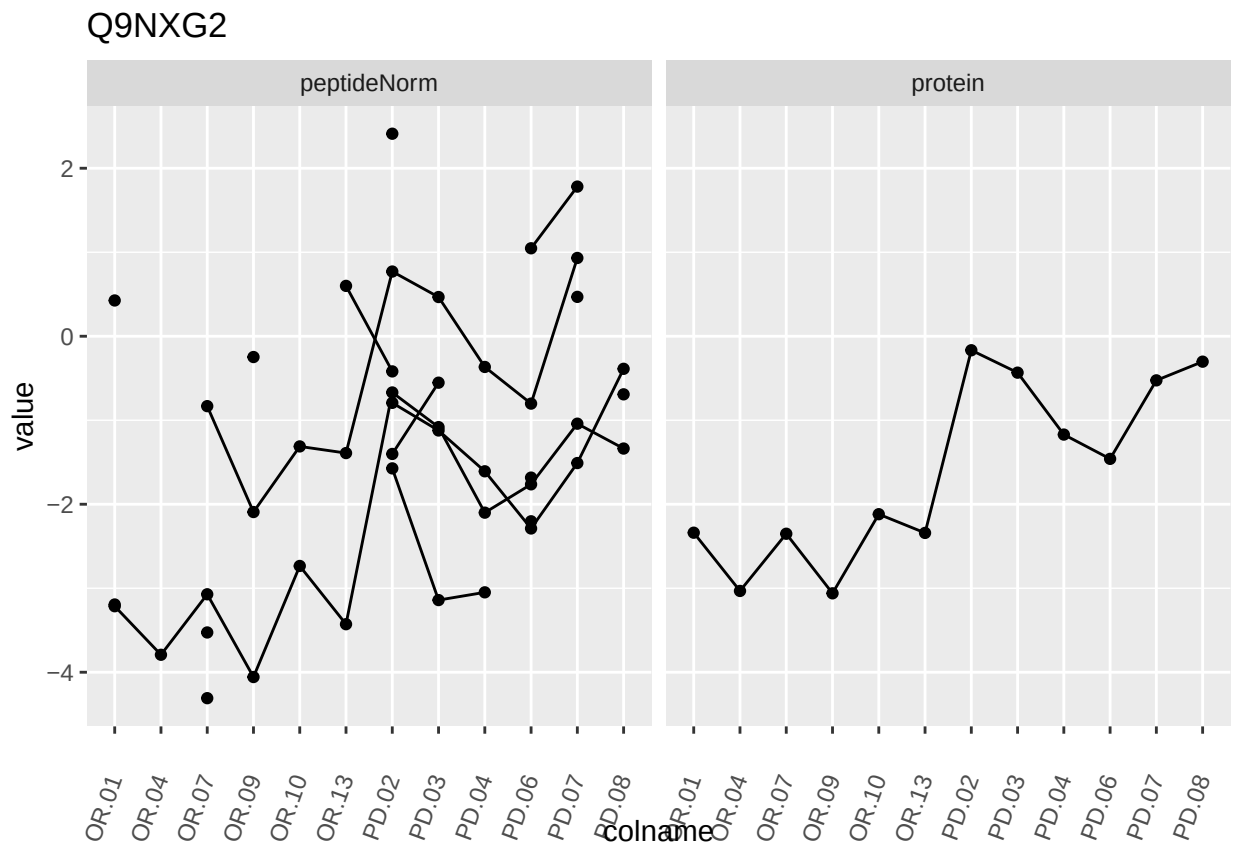


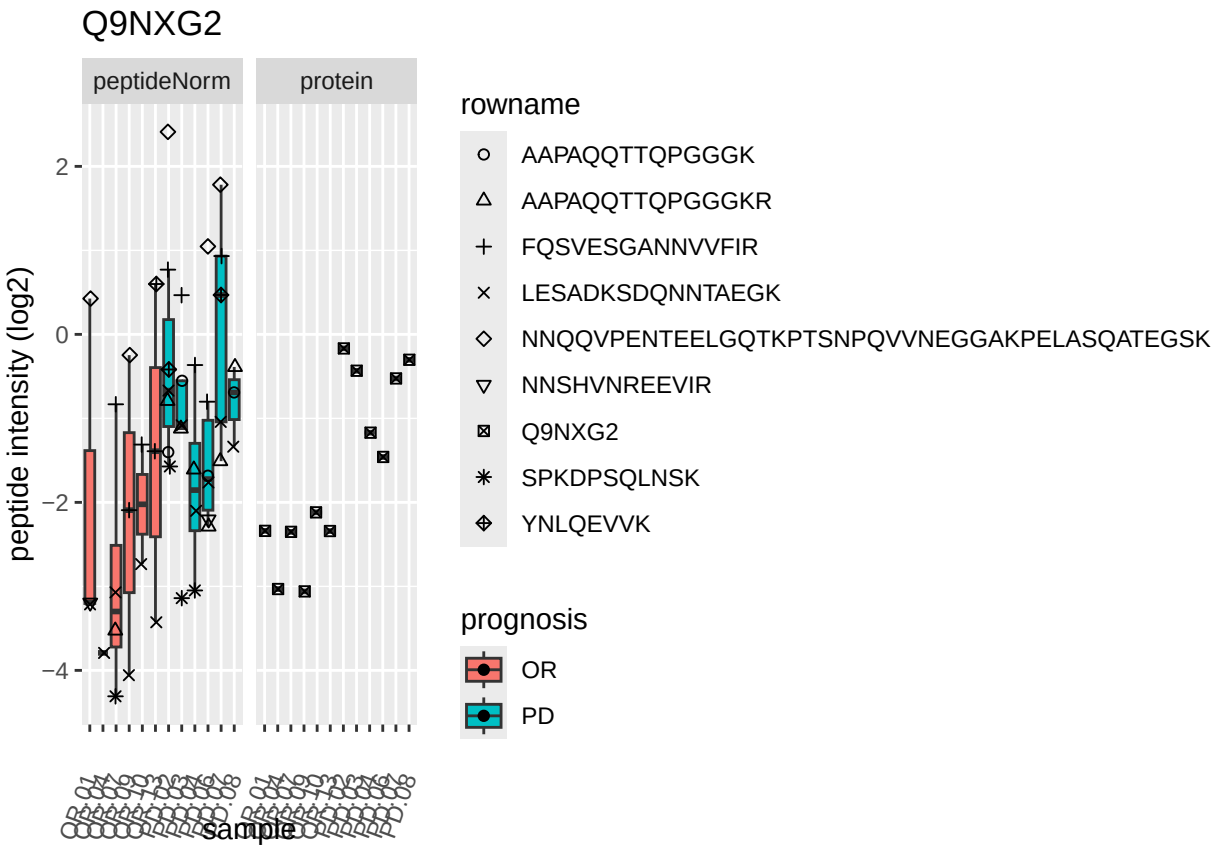
4.3.3 Detail plots

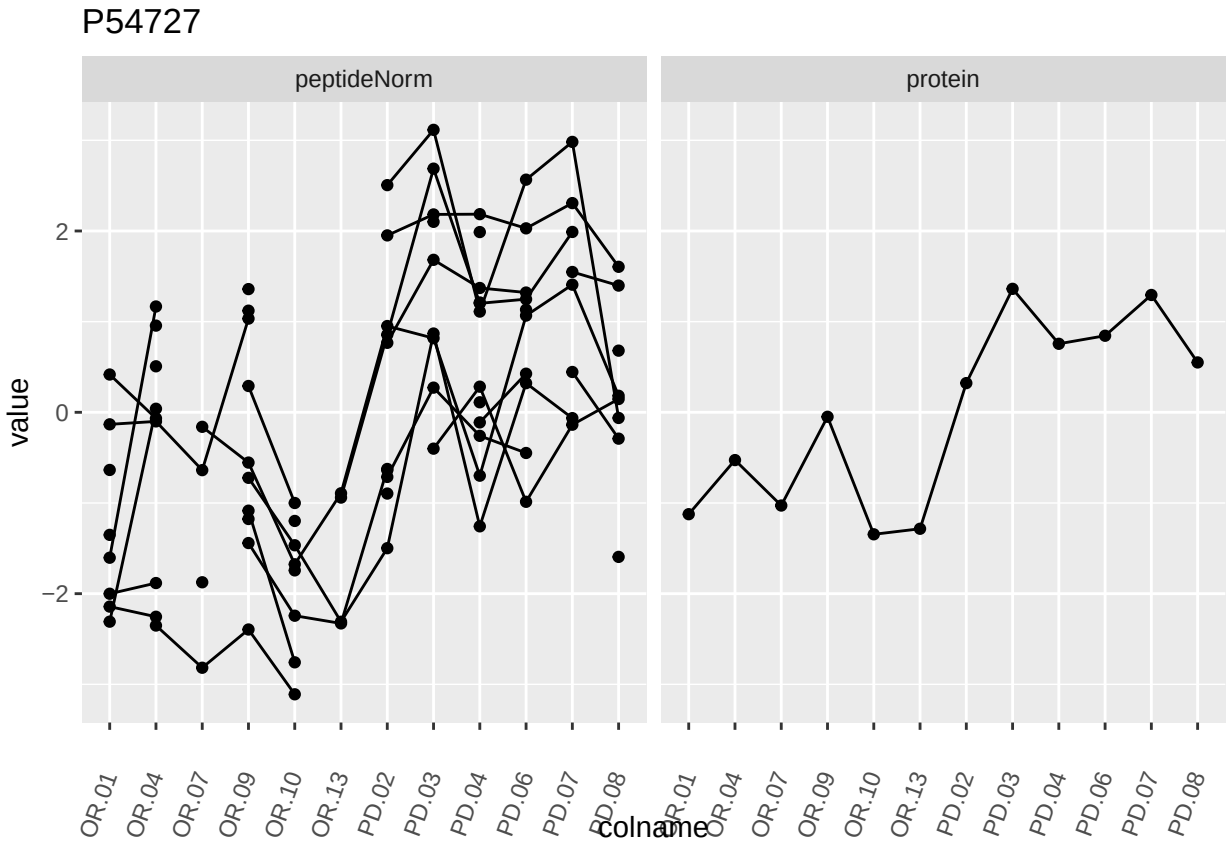
We make detail plots for the top 10 proteins to restrict the number of detail plots.

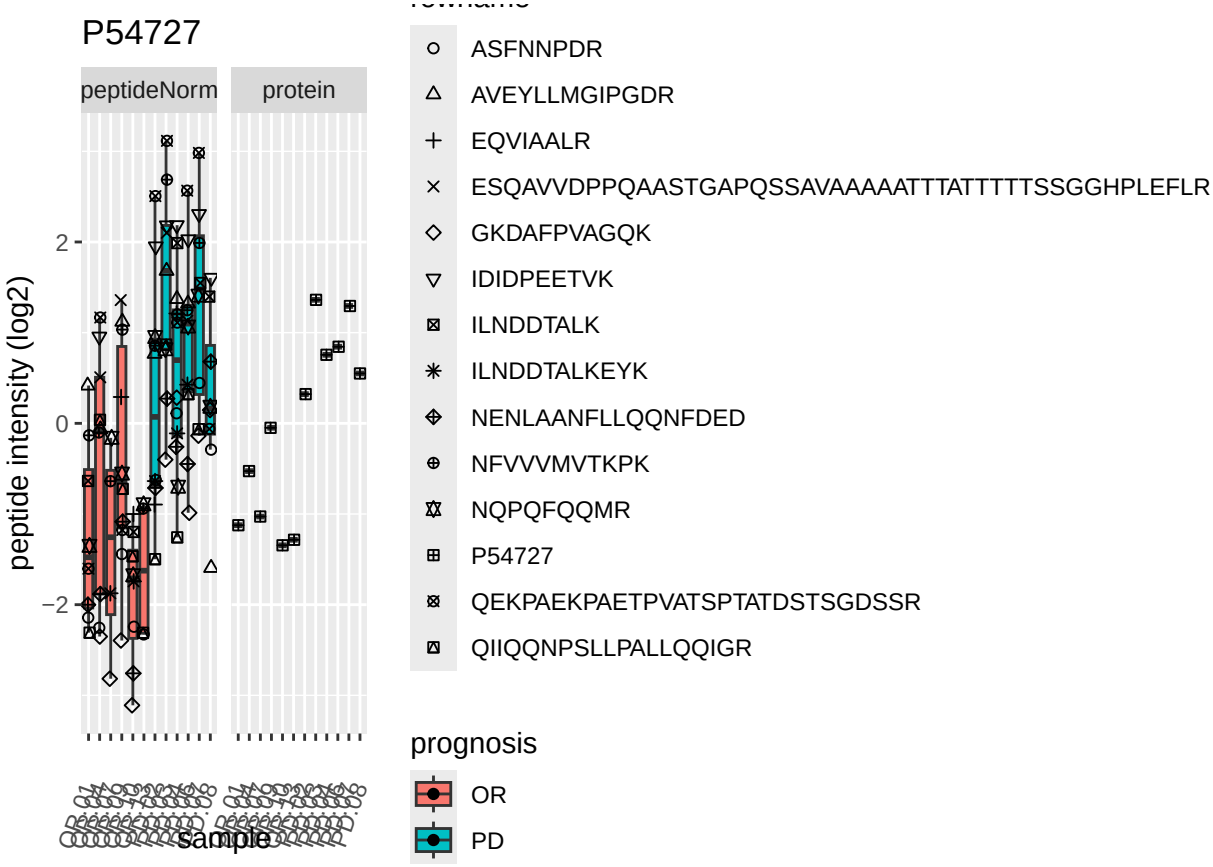
```
for (protName in sigNames)
#for (protName in orderProt[1:10])
{
pePlot <- pe[protName, , c("peptideNorm","protein")]
pePlotDf <- data.frame(longFormat(pePlot))
pePlotDf$assay <- factor(pePlotDf$assay,
                         levels = c("peptideNorm", "protein"))
pePlotDf$prognosis <- as.factor(colData(pePlot)[pePlotDf$colname, "prognosis"])

# plotting
p1 <- ggplot(data = pePlotDf,
       aes(x = colname, y = value, group = rowname)) +
    geom_line() +
    geom_point() +
    theme(axis.text.x = element_text(angle = 70, hjust = 1, vjust = 0.5)) +
    facet_grid(~assay) +
    ggtitle(protName)
print(p1)

# plotting 2
p2 <- ggplot(pePlotDf, aes(x = colname, y = value, fill = prognosis)) +
  geom_boxplot(outlier.shape = NA) +
```

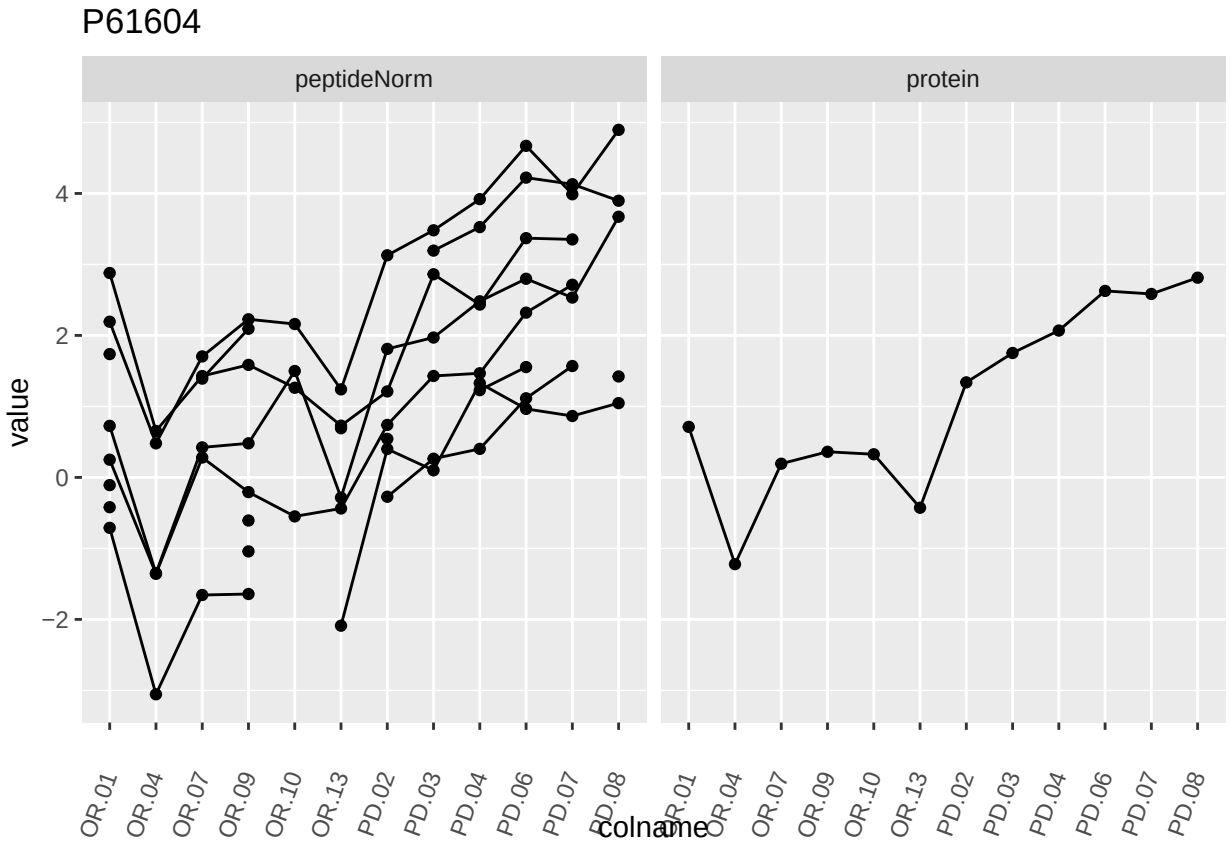
```
  geom_point(
    position = position_jitter(width = .1),
    aes(shape = rowname)) +
  scale_shape_manual(values = 1:nrow(pePlotDf)) +
  labs(title = protName, x = "sample", y = "peptide intensity (log2)") +
  theme(axis.text.x = element_text(angle = 70, hjust = 1, vjust = 0.5)) +
  facet_grid(~assay)
print(p2)
}
```

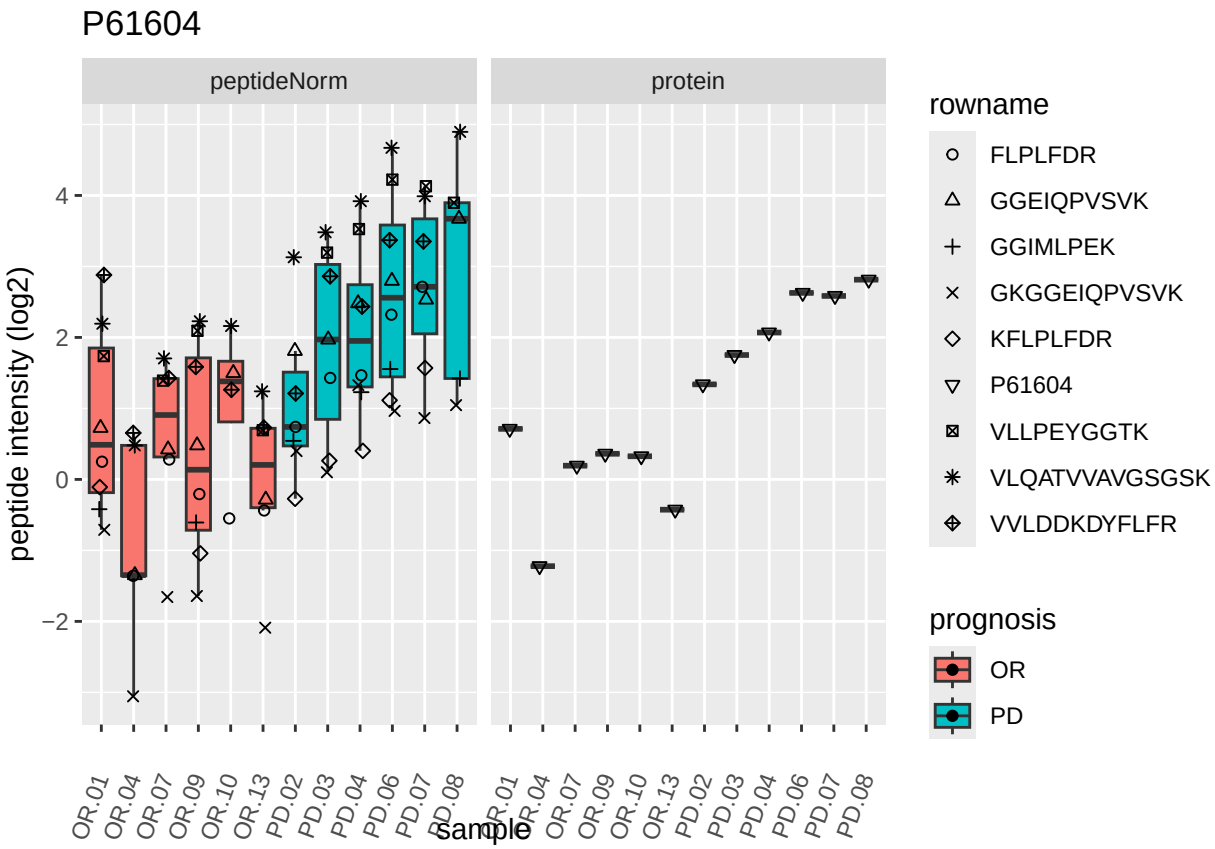


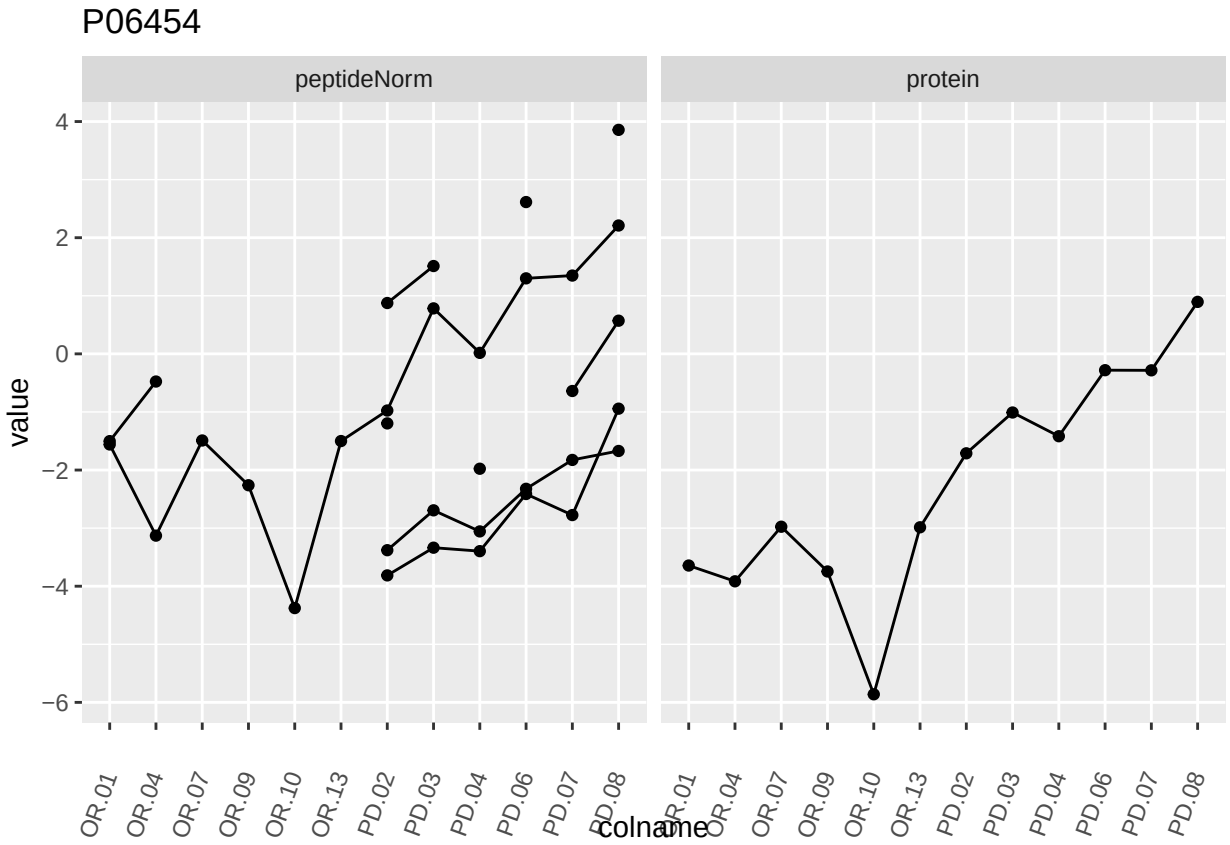

Q9NXG2
peptideNorm
protein
peptide intensity (log2)
2
0
−2
−4
sample
rowname
AAPAQQTTQPGGGK
AAPAQQTTQPGGGKR
FQSVESGANNVVFIR
LESADKSDQNNTAEGK
NNQQVPENTEELGQTKPTSNPQVVNEGGAKPELASQATEGSK
NNSHVNREEVIR
Q9NXG2
SPKDPSQLNSK
YNLQEVVK
prognosis
OR
PD

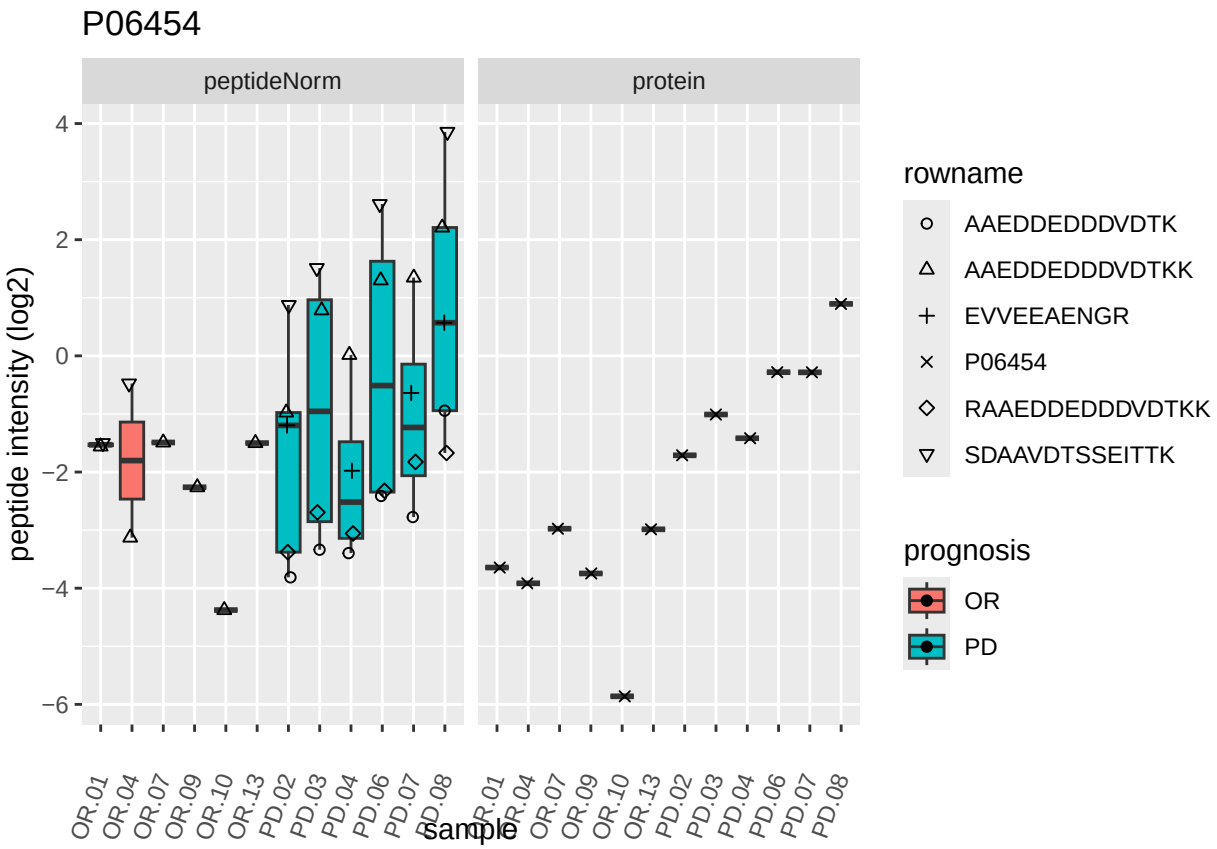

P54727
peptideNorm
protein
value
2
0
−2
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

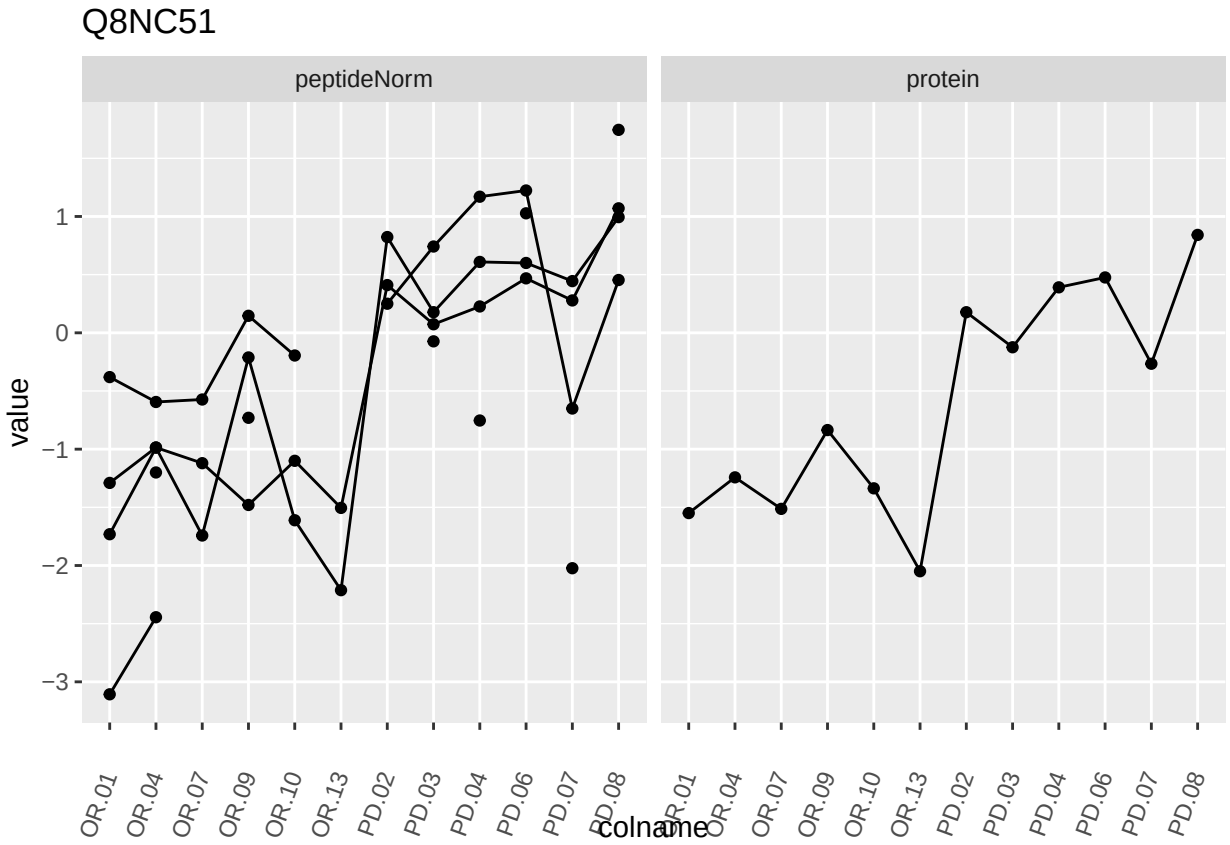

P54727
peptideNorm
protein
peptide intensity (log2)
2
0
−2
sample
ASFNNPDR
AVEYLLMGIPGDR
EQVIAALR
ESQAVVDPPQAASTGAPQSSAVAAAAATTTATTTTTSSGGHPLEFLR
GKDAFPVAGQK
IDIDPEETVK
ILNDDTALK
ILNDDTALKEYK
NENLAANFLLQQNFDED
NFVVVMVTKPK
NQPQFQQMR
P54727
QEKPAEKPAETPVATSPTATDSTSGDSSR
QIIQQNPSLLPALLQQIGR
prognosis
OR
PD

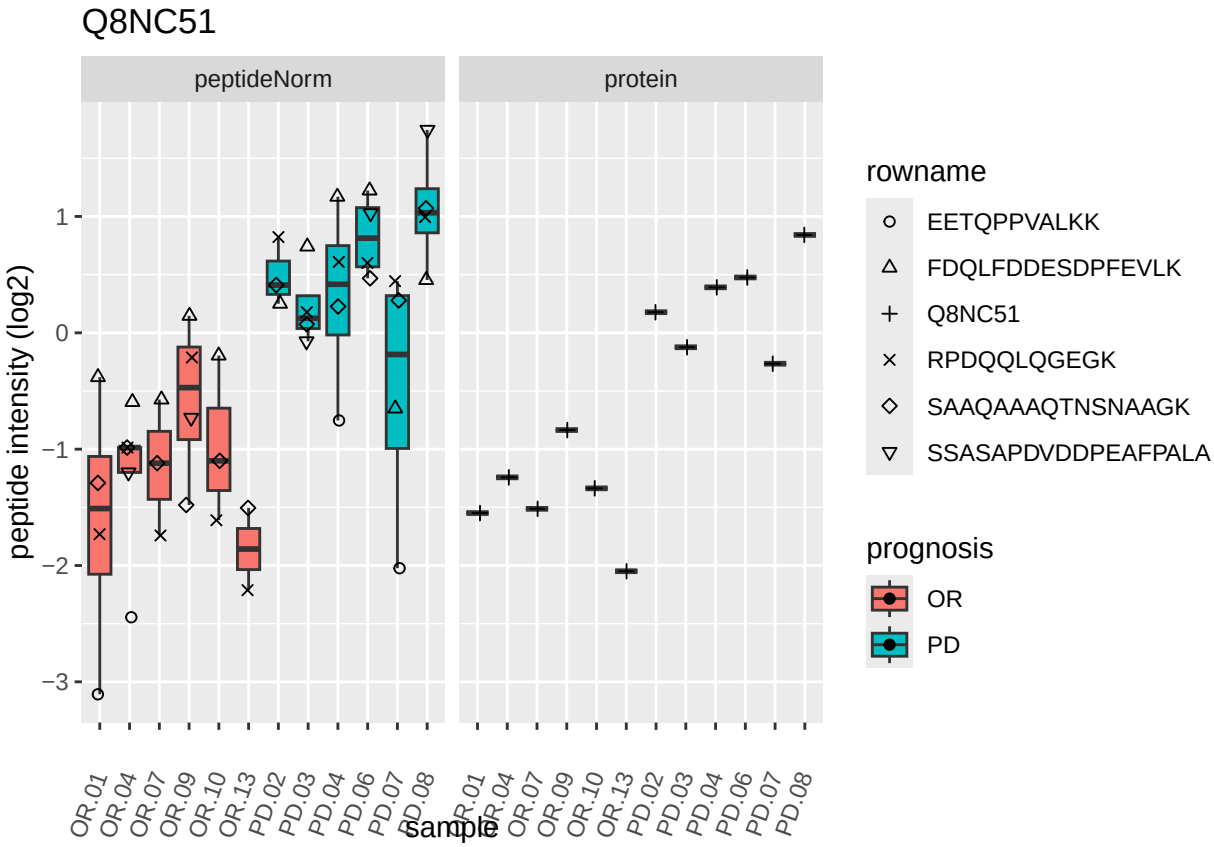

P61604
peptideNorm
protein
value
4
2
0
−2
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

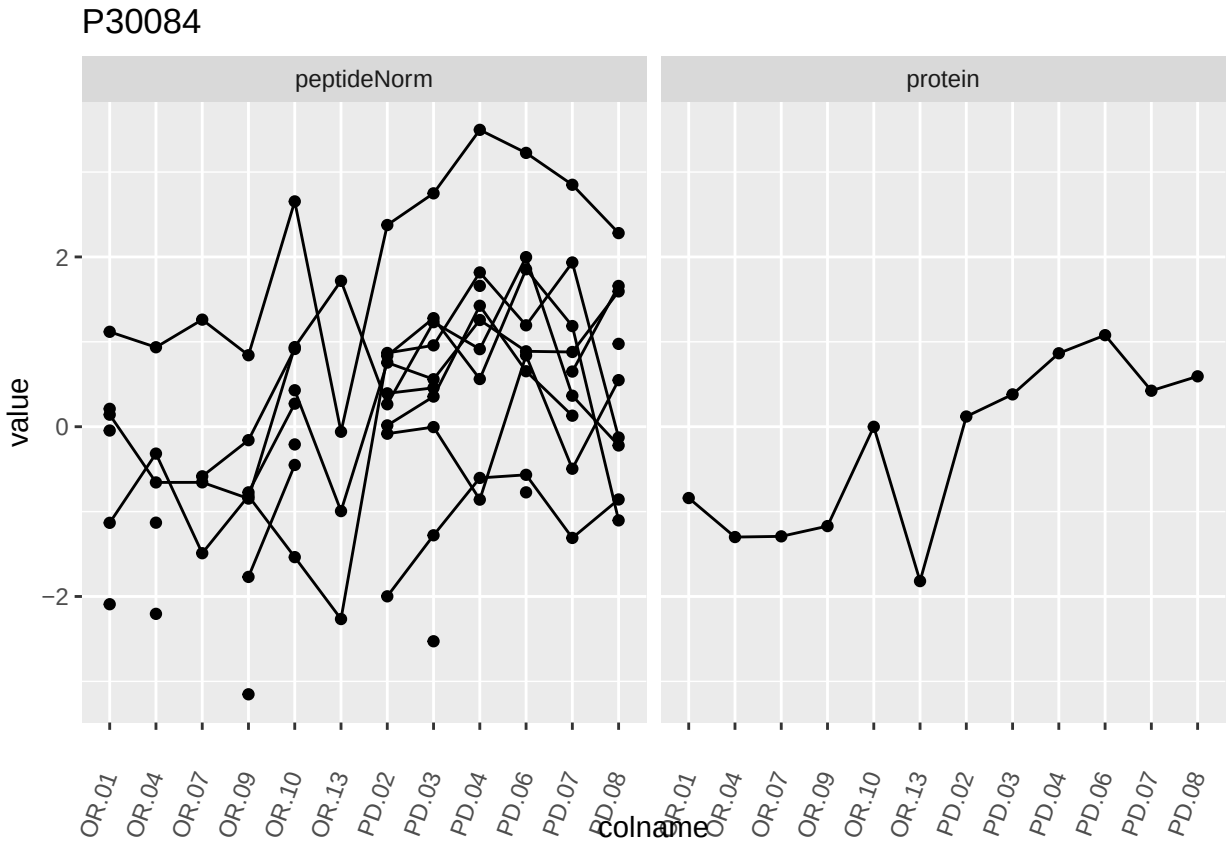

P61604
peptideNorm
protein
peptide intensity (log2)
4
2
0
−2
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
rowname
FLPLFDR
GGEIQPVSVK
GGIMLPEK
GKGGEIQPVSVK
KFLPLFDR
P61604
VLLPEYGGTK
VLQATVVAVGSGSK
VVLDDKDYFLFR
prognosis
OR
PD

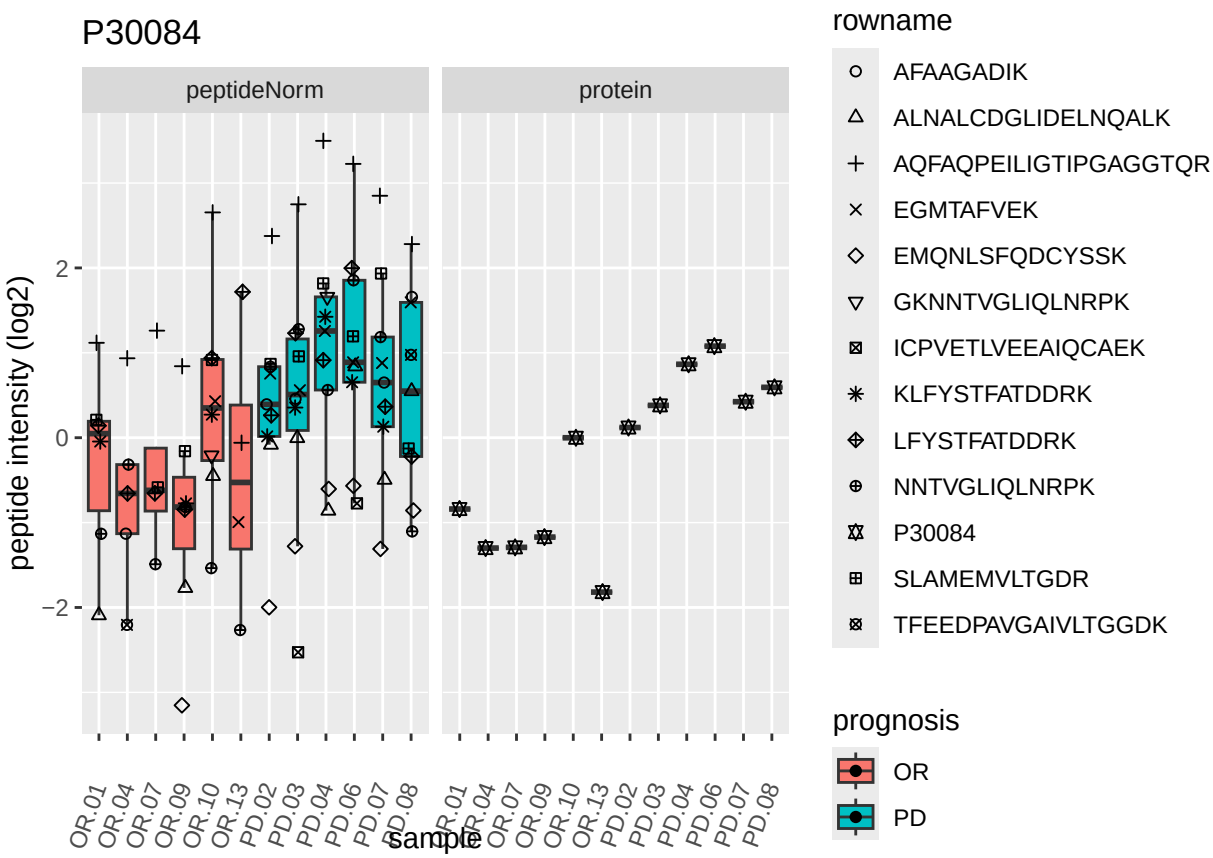

P06454
peptideNorm
protein
value
4
2
0
−2
−4
−6
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

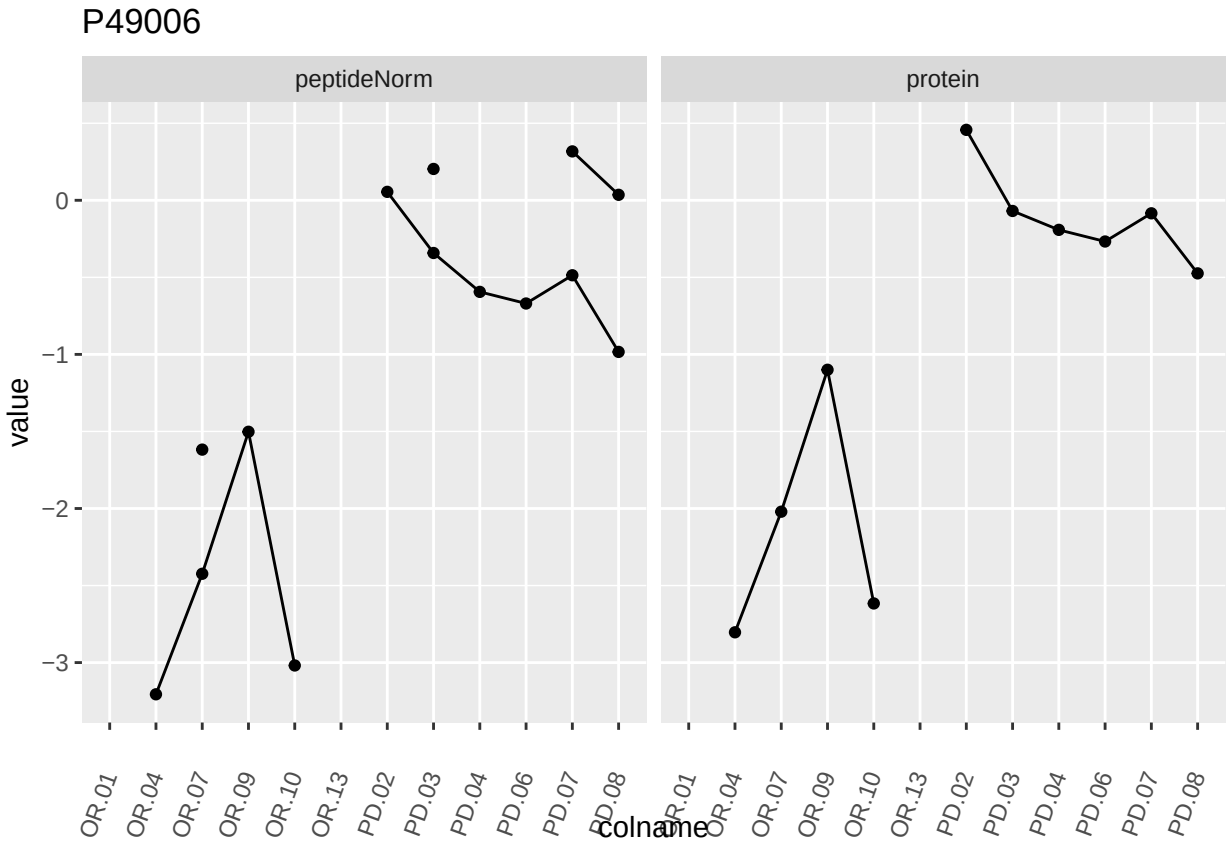

P06454
peptideNorm
protein
peptide intensity (log2)
4
2
0
−2
−4
−6
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
rowname
AAEDDEDDDVDTK
AAEDDEDDDVDTKK
EVVEEAENGR
P06454
RAAEDDEDDDVDTKK
SDAAVDTSSEITTK
prognosis
OR
PD

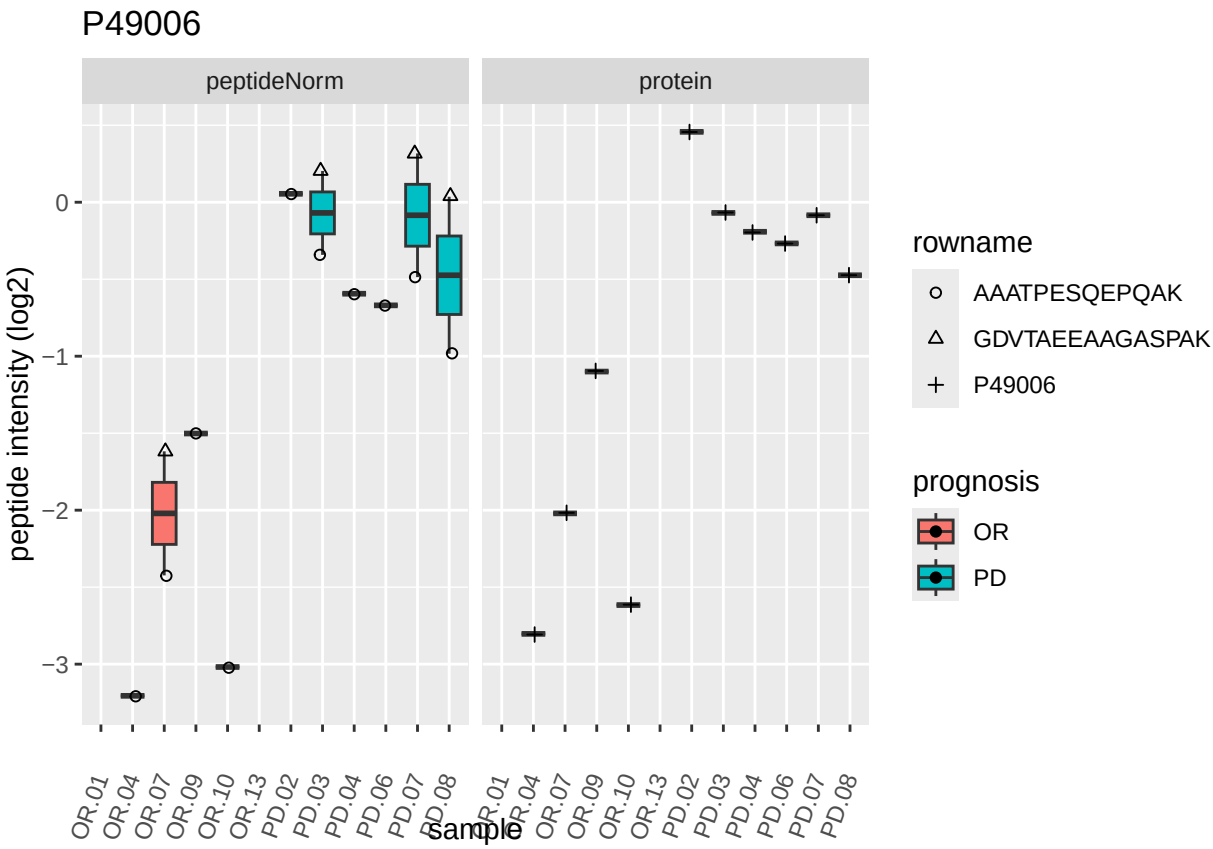

Q8NC51
peptideNorm
protein
value
1
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

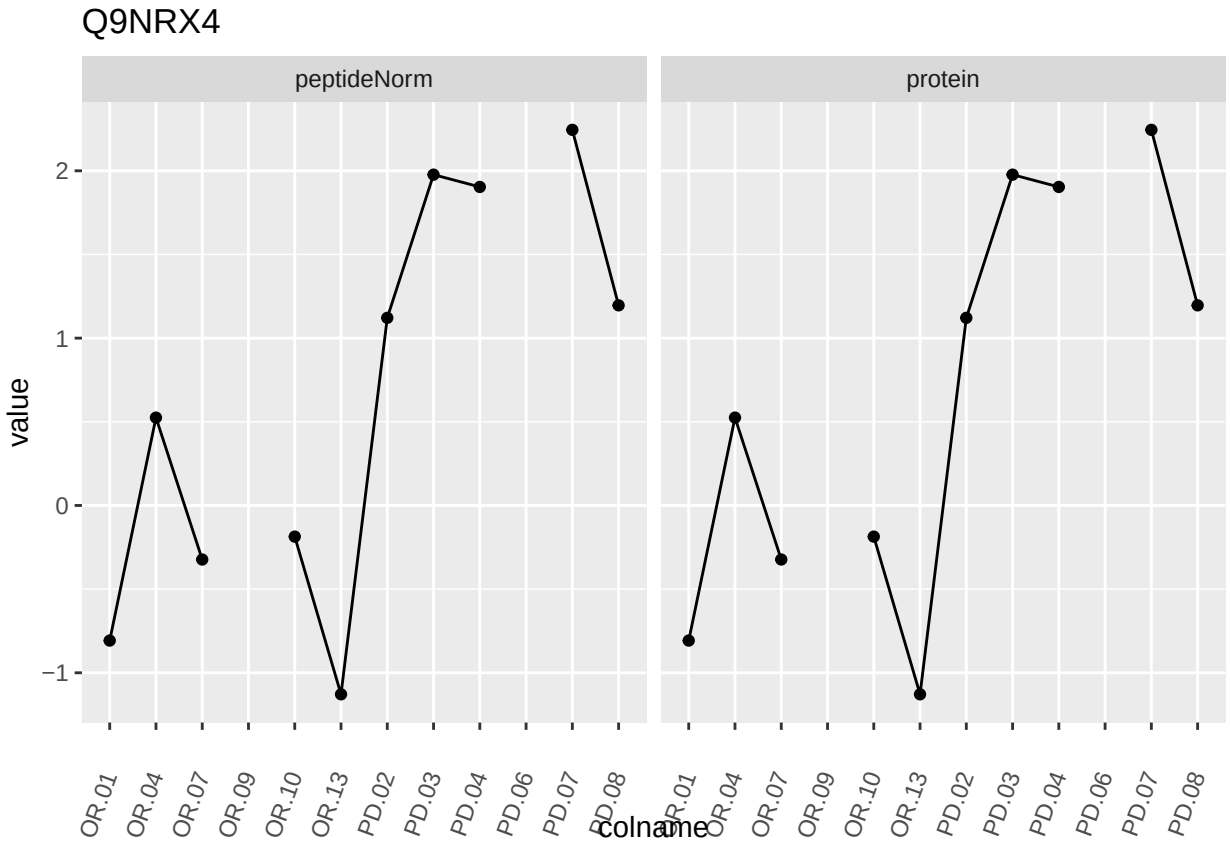

Q8NC51
peptideNorm
protein
peptide intensity (log2)
1
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
rowname
EETQPPVALKK
FDQLFDDESDPFEVLK
Q8NC51
RPDQQLQGEGK
SAAQAAAQTNSNAAGK
SSASAPDVDDPEAFPALA
prognosis
OR
PD

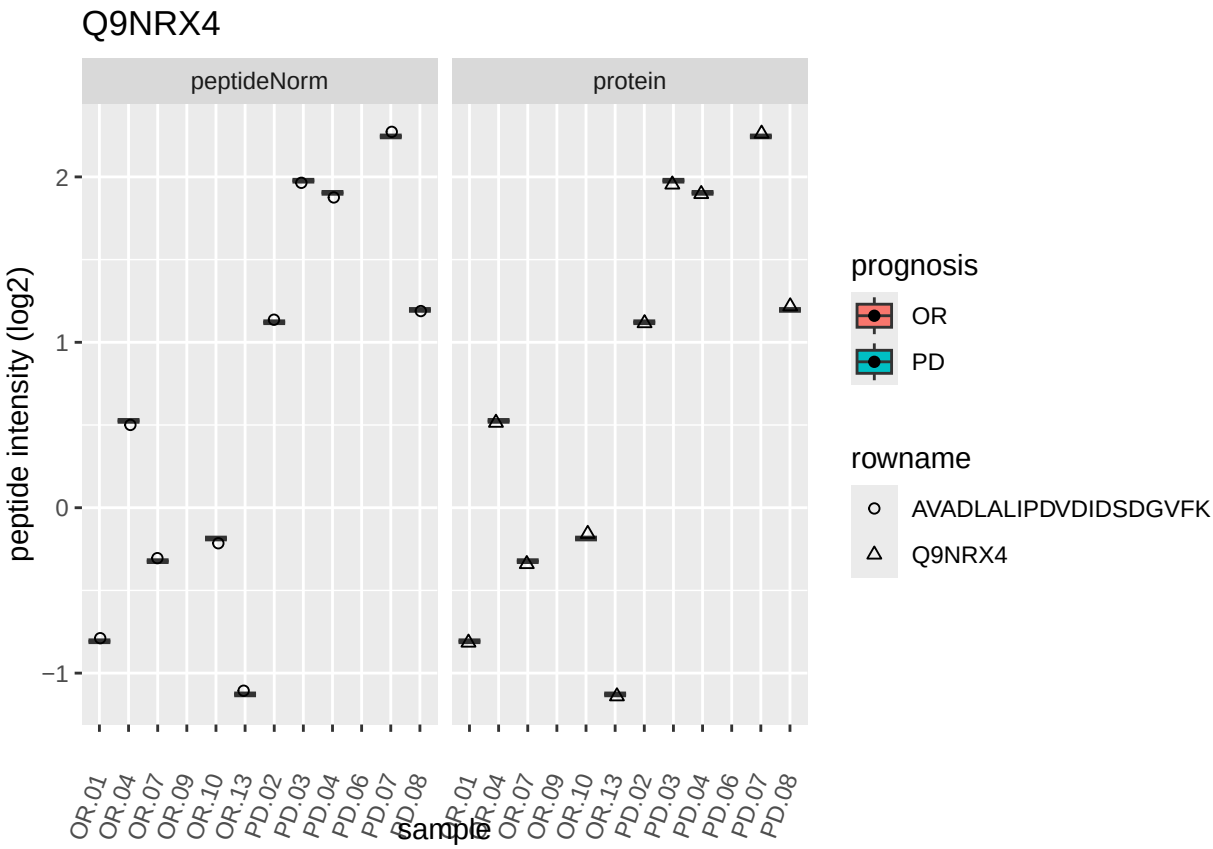

P30084
peptideNorm
protein
value
2
0
−2
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

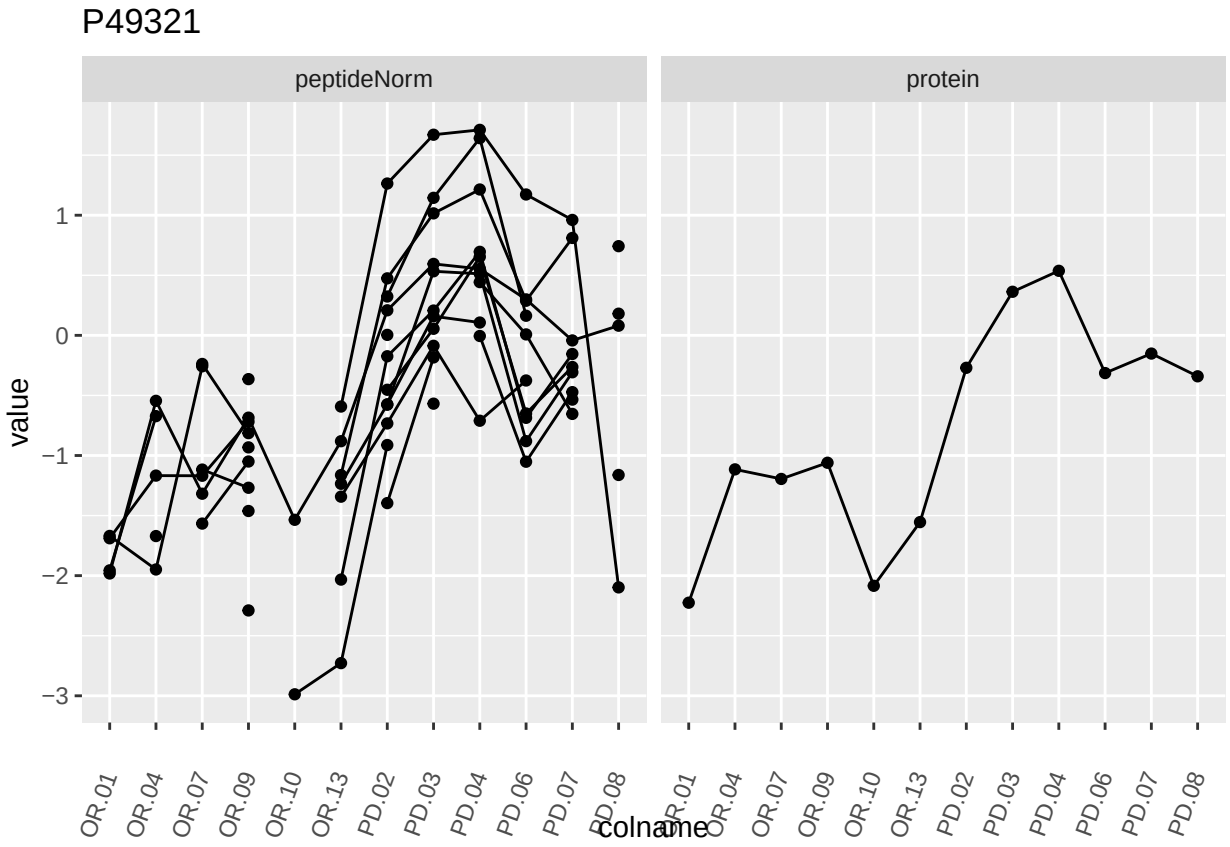

P30084
peptideNorm
protein
peptide intensity (log2)
2
0
−2
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
rowname
AFAAGADIK
ALNALCDGLIDELNQALK
AQFAQPEILIGTIPGAGGTQR
EGMTAFVEK
EMQNLSFQDCYSSK
GKNNTVGLIQLNRPK
ICPVETLVEEAIQCAEK
KLFYSTFATDDRK
LFYSTFATDDRK
NNTVGLIQLNRPK
P30084
SLAMEMVLTGDR
TFEEDPAVGAIVLTGGDK
prognosis
OR
PD

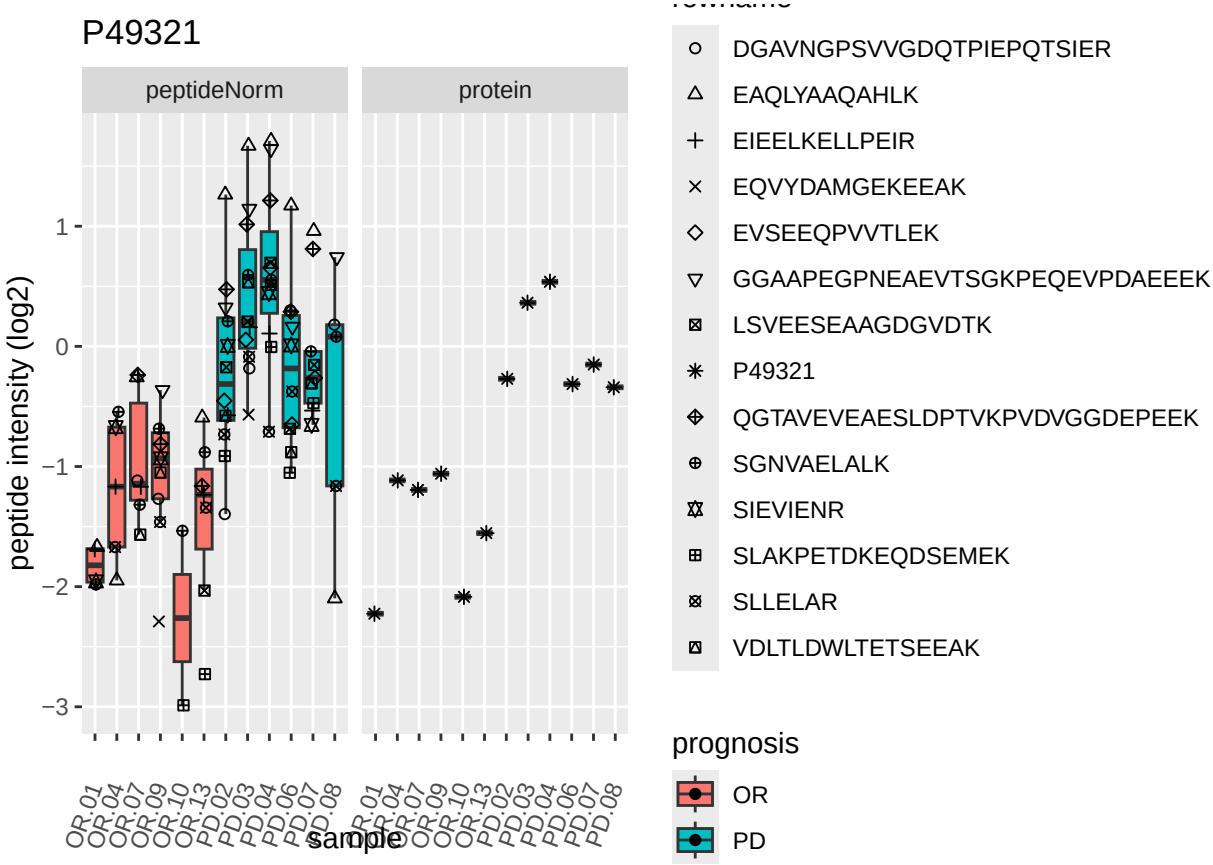

P49006
peptideNorm
protein
value
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

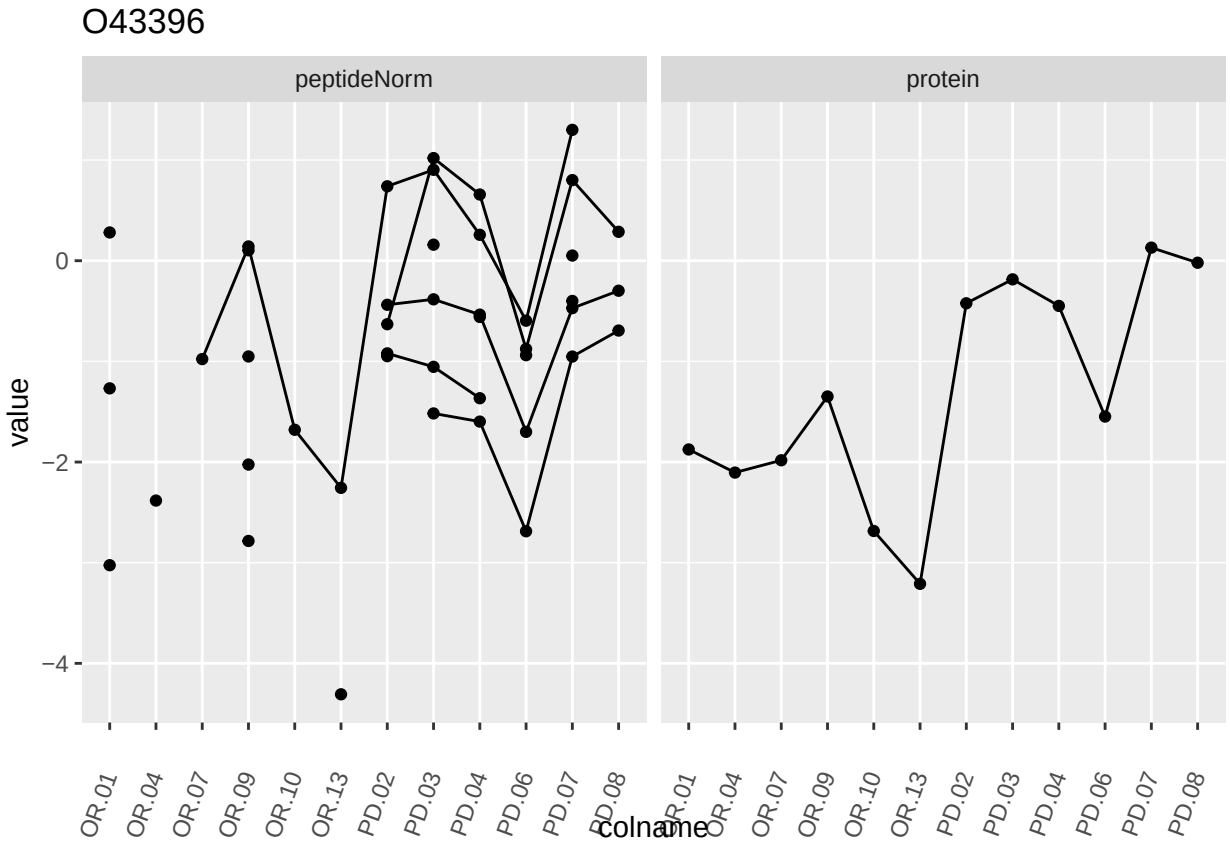

P49006
peptideNorm
protein
peptide intensity (log2)
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
rowname
AAATPESQEPQAK
GDVTAEEAAGASPAK
P49006
prognosis
OR
PD

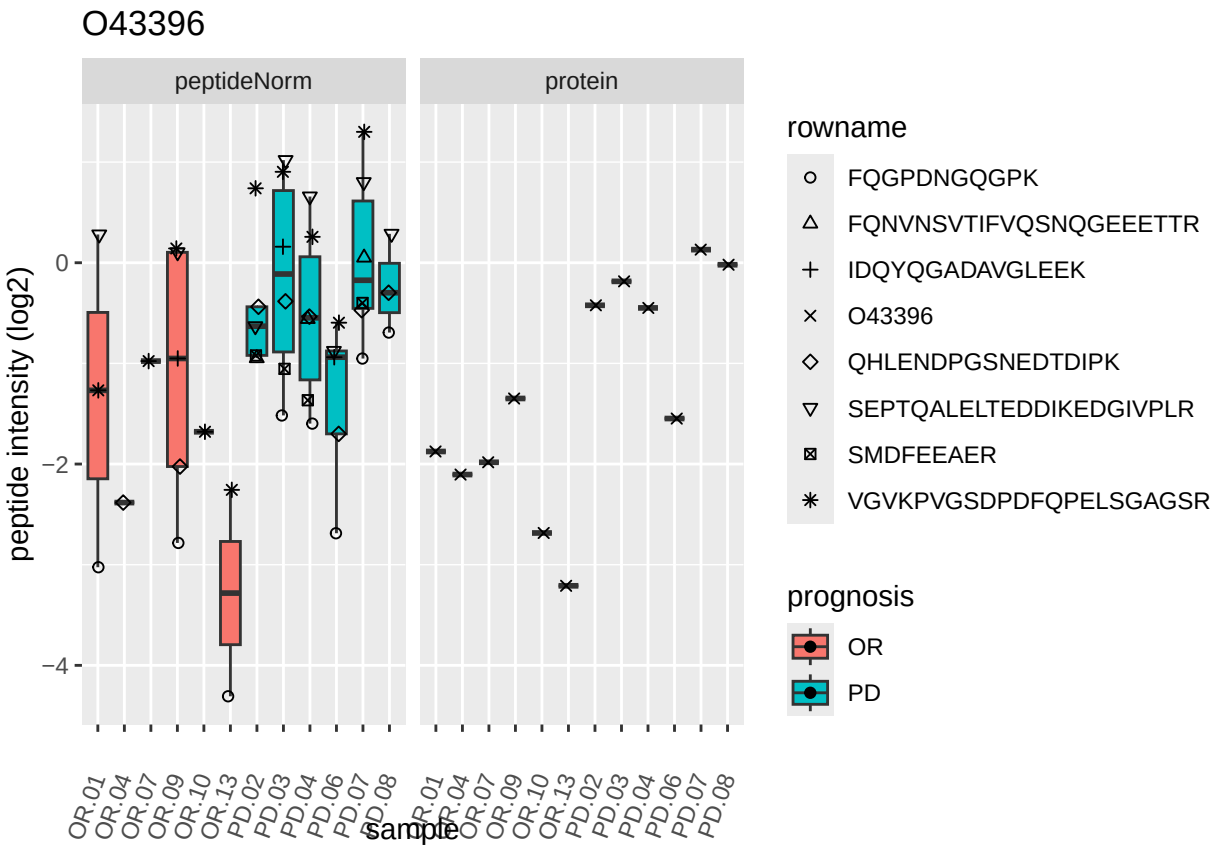

Q9NRX4
peptideNorm
protein
value
2
1
0
−1
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

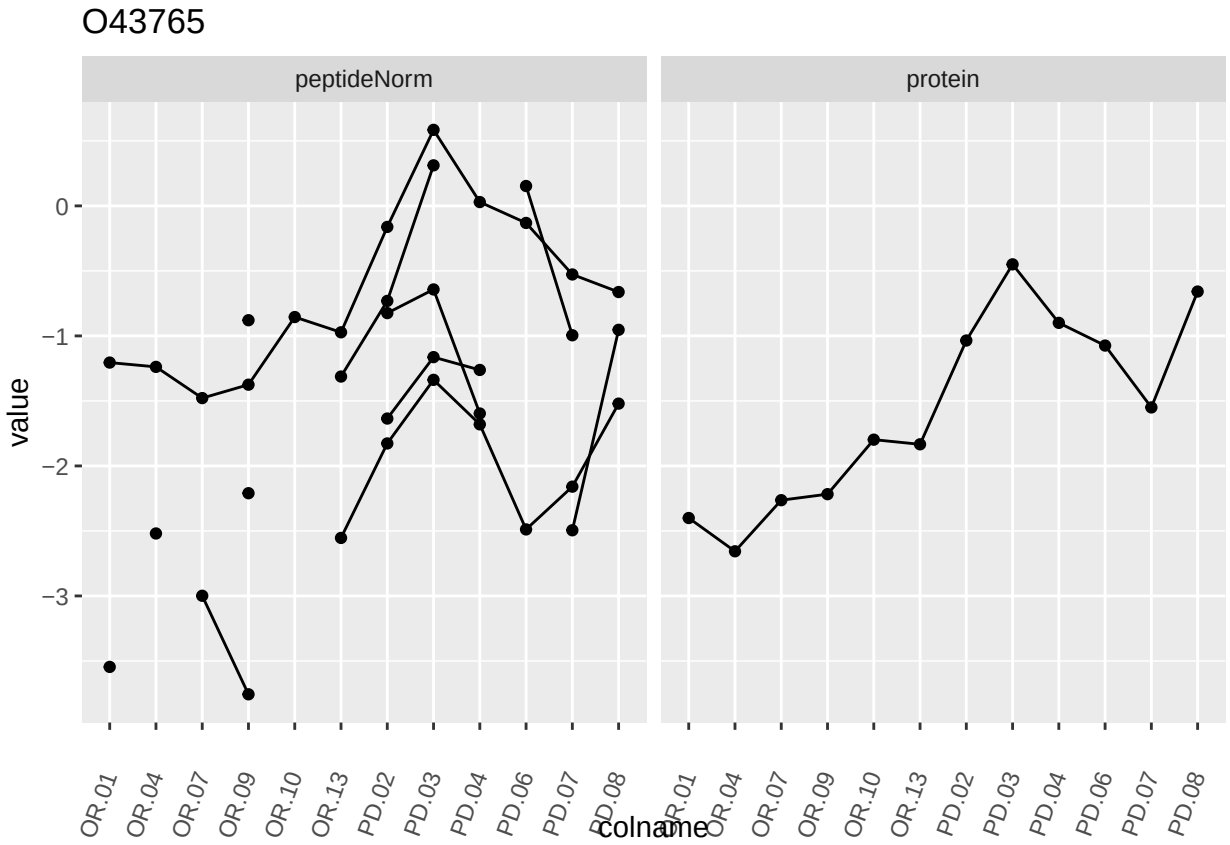

Q9NRX4
peptideNorm
protein
peptide intensity (log2)
2
1
0
−1
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
prognosis
OR
PD
rowname
AVADLALIPDVDIDSDGVFK
Q9NRX4

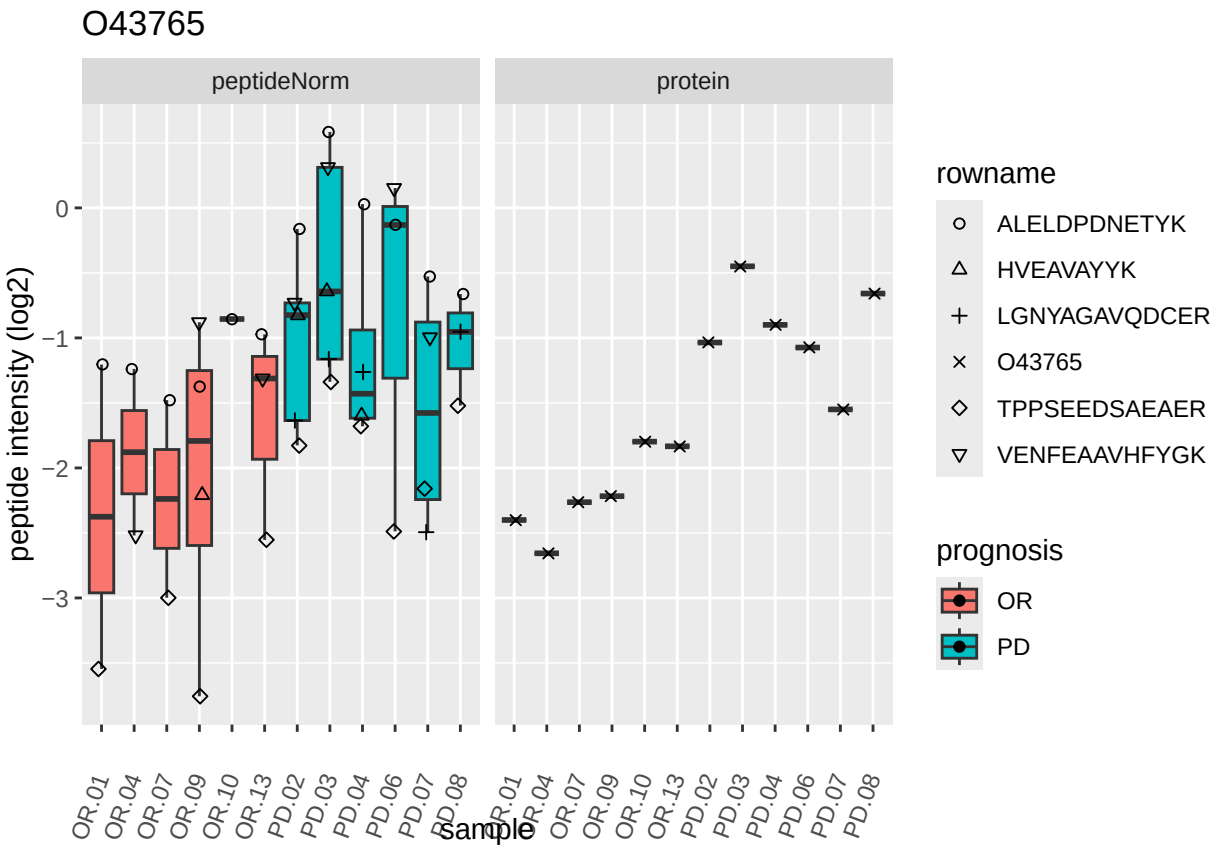

P49321
peptideNorm
protein
value
1
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

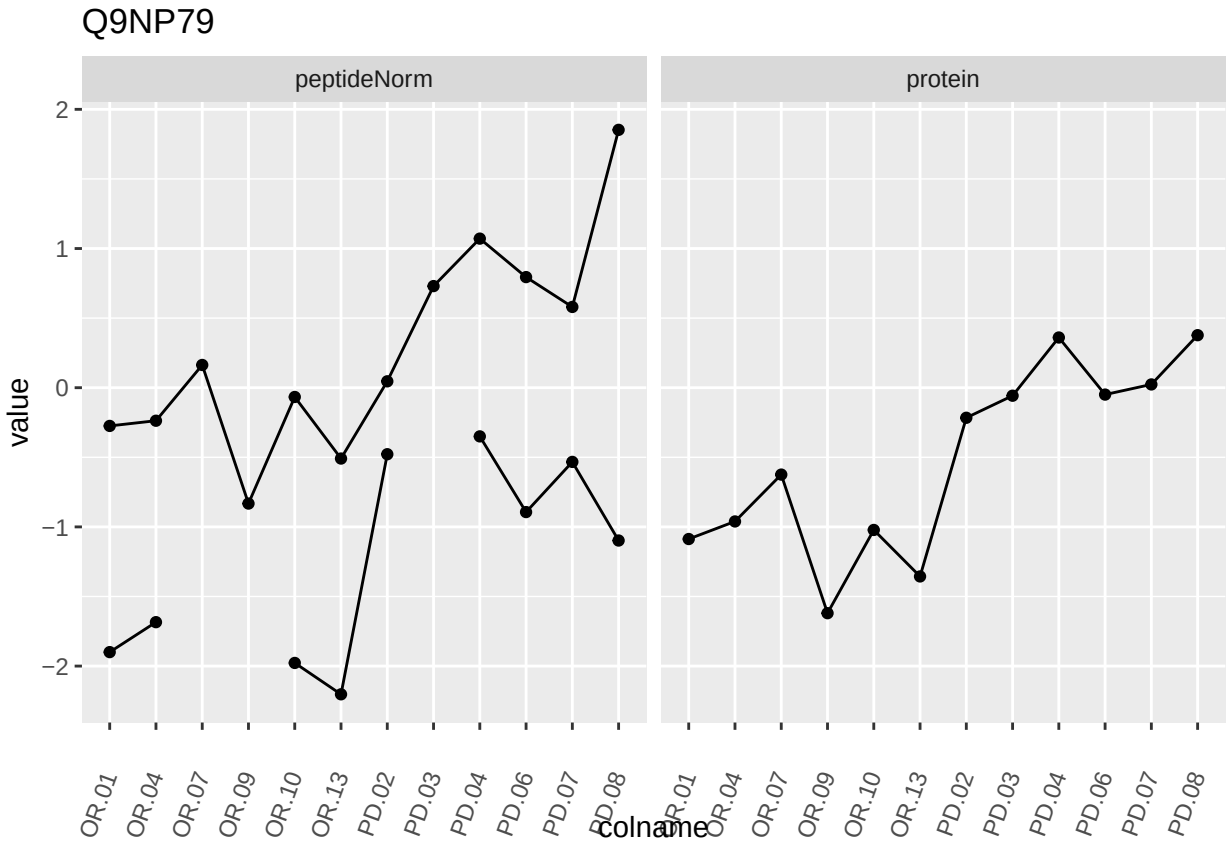

P49321
peptideNorm
protein
peptide intensity (log2)
1
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
DGAVNGPSVVGDQTPIEPQTSIER
EAQLYAAQAHLK
EIEELKELLPEIR
EQVYDAMGEKEEAK
EVSEEQPVVTLEK
GGAAPEGPNEAEVTSGKPEQEVPDAEEEK
LSVEESEAAGDGVDTK
P49321
QGTAVEVEAESLDPTVKPVDVGGDEPEEK
SGNVAELALK
SIEVIENR
SLAKPETDKEQDSEMEK
SLLELAR
VDLTLDWLTETSEEAK
prognosis
OR
PD

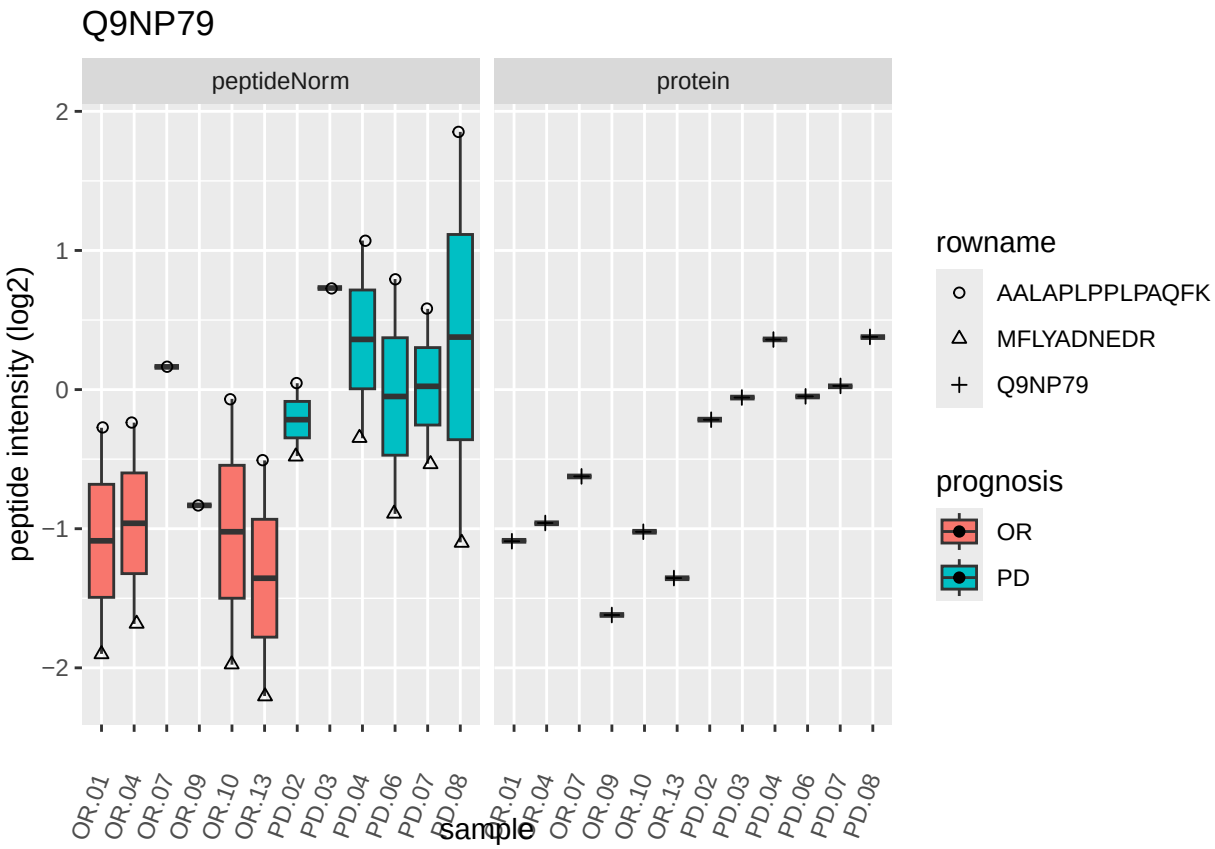

O43396
peptideNorm
protein
value
0
−2
−4
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

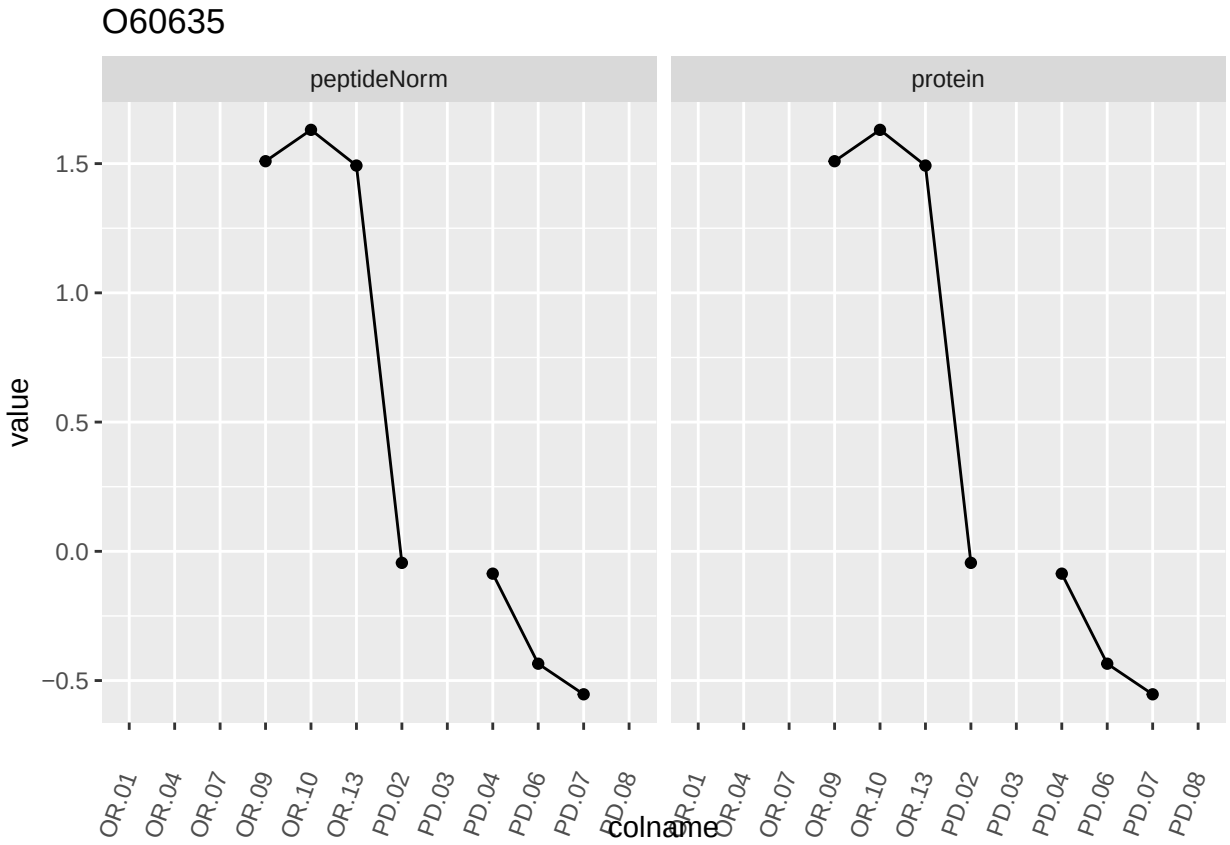

O43396
peptideNorm
protein
peptide intensity (log2)
0
−2
−4
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
rowname
FQGPDNGQGPK
FQNVNSVTIFVQSNQGEEETTR
IDQYQGADAVGLEEK
O43396
QHLENDPGSNEDTDIPK
SEPTQALELTEDDIKEDGIVPLR
SMDFEEAER
VGVKPVGSDPDFQPELSGAGSR
prognosis
OR
PD

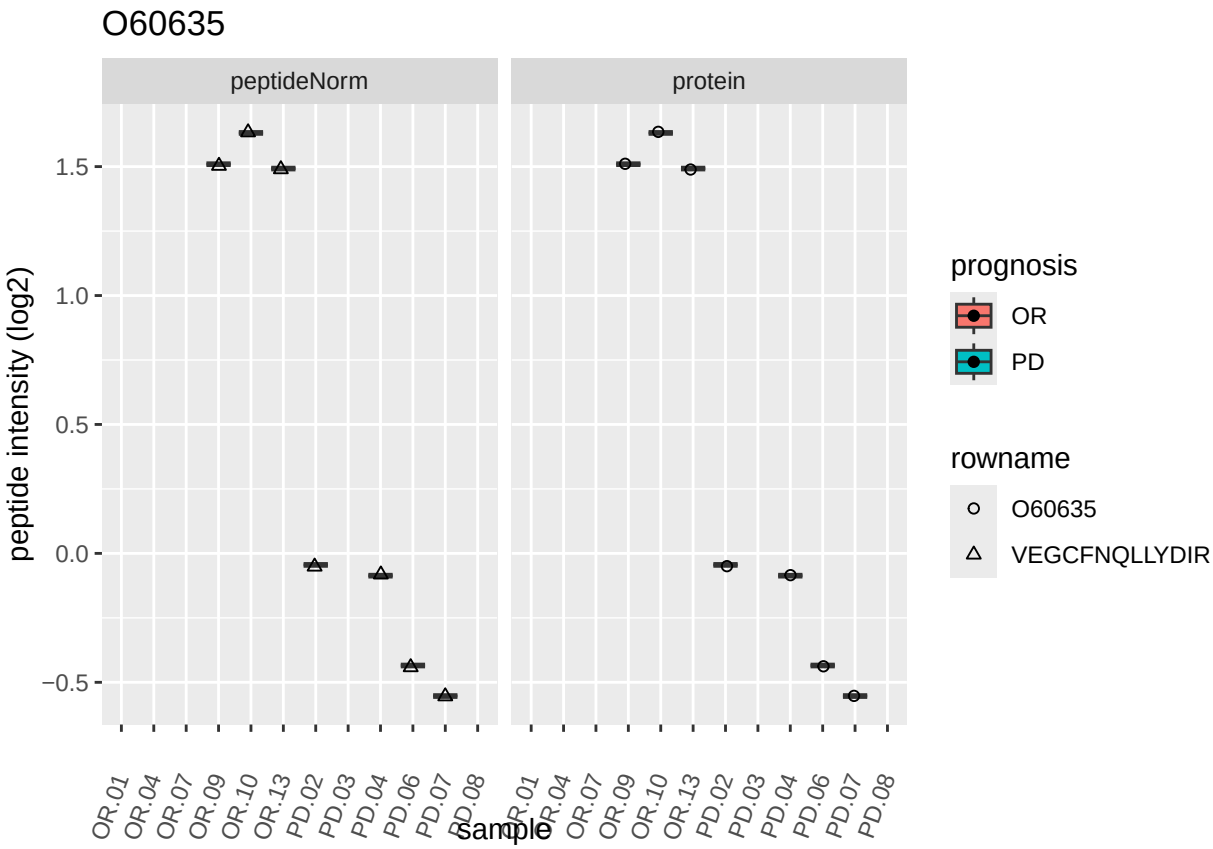

O43765
peptideNorm
protein
value
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

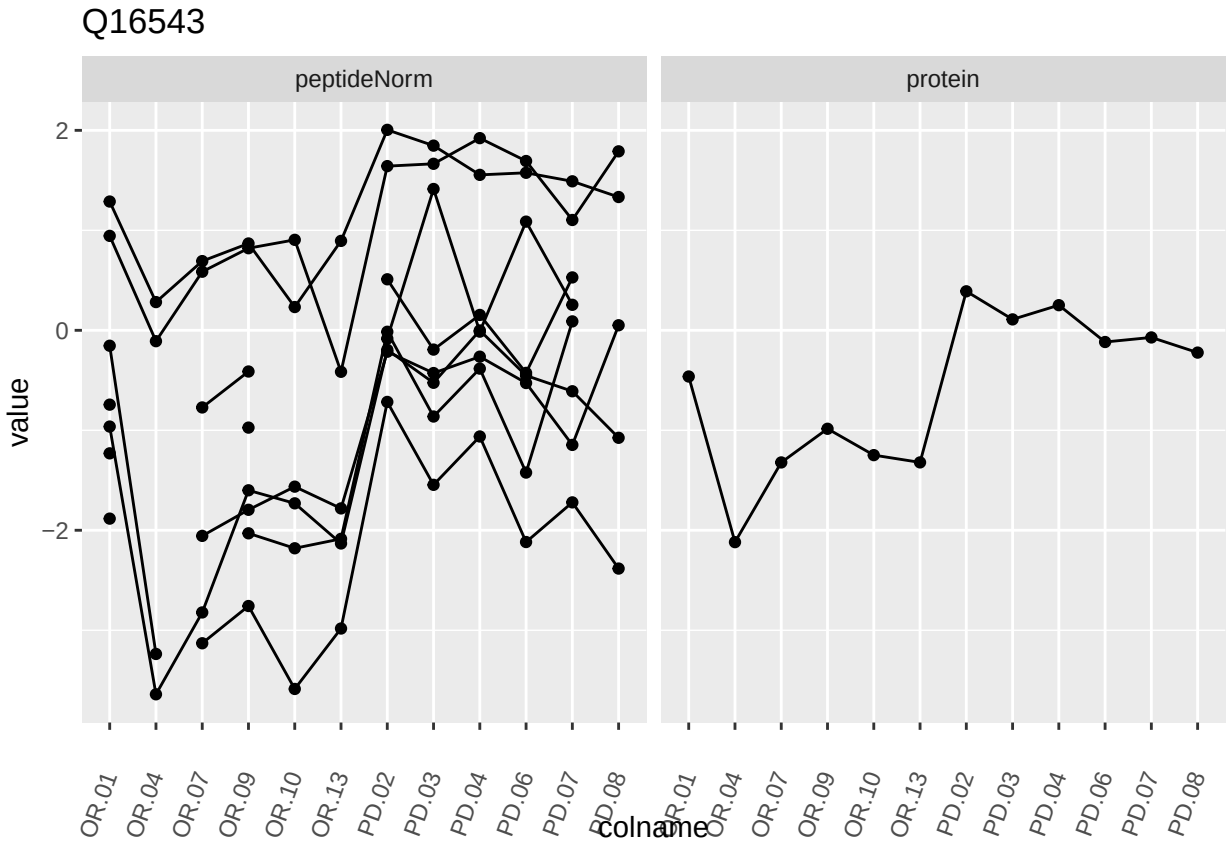

O43765
peptideNorm
protein
peptide intensity (log2)
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
Sample
rowname
ALELDPDNETYK
HVEAVAYYK
LGNYAGAVQDCER
O43765
TPPSEEDSAEAER
VENFEAAVHFYGK
prognosis
OR
PD

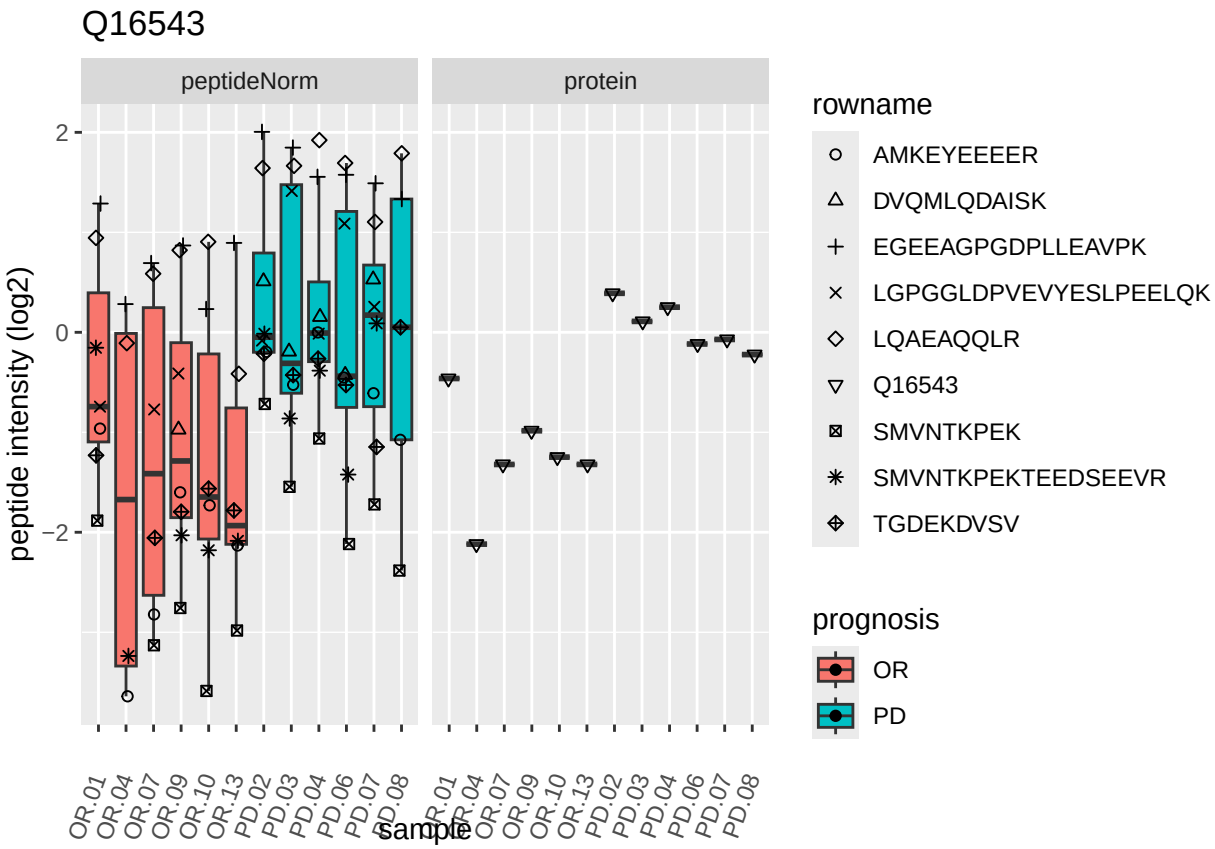

Q9NP79
peptideNorm
protein
value
2
1
0
−1
−2
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

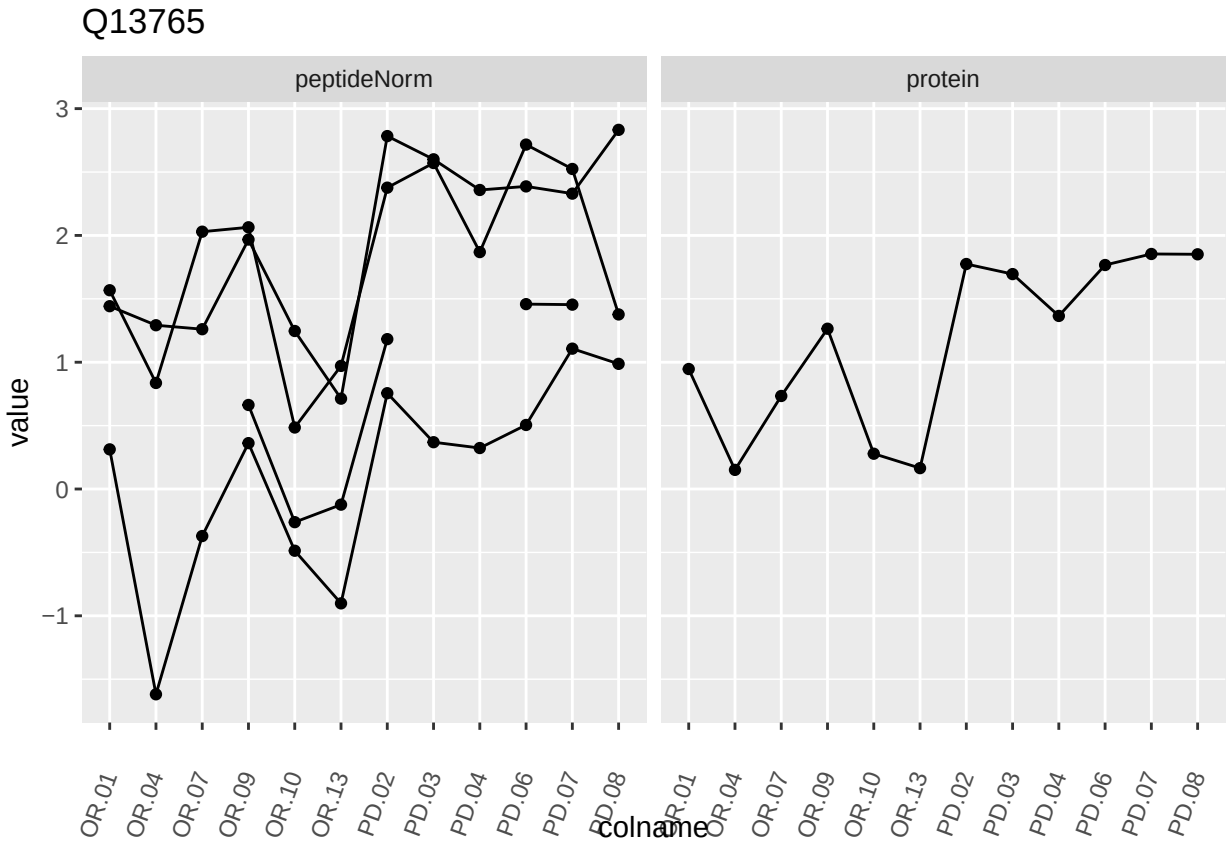

Q9NP79
peptideNorm
protein
peptide intensity (log2)
2
1
0
−1
−2
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
rowname
AALAPLPPLPAQFK
MFLYADNEDR
Q9NP79
prognosis
OR
PD

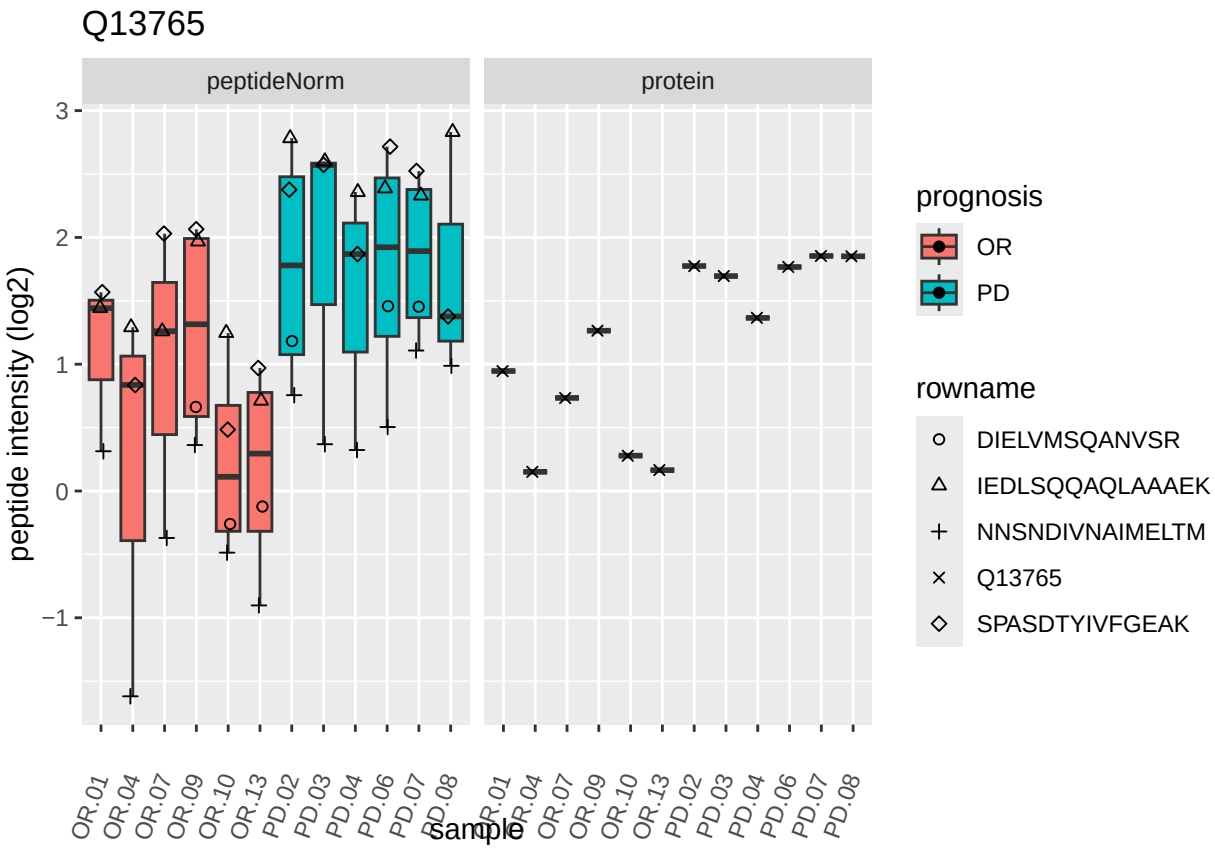

O60635
peptideNorm
protein
value
1.5
1.0
0.5
0.0
−0.5
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

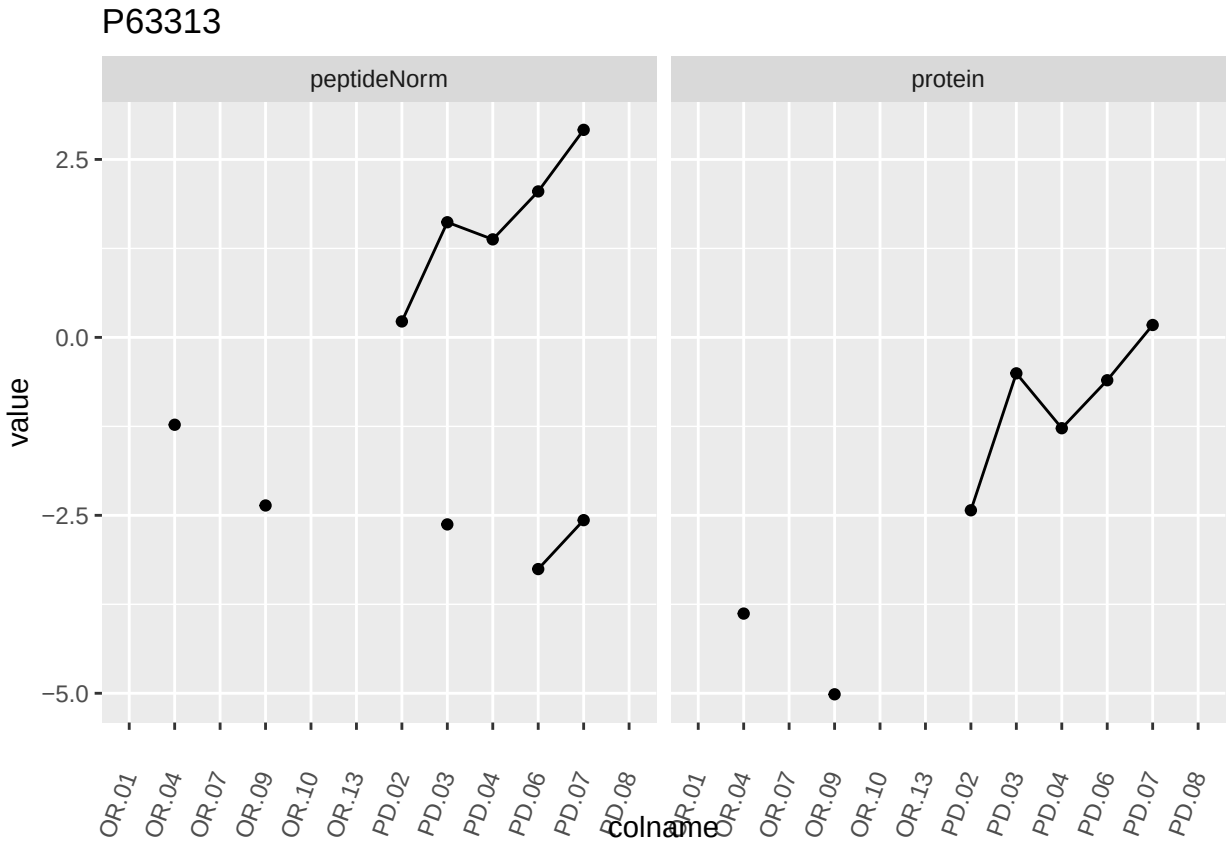

O60635
peptideNorm
protein
peptide intensity (log2)
1.5
1.0
0.5
0.0
−0.5
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
prognosis
OR
PD
rowname
O60635
VEGCFNQLLYDIR

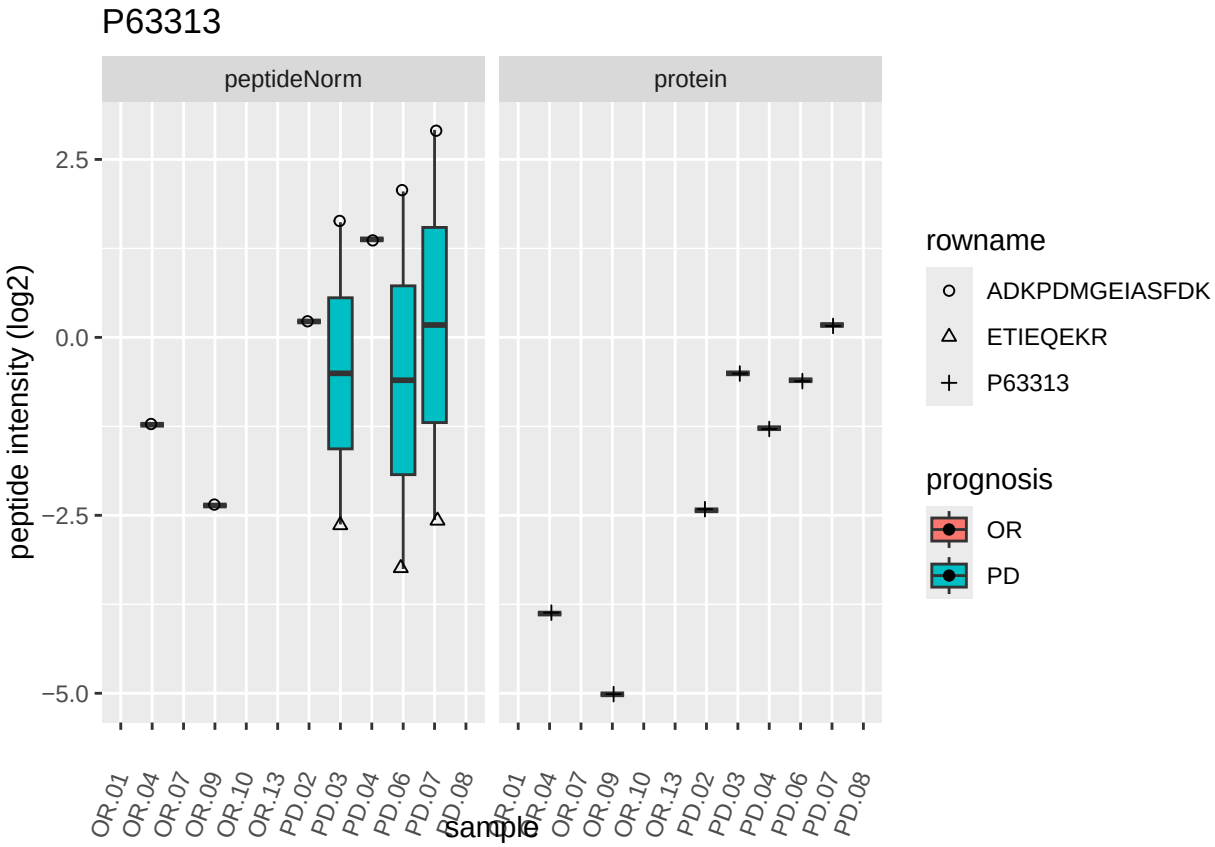

Q16543
peptideNorm
protein
value
2
0
−2
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

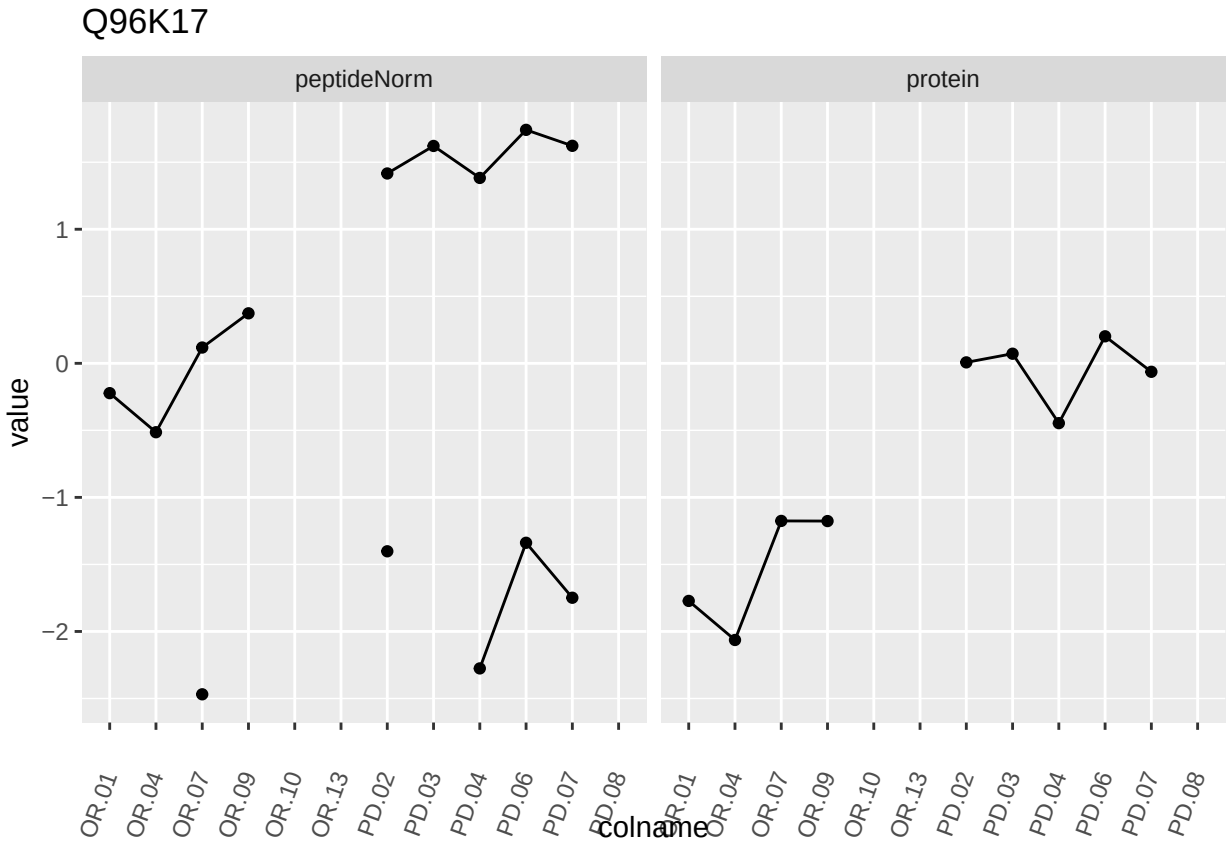

Q16543
peptideNorm
protein
peptide intensity (log2)
2
0
−2
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
rowname
AMKEYEEEER
DVQMLQDAISK
EGEEAGPGDPLLEAVPK
LGPGGLDPVEVYESLPEELQK
LQAEAQQLR
Q16543
SMVNTKPEK
SMVNTKPEKTEEDSEEVR
TGDEKDVSV
prognosis
OR
PD

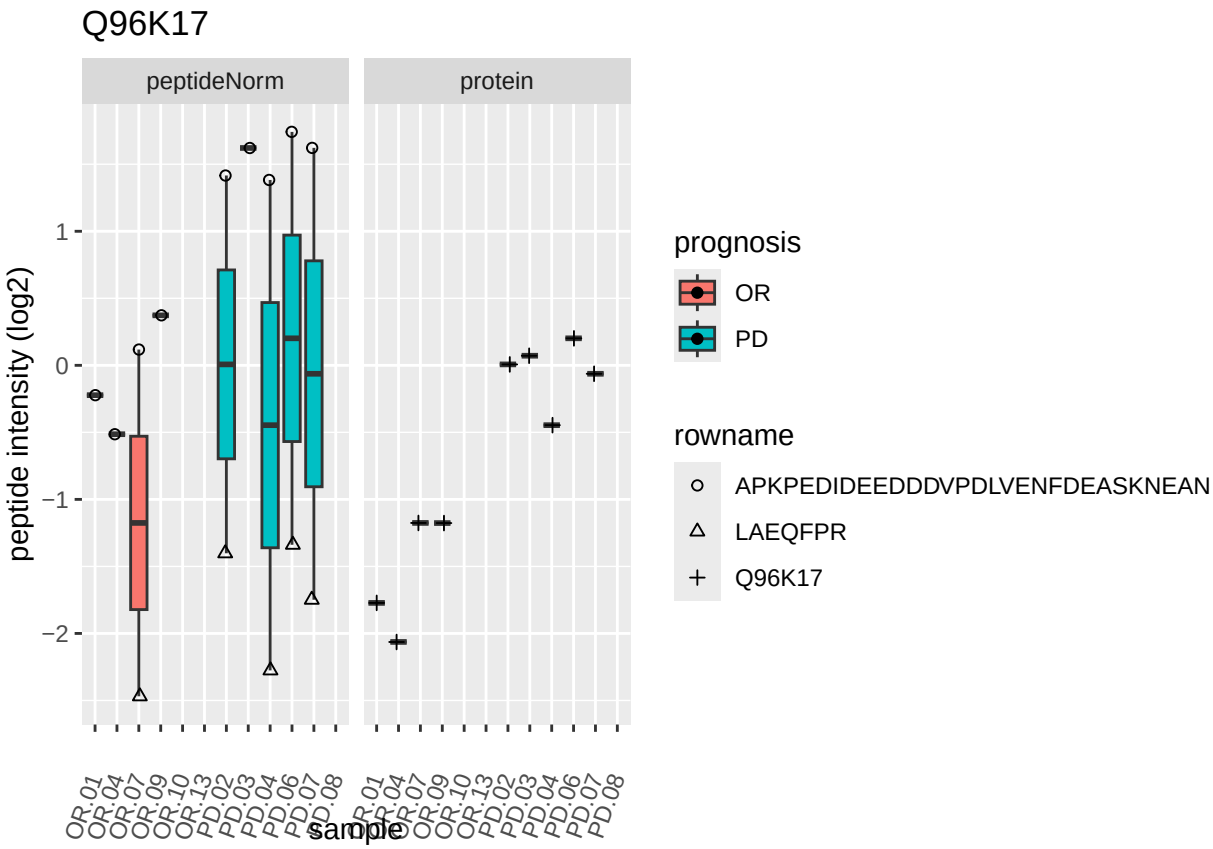

Q13765
peptideNorm
protein
value
3
2
1
0
−1
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

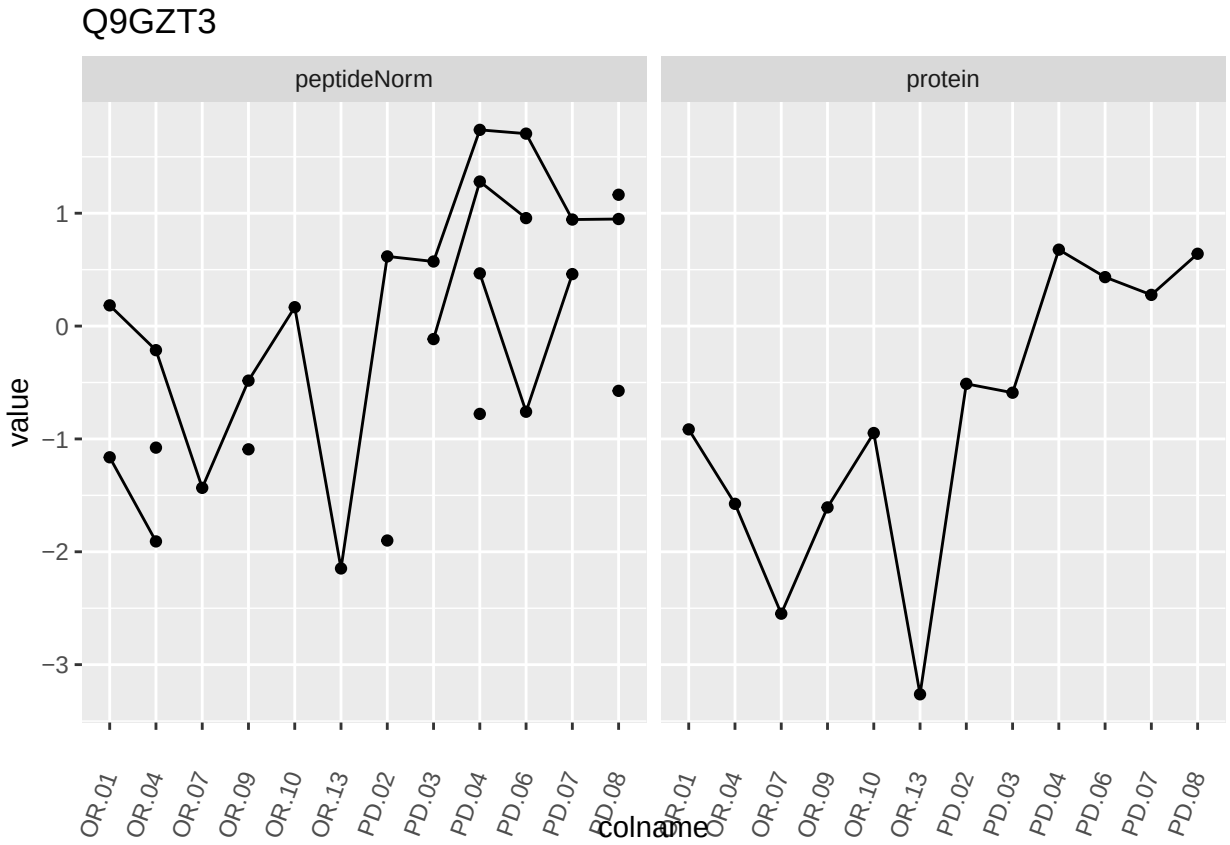

Q13765
peptideNorm
protein
peptide intensity (log2)
3
2
1
0
−1
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
prognosis
OR
PD
rowname
DIELVMSQANVSR
IEDLSQQAQLAAAEK
NNSNDIVNAIMELTM
Q13765
SPASDTYIVFGEAK

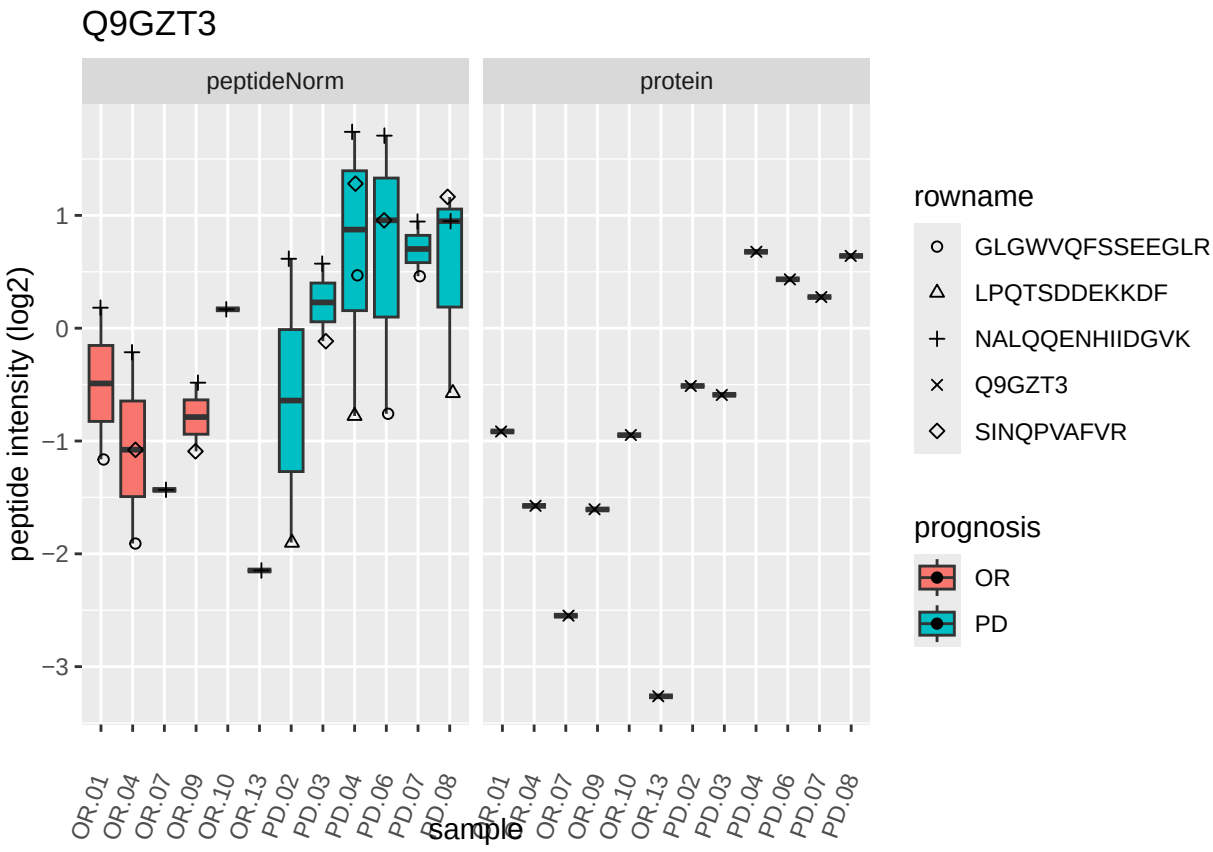

P63313
peptideNorm
protein
2.5
0.0
−2.5
−5.0
value
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

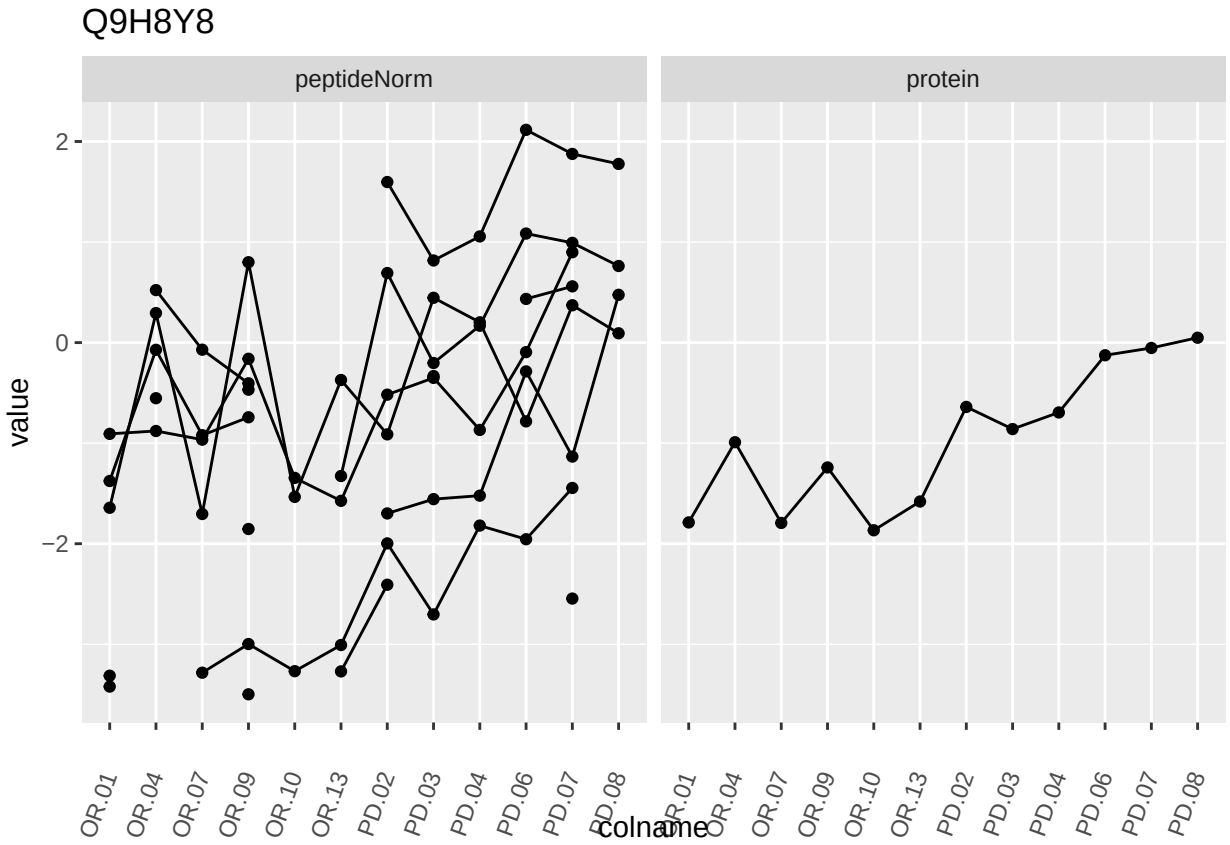

P63313
peptideNorm
protein
peptide intensity (log2)
2.5
0.0
−2.5
−5.0
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
rowname
ADKPDMGEIASFDK
ETIEQEKR
P63313
prognosis
OR
PD

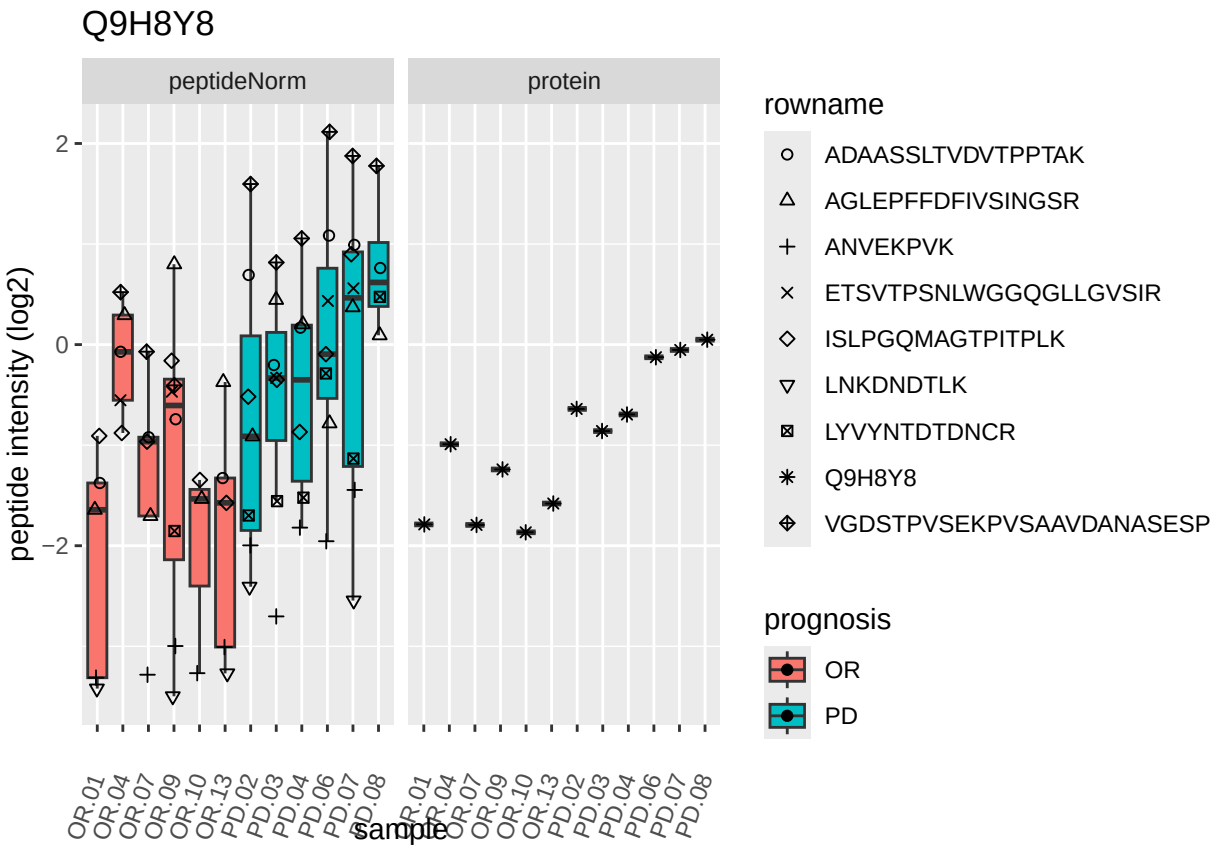

Q96K17
peptideNorm
protein
value
1
0
−1
−2
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

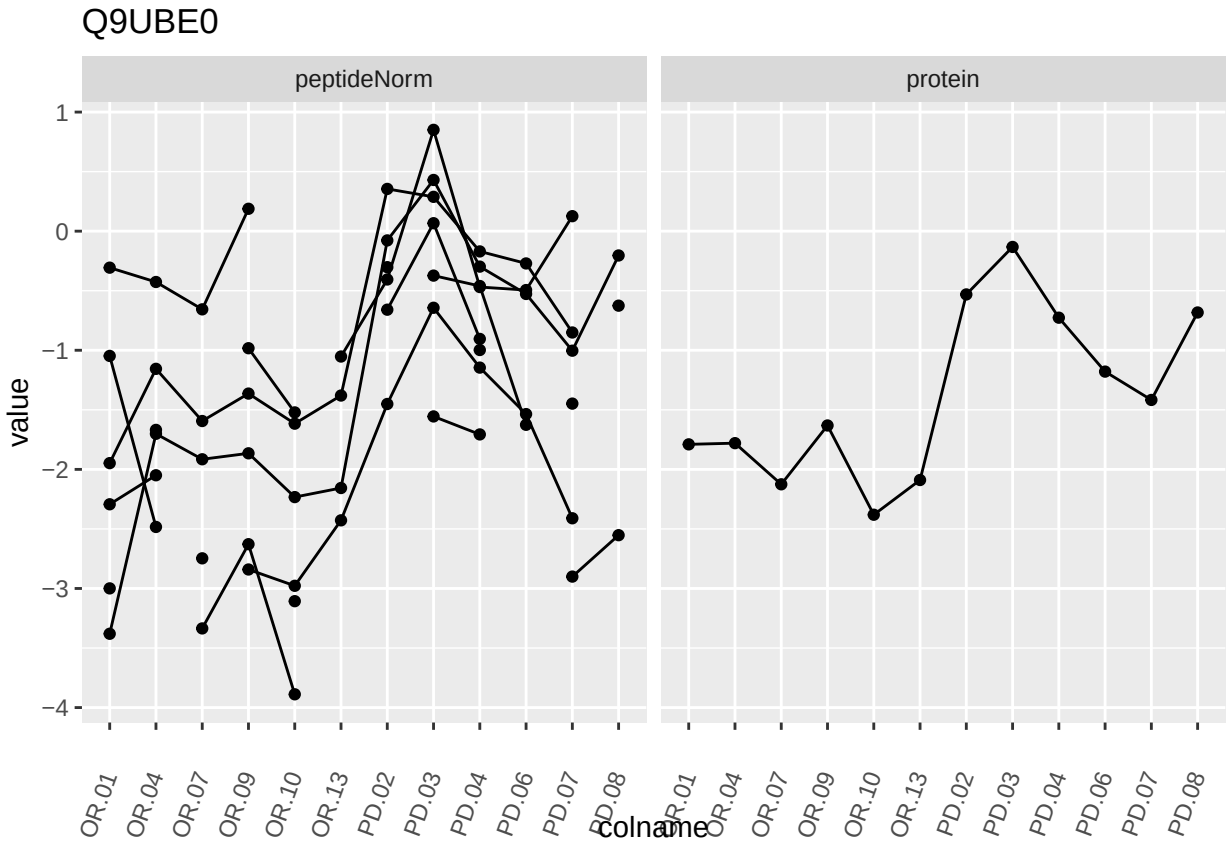

Q96K17
peptideNorm
protein
peptide intensity (log2)
1
0
−1
−2
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
prognosis
OR
PD
rowname
APKPEDIDEEDDDVPDLVENFDEASKNEAN
LAEQFPR
Q96K17

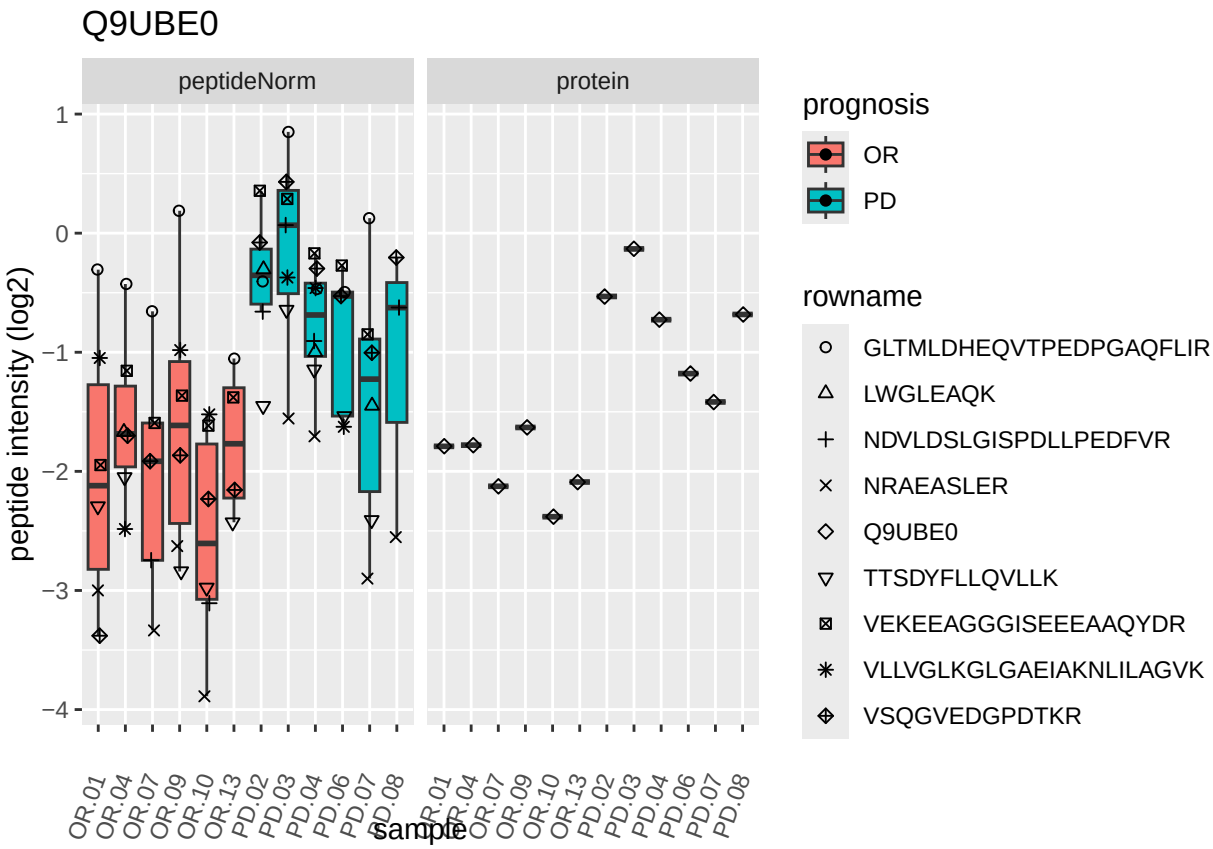

Q9GZT3
peptideNorm
protein
value
1
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

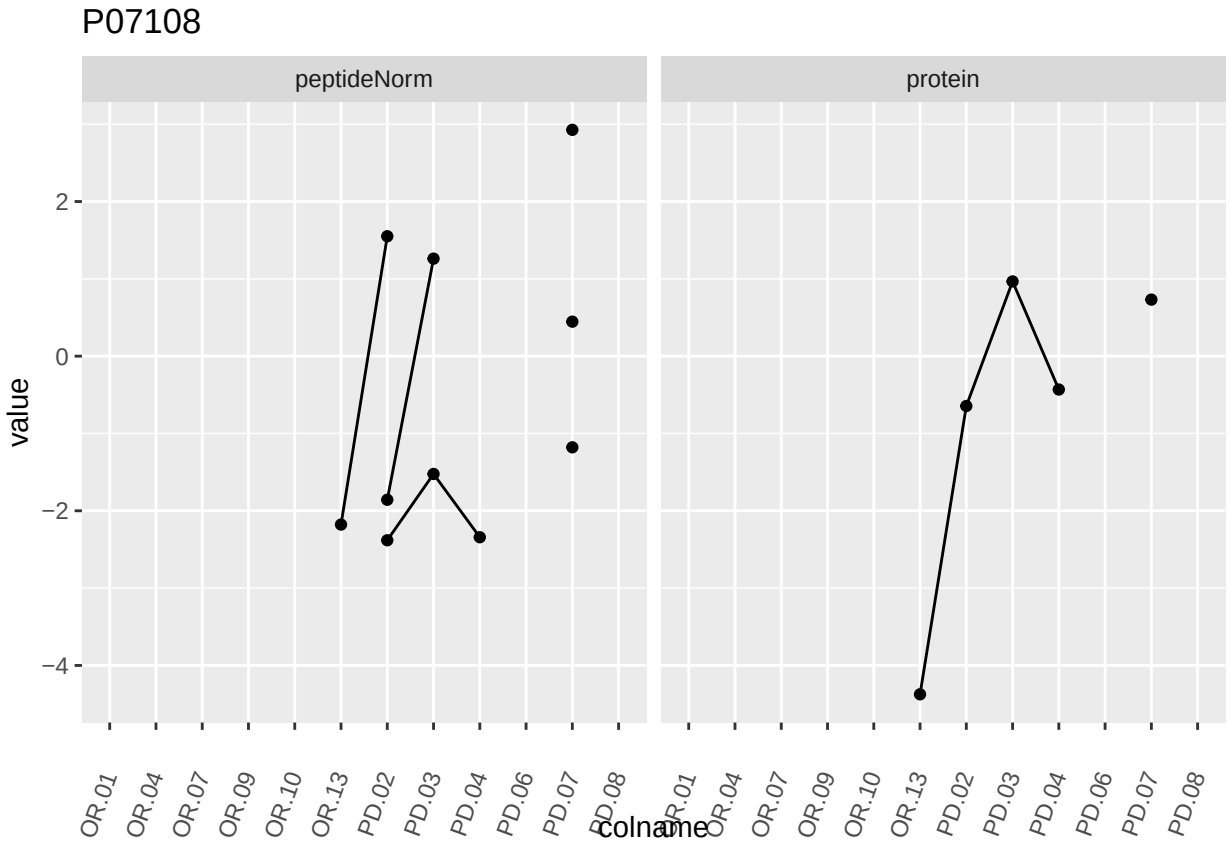

Q9GZT3
peptideNorm
protein
peptide intensity (log2)
1
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
rowname
GLGWVQFSSEEGLR
LPQTSDDEKKDF
NALQQENHIIDGVK
Q9GZT3
SINQPVAFVR
prognosis
OR
PD

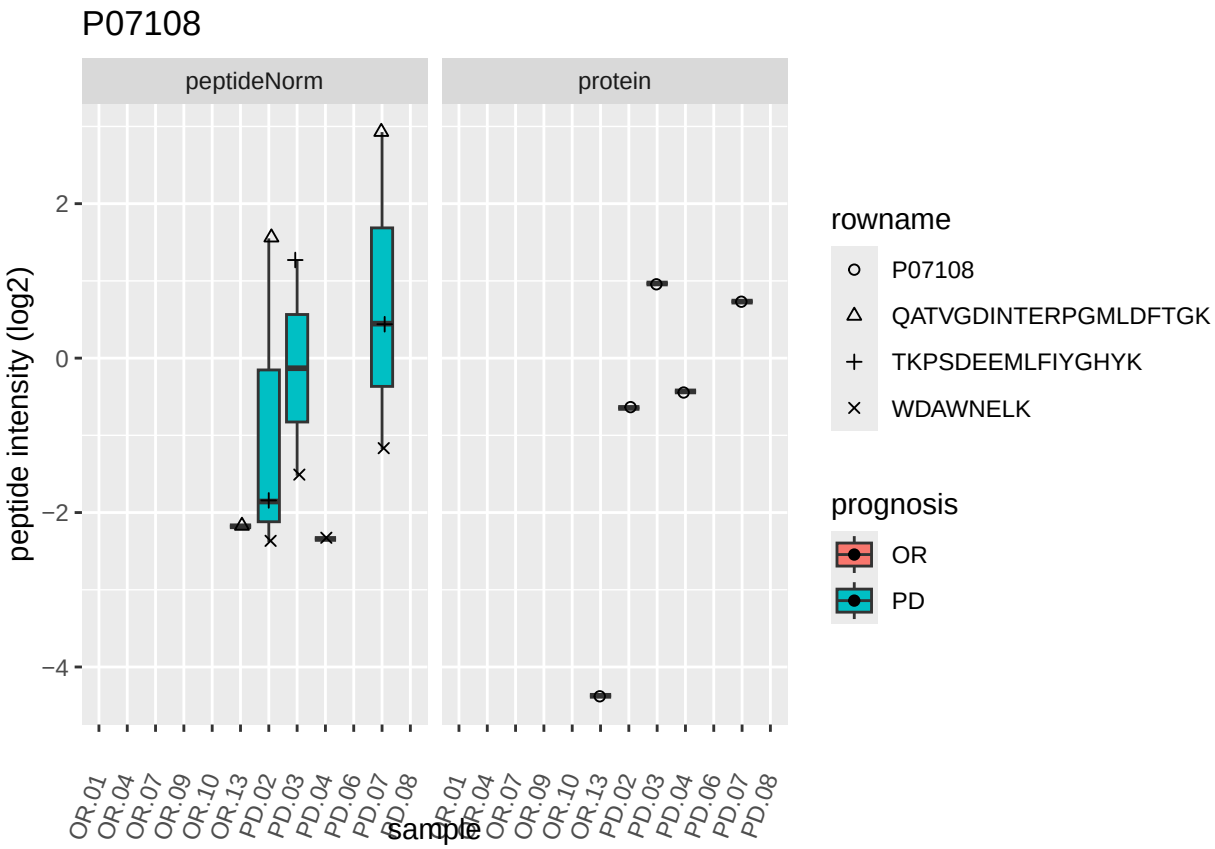

Q9H8Y8
peptideNorm
protein
value
2
0
−2
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

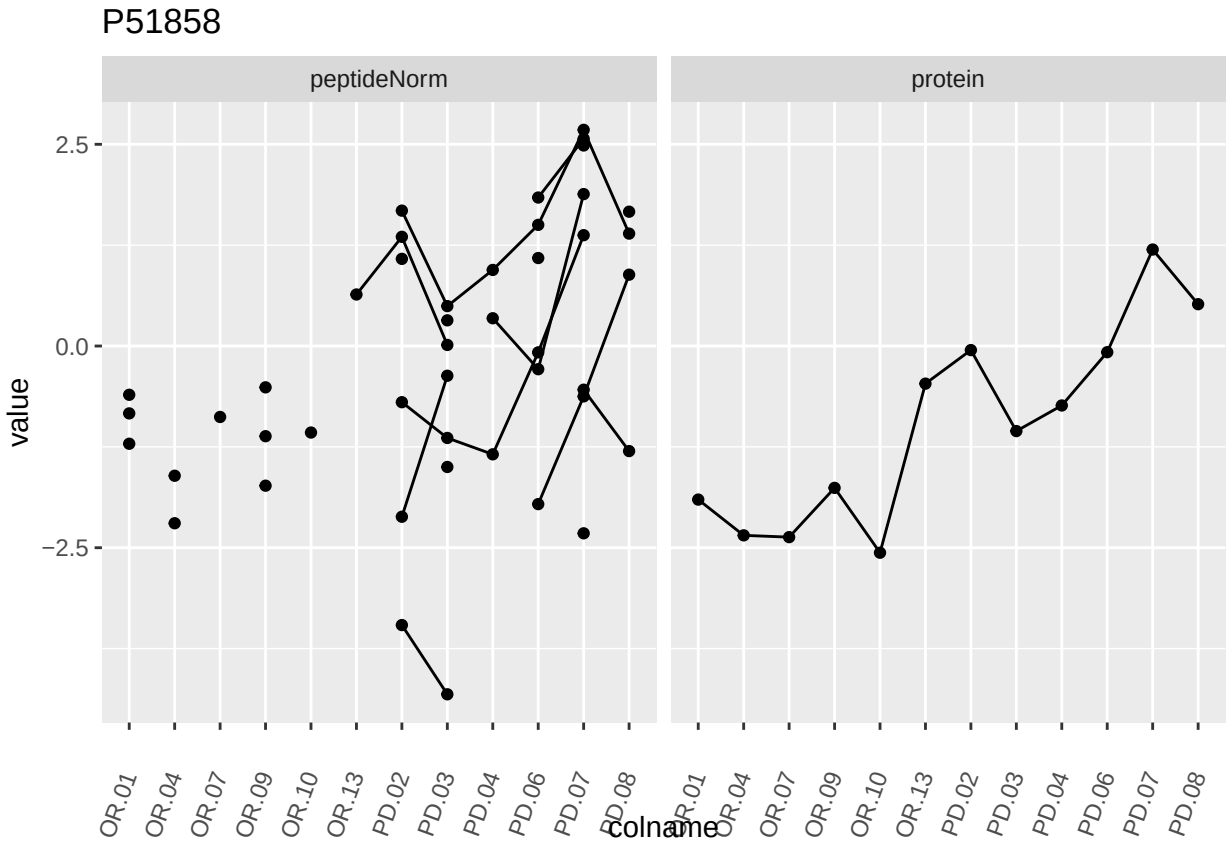

Q9H8Y8
peptideNorm
protein
peptide intensity (log2)
2
0
−2
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
rowname
ADAASSLTVDVTPPTAK
AGLEPFFDFIVSINGSR
ANVEKPVK
ETSVTPSNLWGGQGLLGVSIR
ISLPGQMAGTPITPLK
LNKDNDTLK
LYVYNTDTDNCR
Q9H8Y8
VGDSTPVSEKPVSAAVDANASESP
prognosis
OR
PD

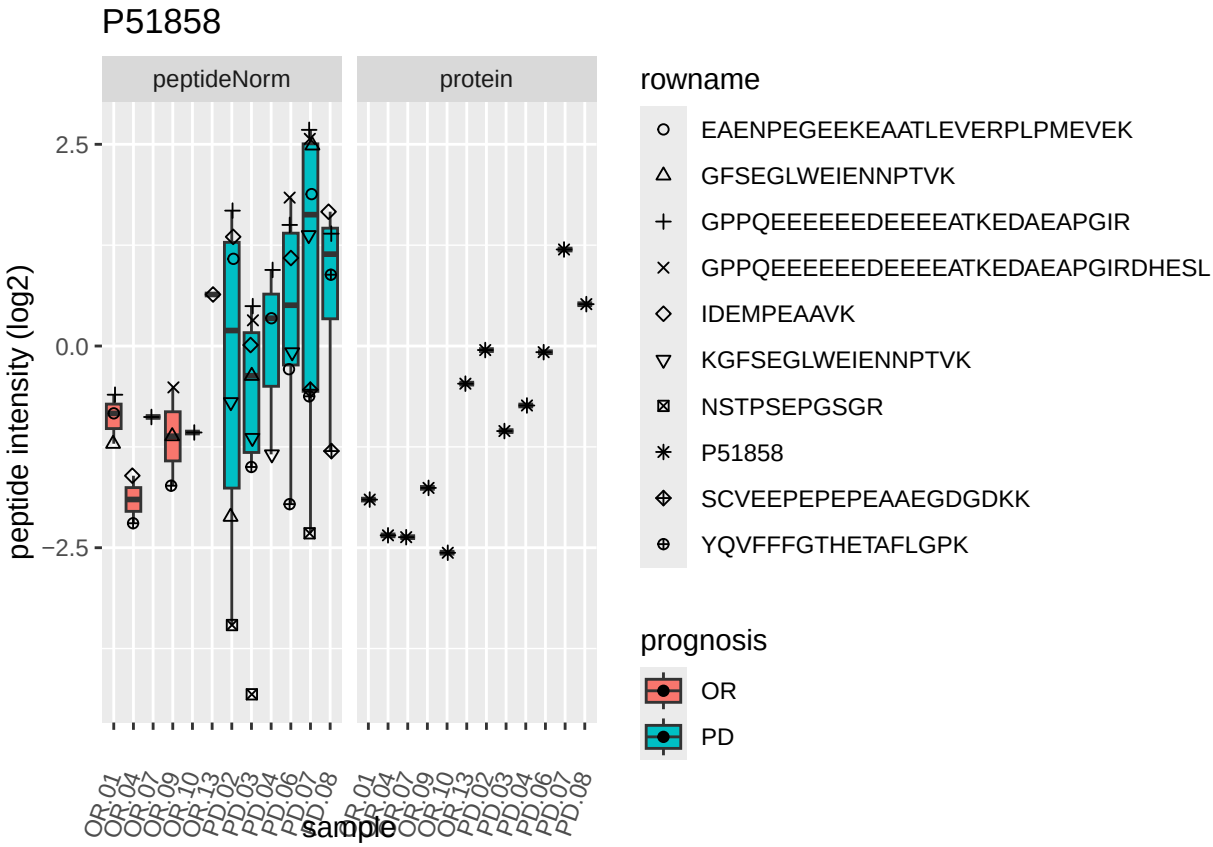

Q9UBE0
peptideNorm
protein
value
1
0
−1
−2
−3
−4
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

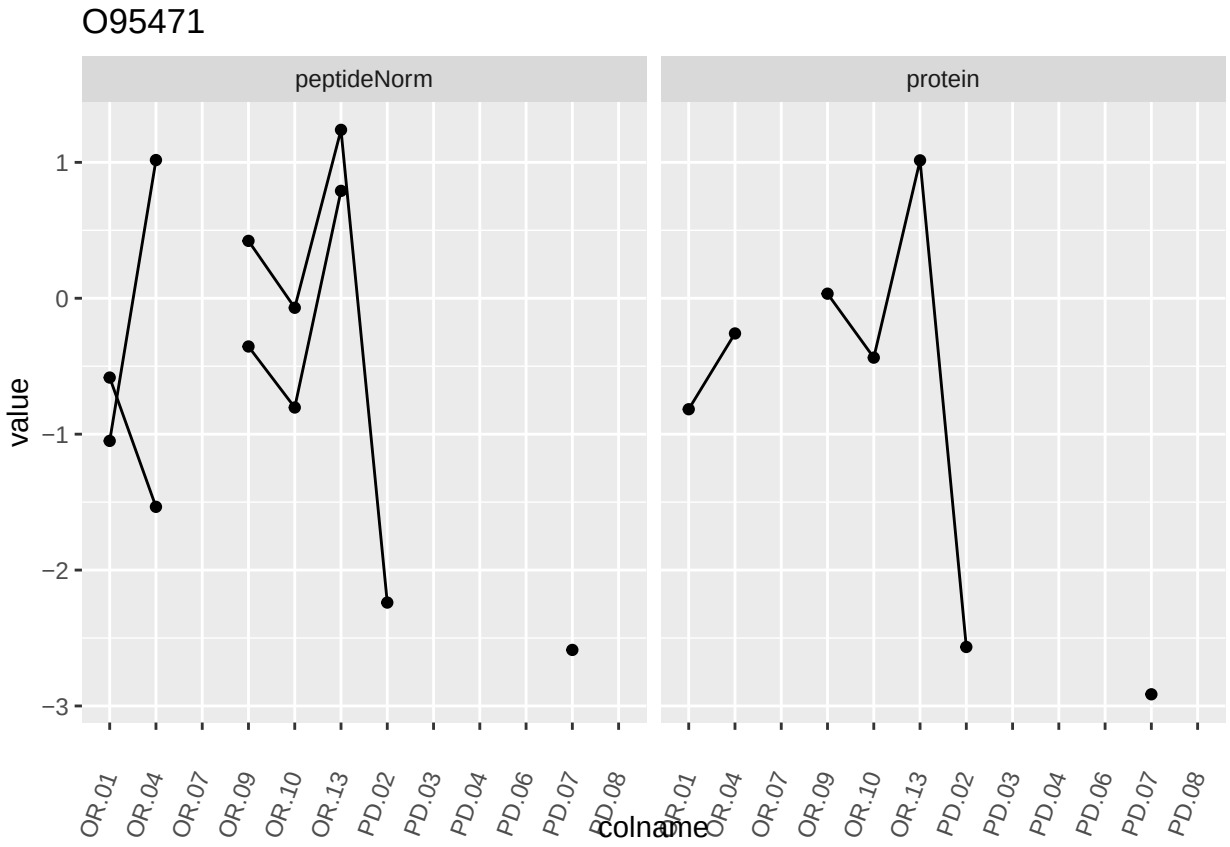

Q9UBE0
peptideNorm
protein
peptide intensity (log2)
1
0
−1
−2
−3
−4
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
prognosis
OR
PD
rowname
GLTMLDHEQVTPEDPGAQFLIR
LWGLEAQK
NDVLDSLGISPDLLPEDFVR
NRAEASLER
Q9UBE0
TTSDYFLLQVLLK
VEKEEAGGGISEEEAAQYDR
VLLVGLKGLGAEIAKNLILAGVK
VSQGVEDGPDTKR

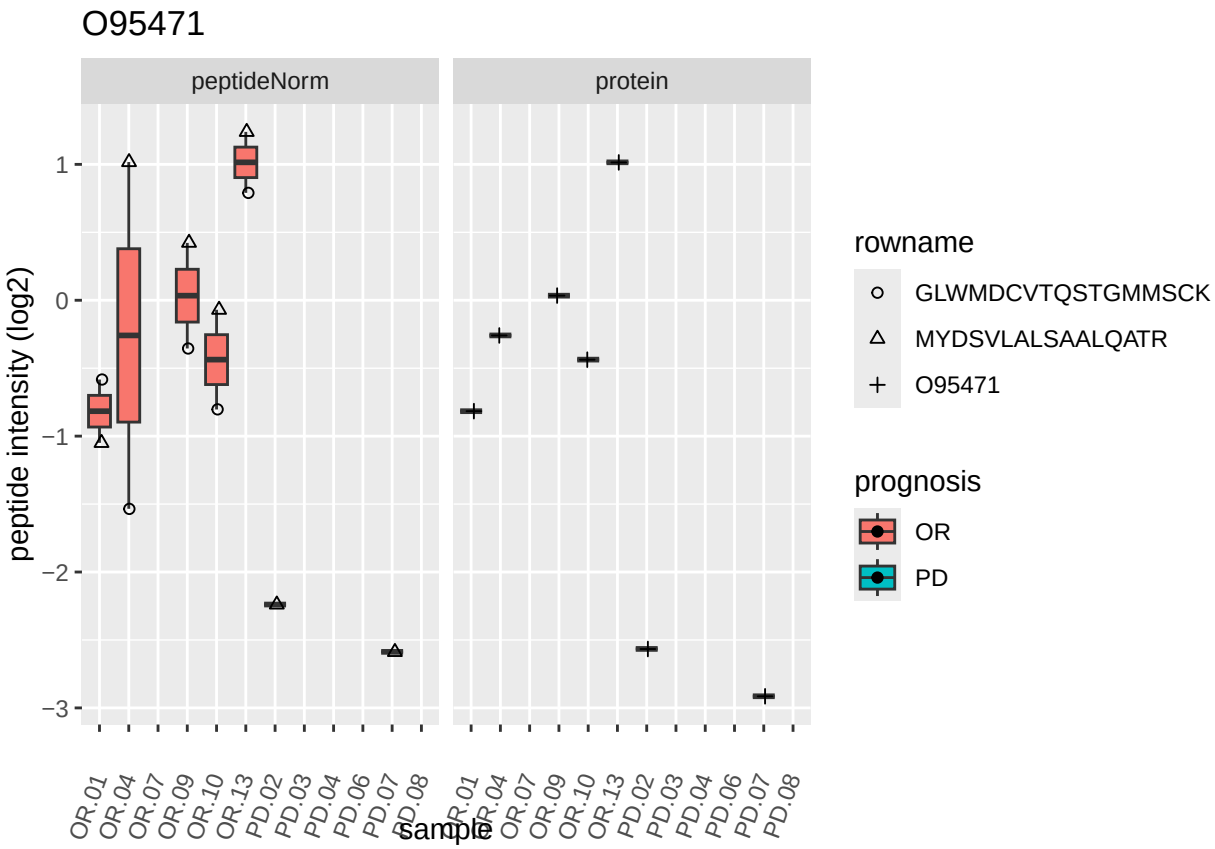

P07108
peptideNorm
protein
value
2
0
−2
−4
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

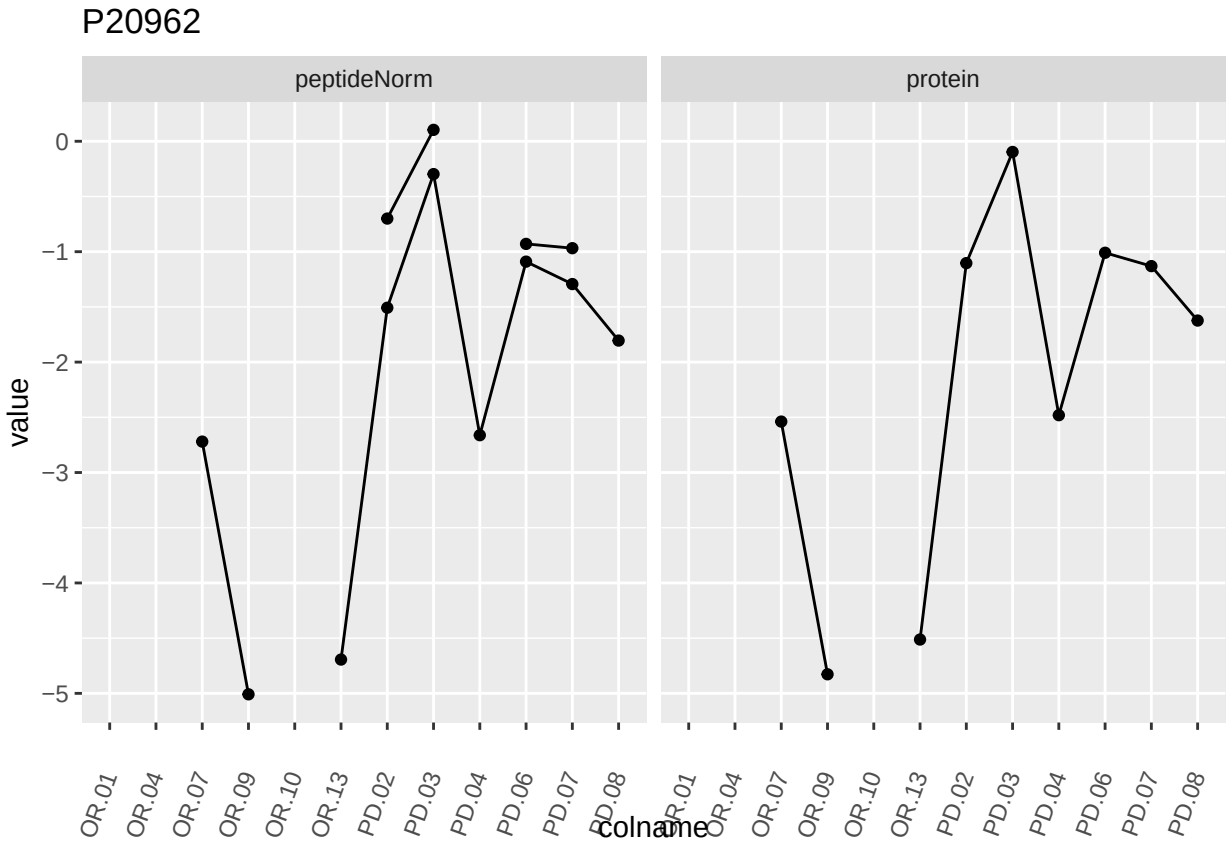

P07108
peptideNorm
protein
peptide intensity (log2)
2
0
−2
−4
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
rowname
P07108
QATVGDINTERPGMLDFTGK
TKPSDEEMLFIYGHYK
WDAWNELK
prognosis
OR
PD

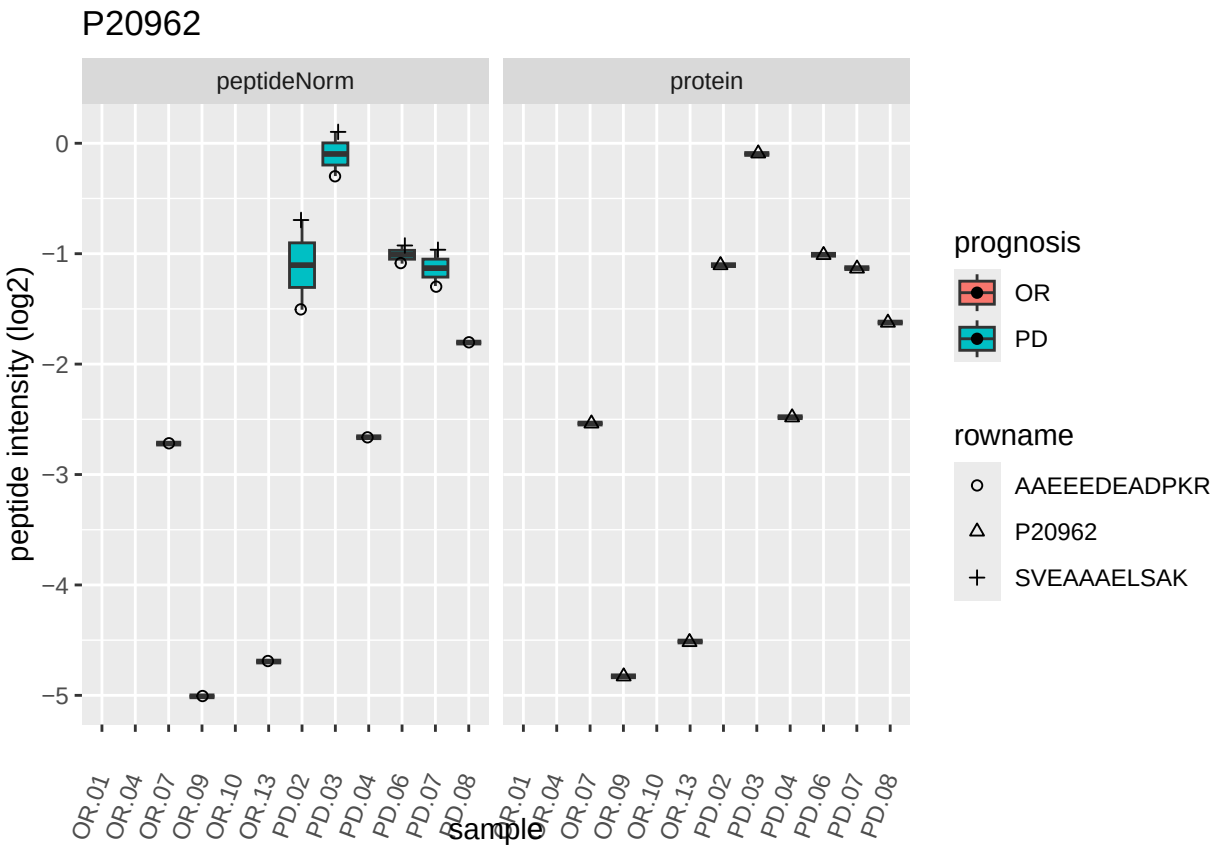

P51858
peptideNorm
protein
value
2.5
0.0
−2.5
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

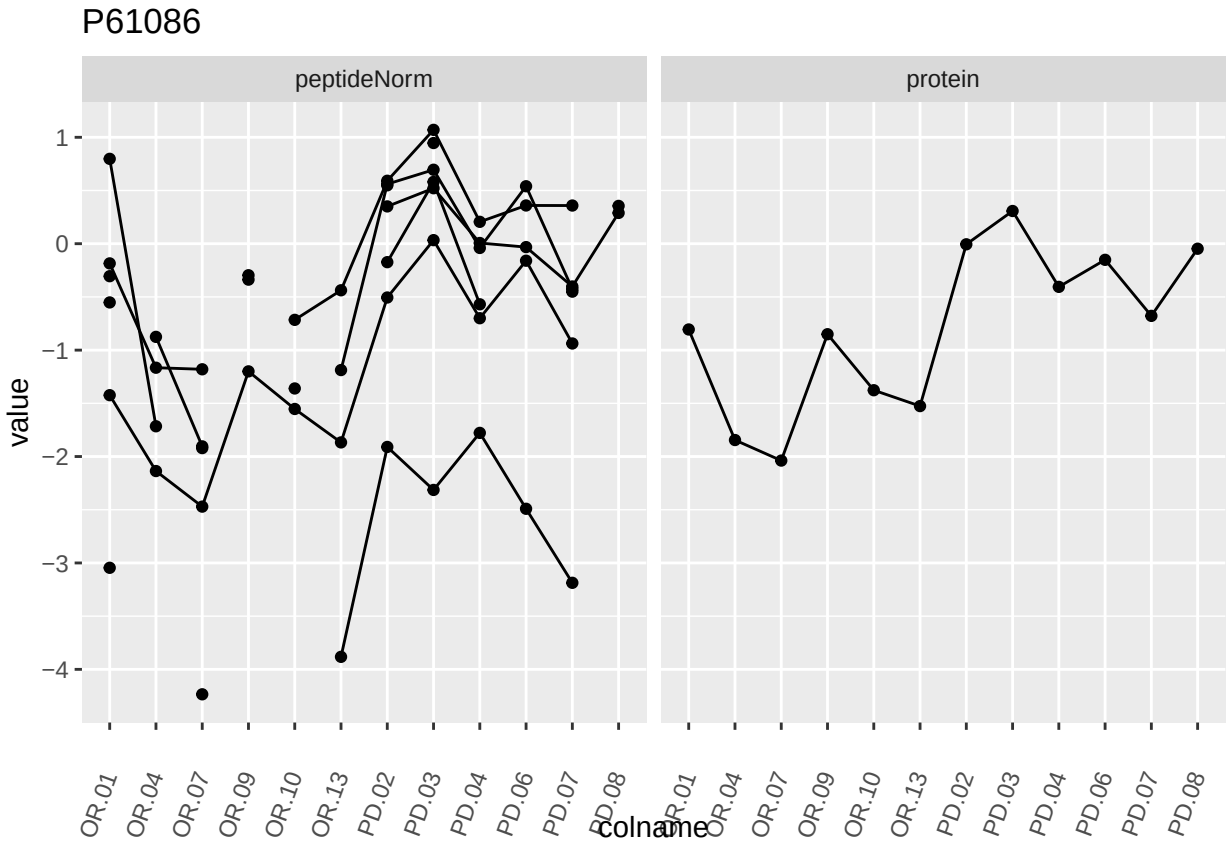

P51858
peptideNorm
protein
peptide intensity (log2)
2.5
0.0
−2.5
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
rowname
EAENPEGEEKEAATLEVERPLPMEVEK
GFSEGLWEIENNPTVK
GPPQEEEEEEDEEEEATKEDAEAPGIR
GPPQEEEEEEDEEEEATKEDAEAPGIRDHESL
IDEMPEAAVK
KGFSEGLWEIENNPTVK
NSTPSEPGSGR
P51858
SCVEEPEPEPEAAEGDGDKK
YQVFFFGTHETAFLGPK
prognosis
OR
PD

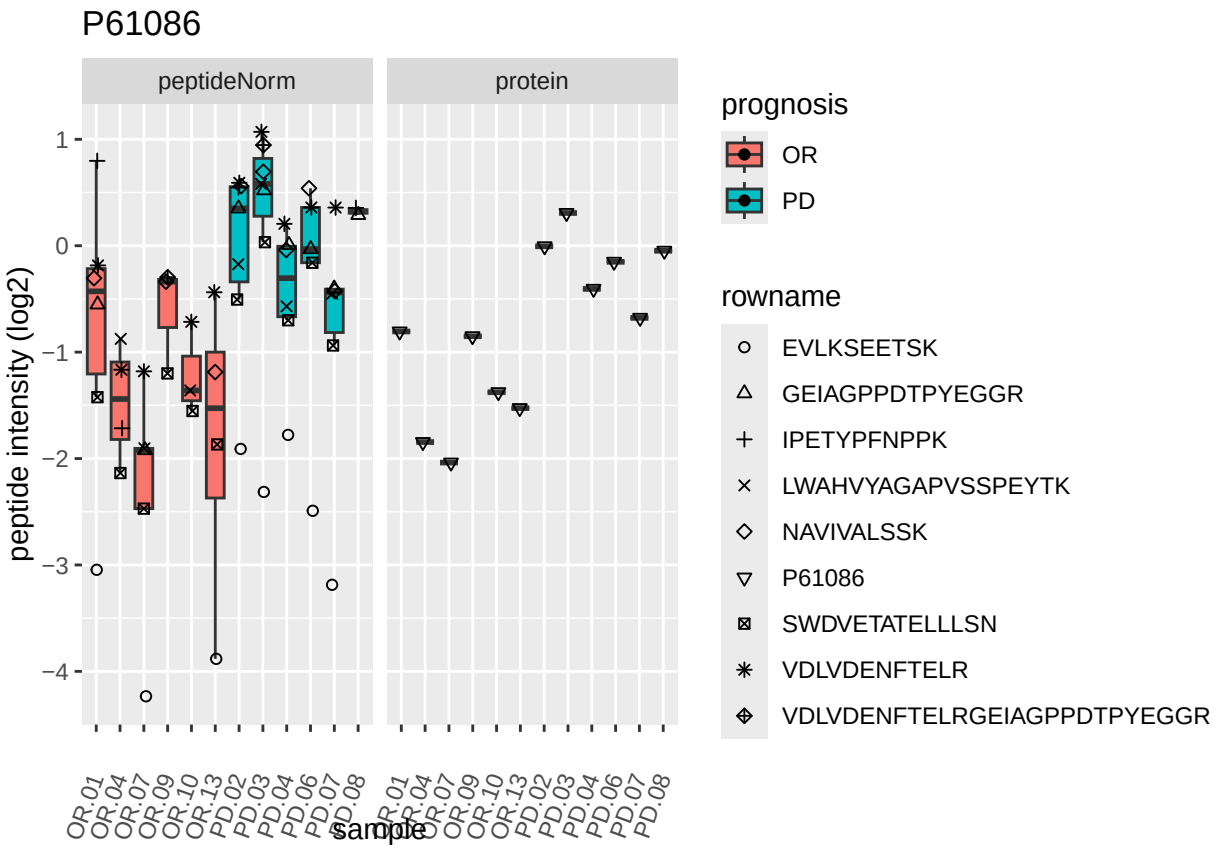

O95471
peptideNorm
protein
value
1
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

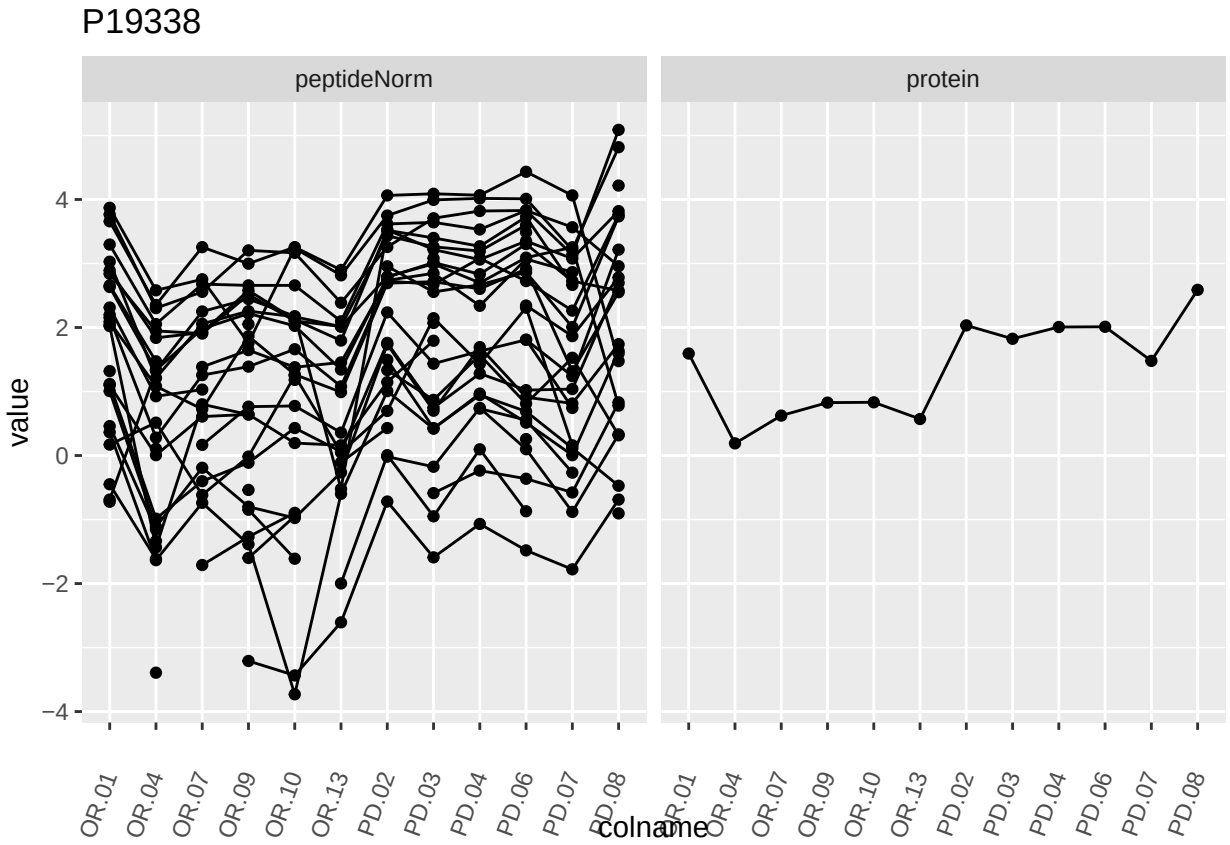

O95471
peptideNorm
protein
peptide intensity (log2)
1
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
rowname
GLWMDCVTQSTGMMSCK
MYDSVLALSAALQATR
O95471
prognosis
OR
PD

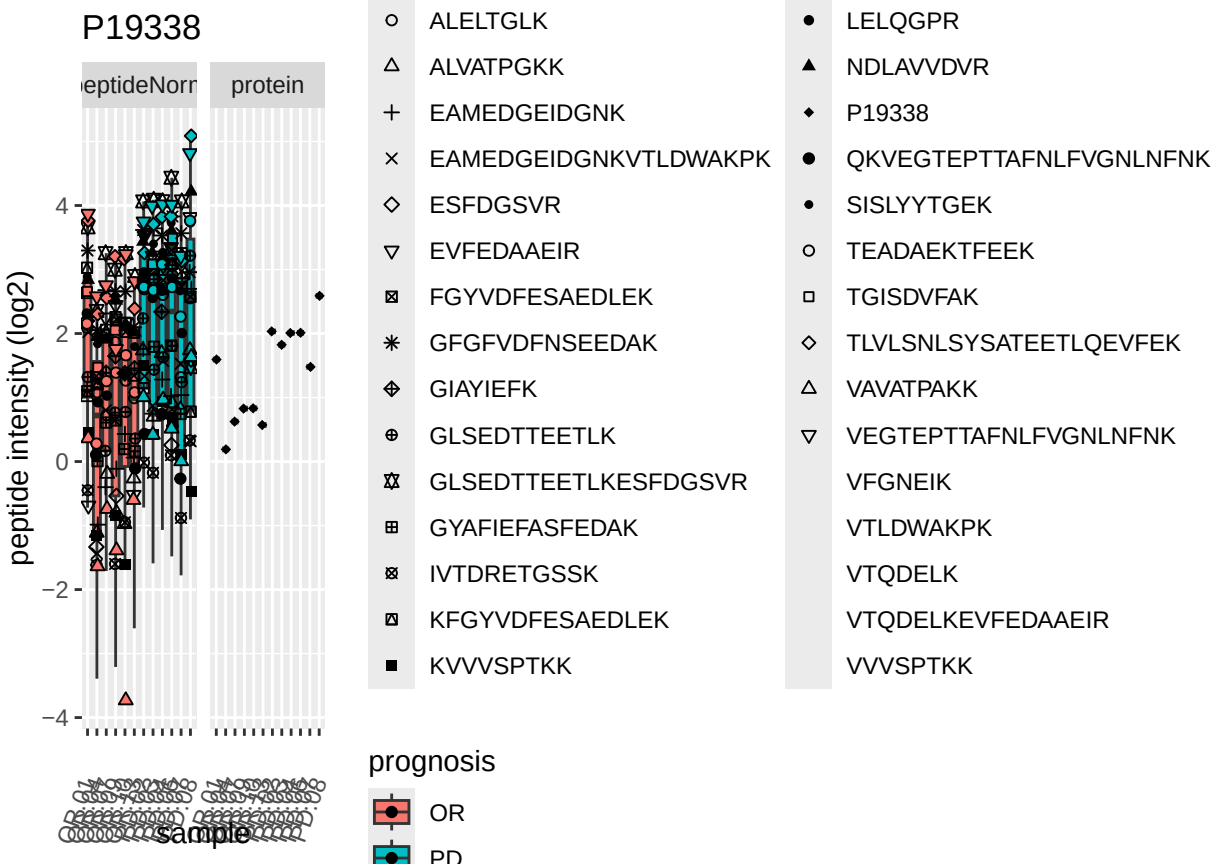

P20962
peptideNorm
protein
value
0
−1
−2
−3
−4
−5
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

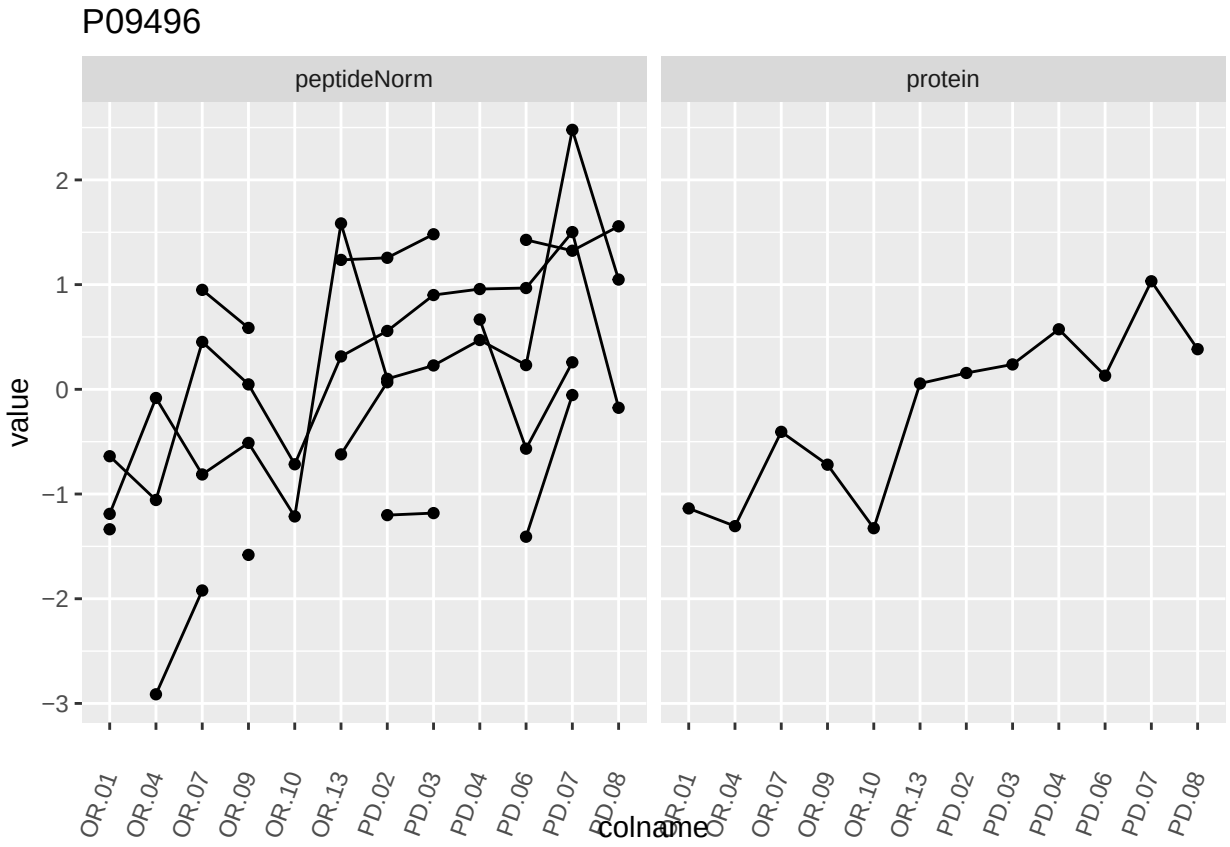

P20962
peptideNorm
protein
peptide intensity (log2)
0
−1
−2
−3
−4
−5
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
prognosis
OR
PD
rowname
AAEEEDEADPKR
P20962
SVEAAAELSAK

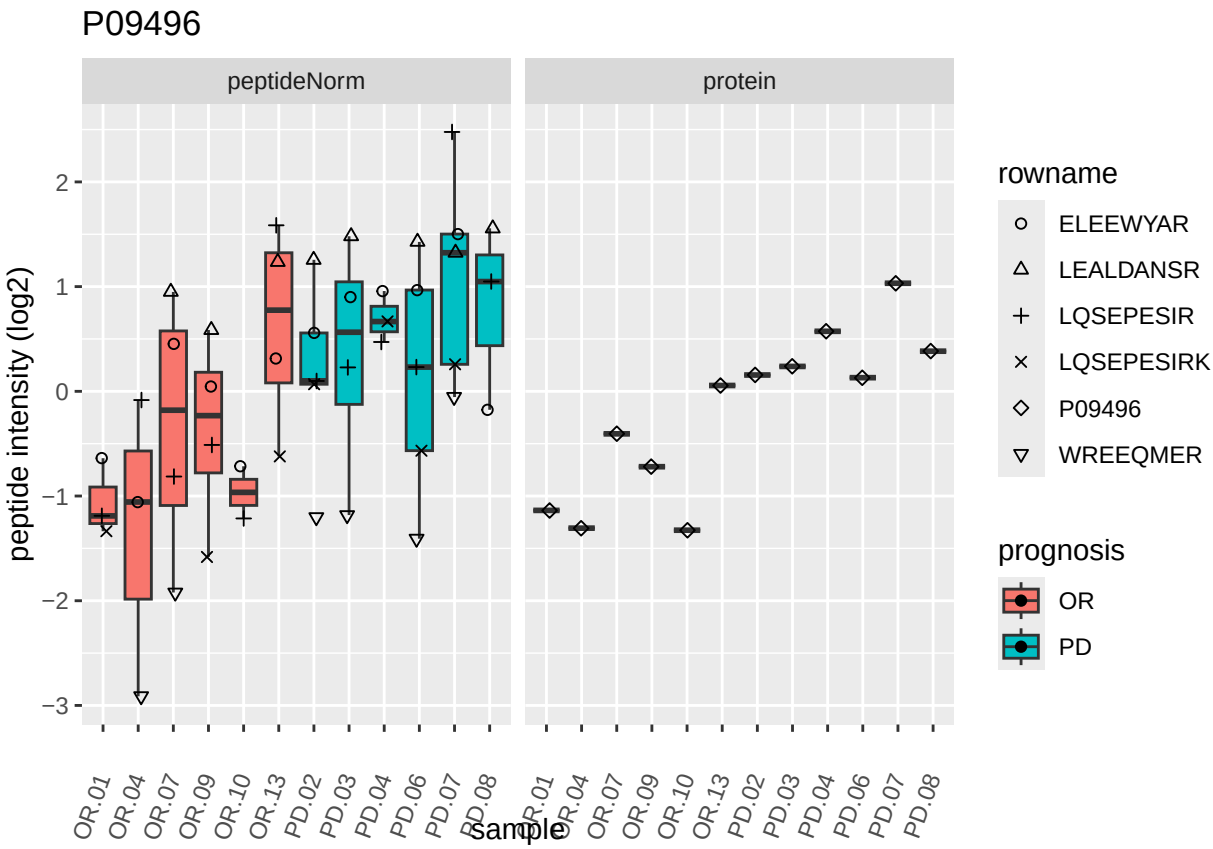

P61086
peptideNorm
protein
value
1
0
−1
−2
−3
−4
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

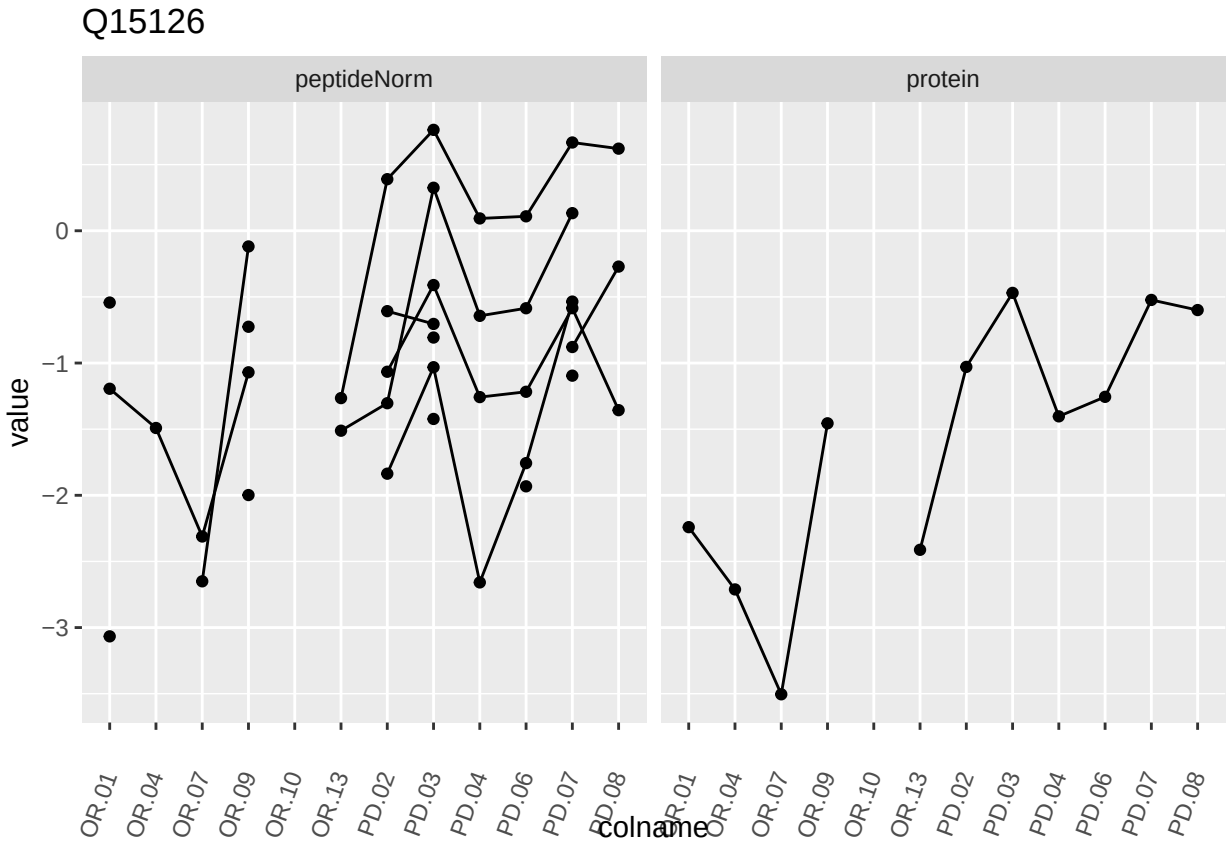

P61086
peptideNorm
protein
peptide intensity (log2)
1
0
−1
−2
−3
−4
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
prognosis
OR
PD
rowname
EVLKSEETSK
GEIAGPPDTPYEGGR
IPETYPFNPPK
LWAHVYAGAPVSSPEYTK
NAVIVALSSK
P61086
SWDVETATELLLSN
VDLVDENFTELR
VDLVDENFTELRGEIAGPPDTPYEGGR

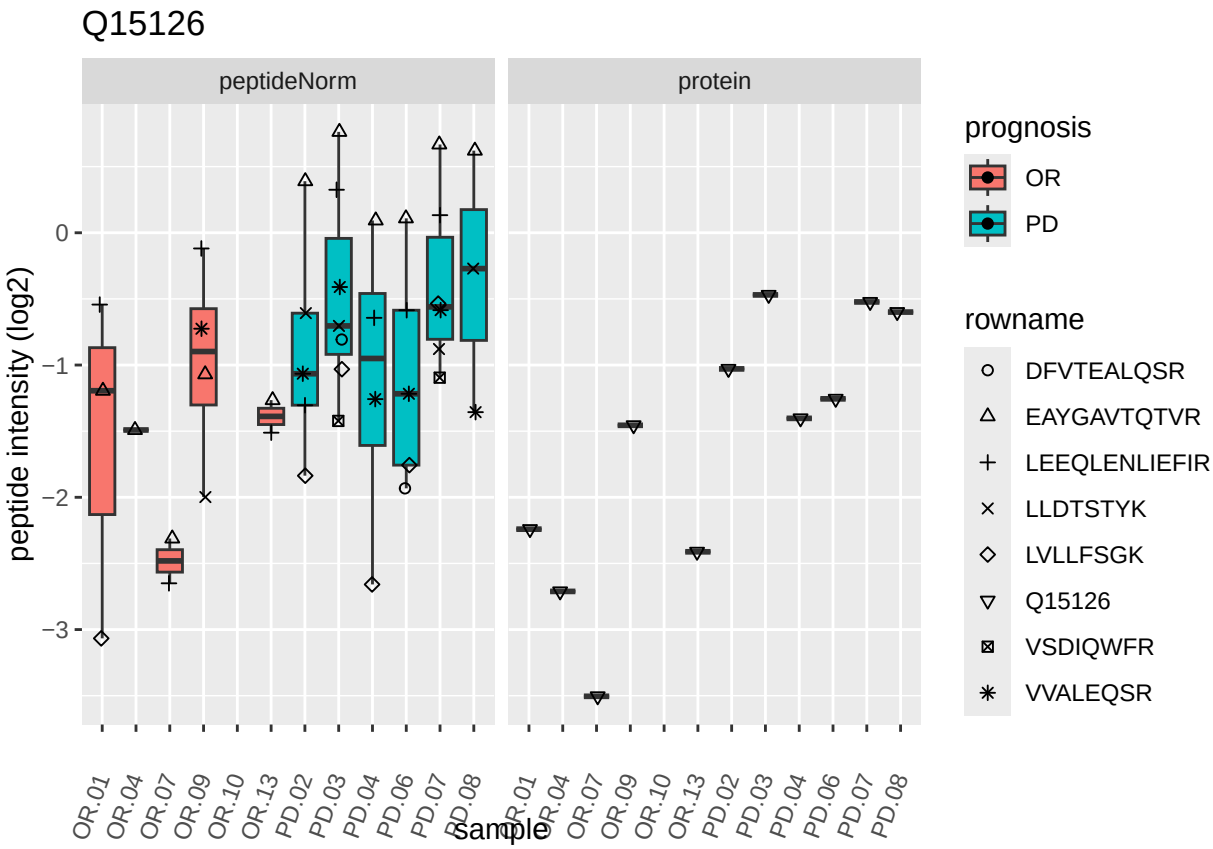

P19338
peptideNorm
protein
value
4
2
0
−2
−4
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

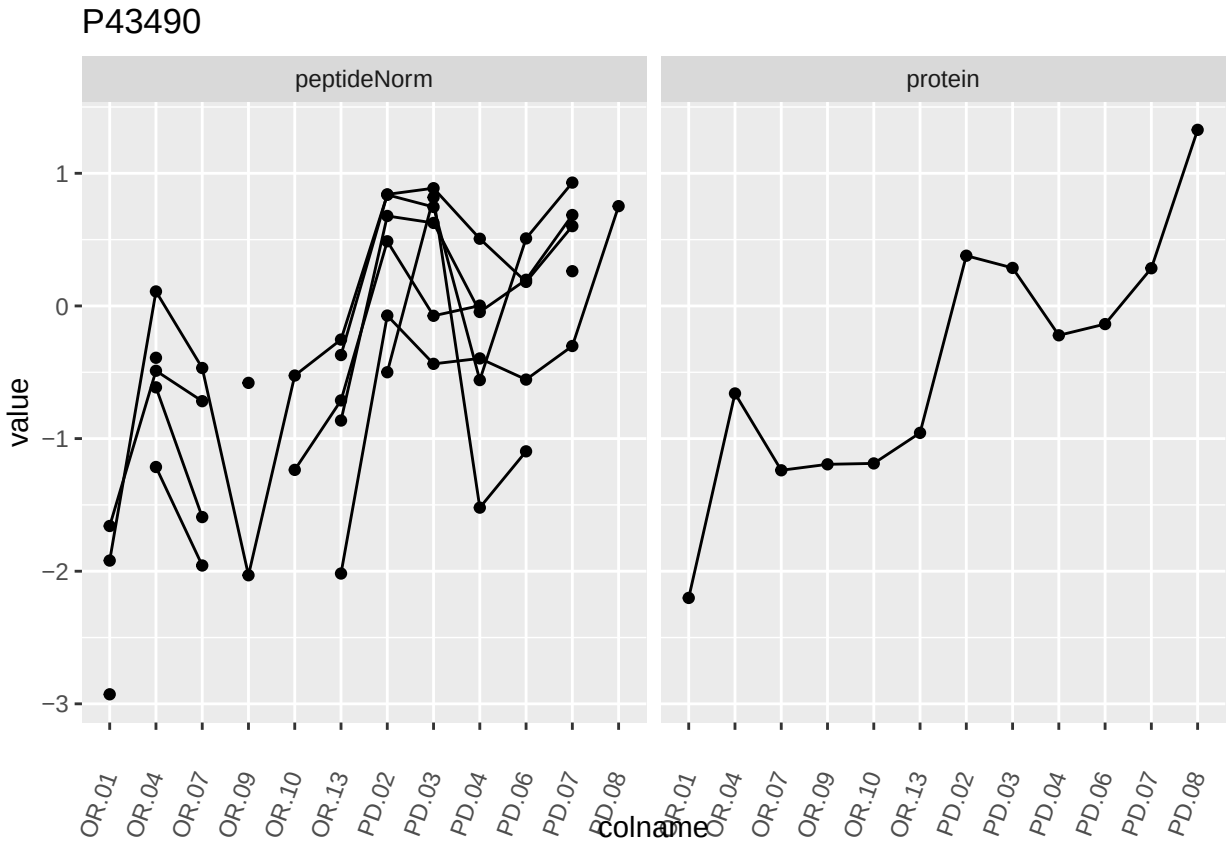

P19338
eptideNorm
protein
peptide intensity (log2)
4
2
0
−2
−4
sample
ALELTGLK
ALVATPGKK
EAMEDGEIDGNK
EAMEDGEIDGNKVTLDWAKPK
ESFDGSVR
EVFEDAAEIR
FGYVDFESAEDLEK
GFGFVDFNSEEDAK
GIAYIEFK
GLSEDTTEETLK
GLSEDTTEETLKESFDGSVR
GYAFIEFASFEDAK
IVTDRETGSSK
KFGYVDFESAEDLEK
KVVVSPTKK
LELQGPR
NDLAVVDVR
P19338
QKVEGTEPTTAFNLFVGNLNFNK
SISLYYTGEK
TEADAEKTFEEK
TGISDVFAK
TLVLSNLSYSATEETLQEVFEK
VAVATPAKK
VEGTEPTTAFNLFVGNLNFNK
VFGNEIK
VTLDWAKPK
VTQDELK
VTQDELKEVFEDAAEIR
VVVSPTKK
prognosis
OR
PD

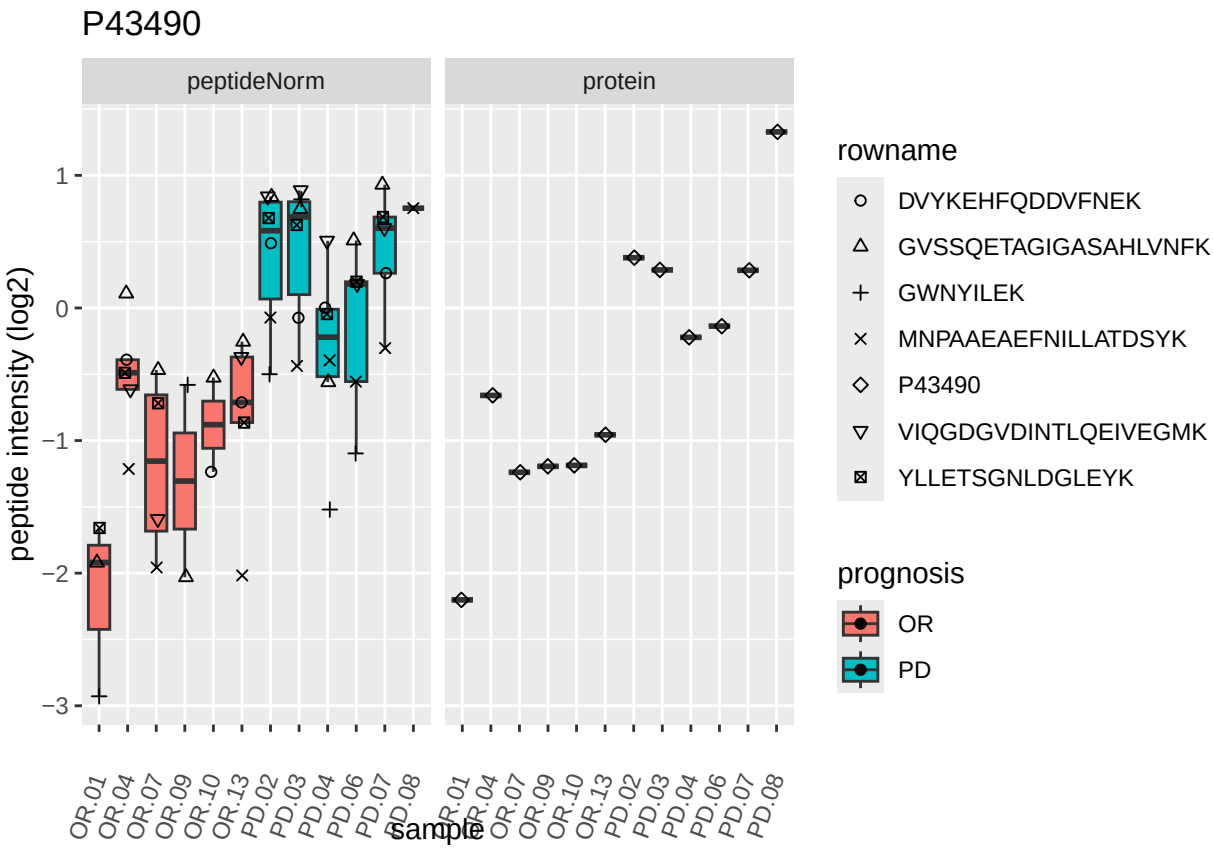

P09496
peptideNorm
protein
value
2
1
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

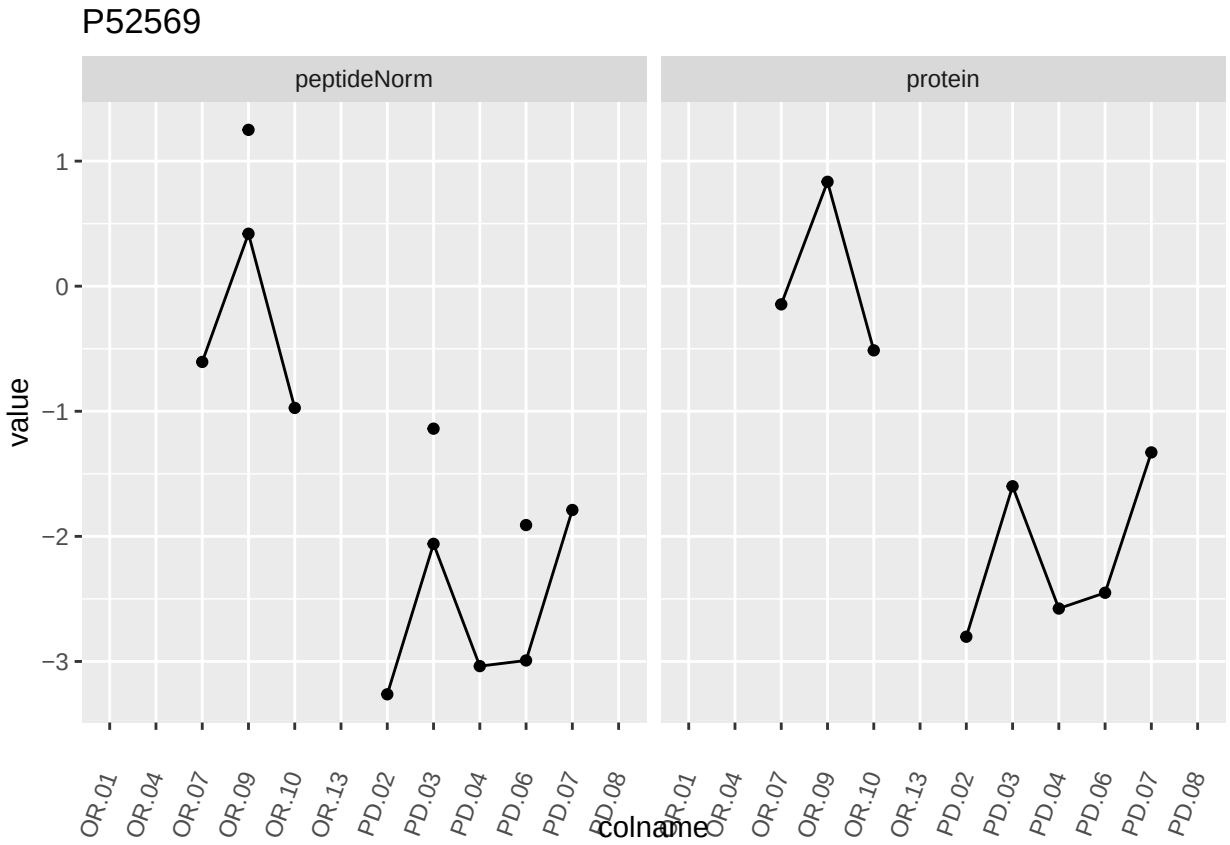

P09496
peptideNorm
protein
peptide intensity (log2)
2
1
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
rowname
ELEEWYAR
LEALDANSR
LQSEPESIR
LQSEPESIRK
P09496
WREEQMER
prognosis
OR
PD

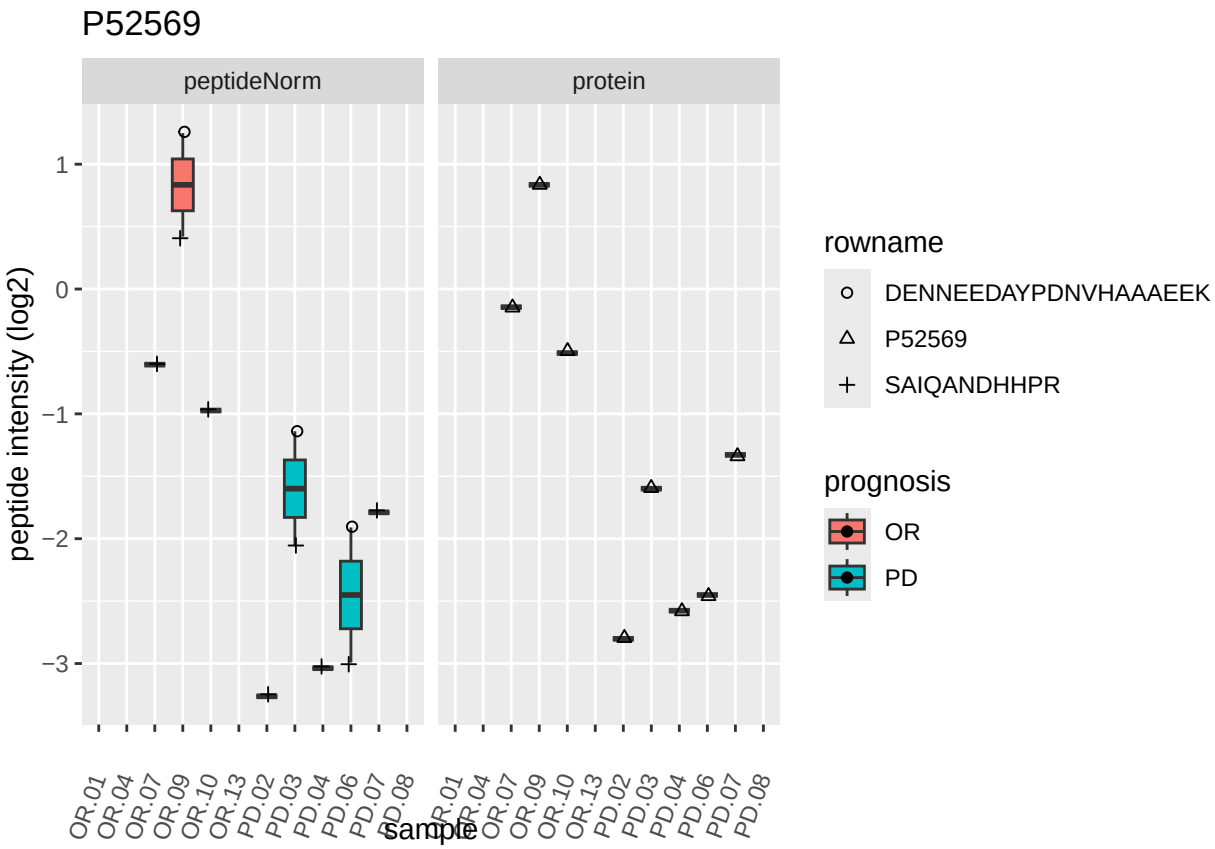

Q15126
peptideNorm
protein
value
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

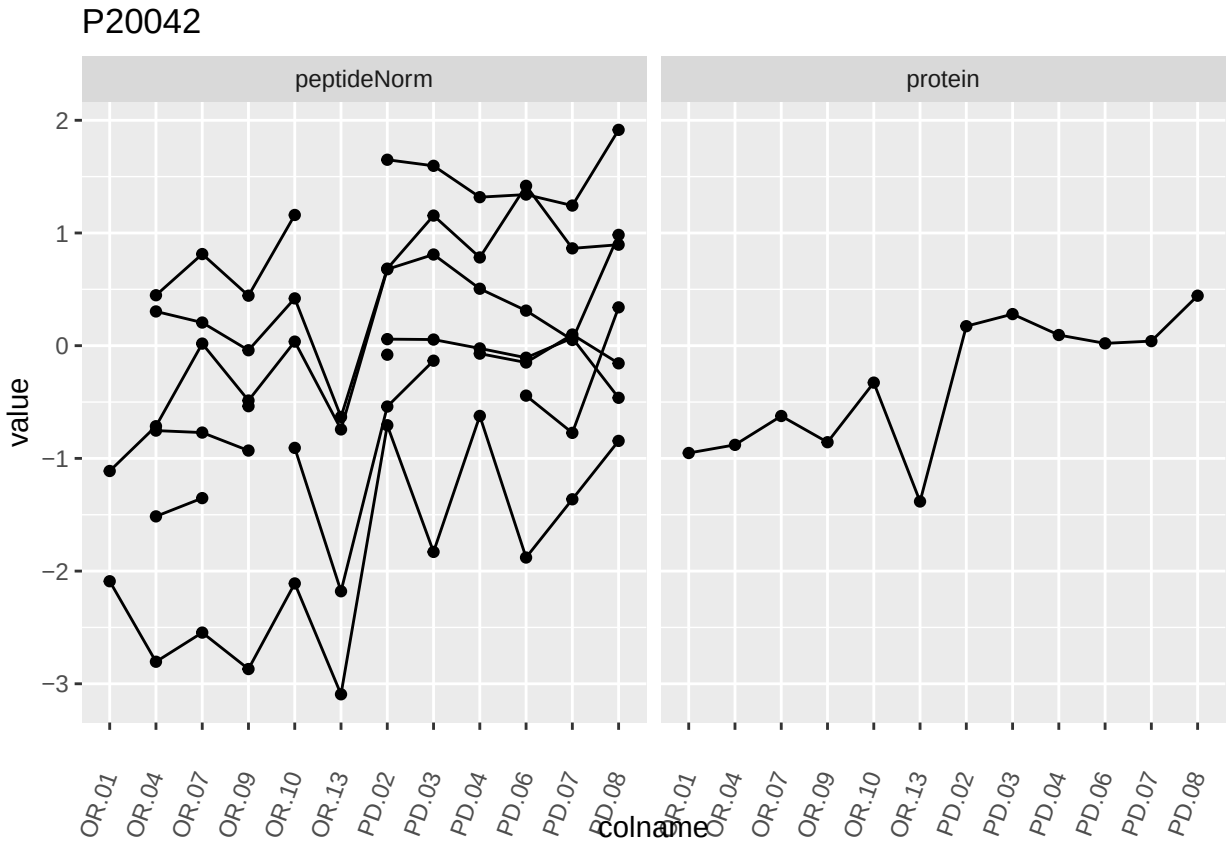

Q15126
peptideNorm
protein
peptide intensity (log2)
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
Sample
prognosis
OR
PD
rowname
DFVTEALQSR
EAYGAVTQTVR
LEEQLENLIEFIR
LLDTSTYK
LVLLFSGK
Q15126
VSDIQWFR
VVALEQSR

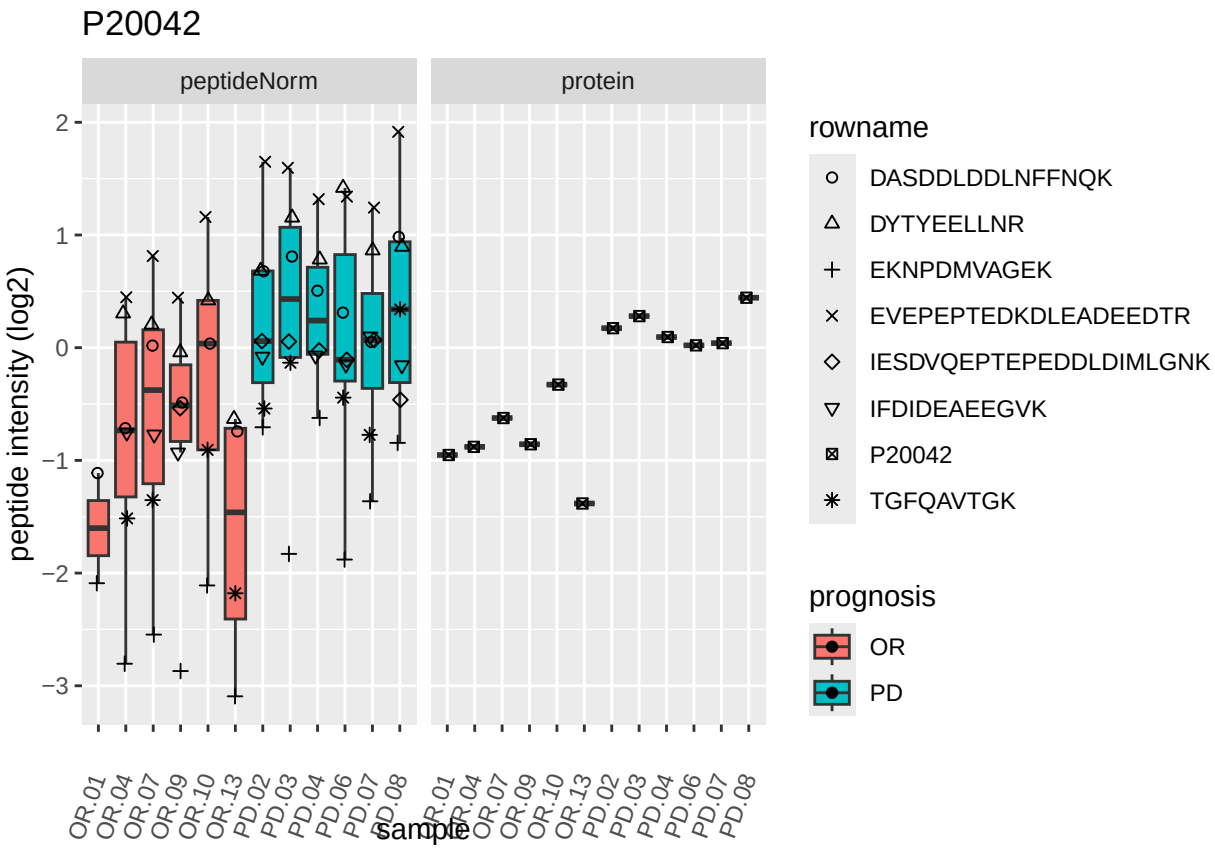

P43490
peptideNorm
protein
value
1
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

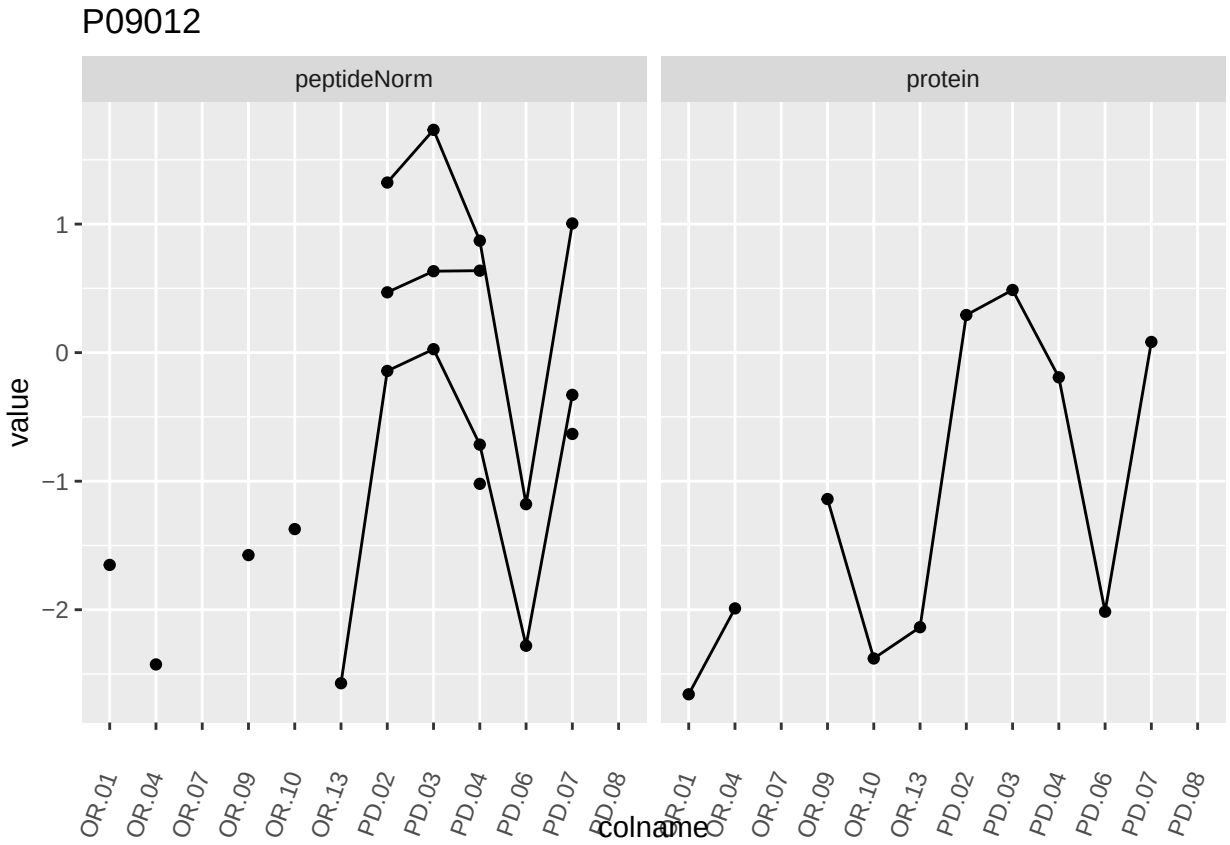

P43490
peptideNorm
protein
peptide intensity (log2)
1
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
rowname
DVYKEHFQDDVFNEK
GVSSQETAGIGASAHLVNFK
GWNYILEK
MNPAAEAEFNILLATDSYK
P43490
VIQGDGVDINTLQEIVEGMK
YLLETSGNLDGLEYK
prognosis
OR
PD

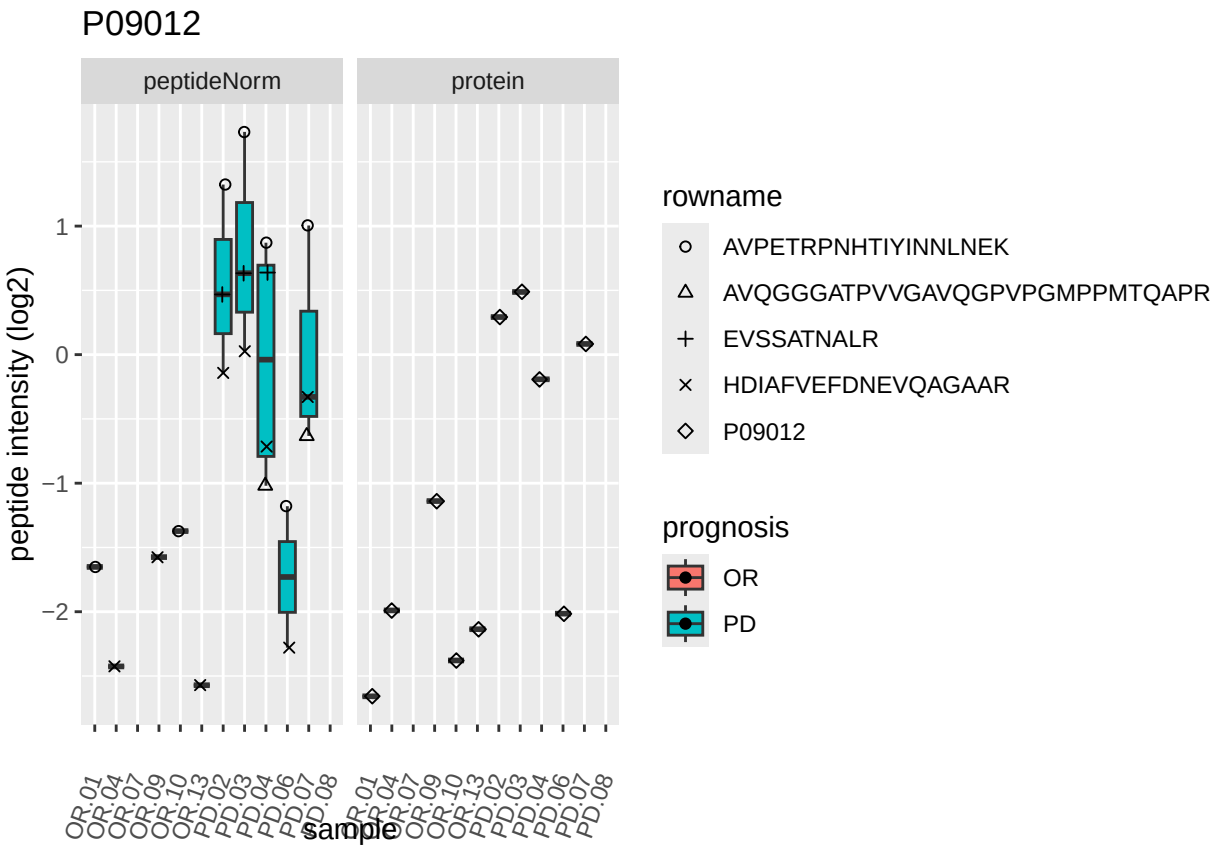

P52569
peptideNorm
protein
value
1
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

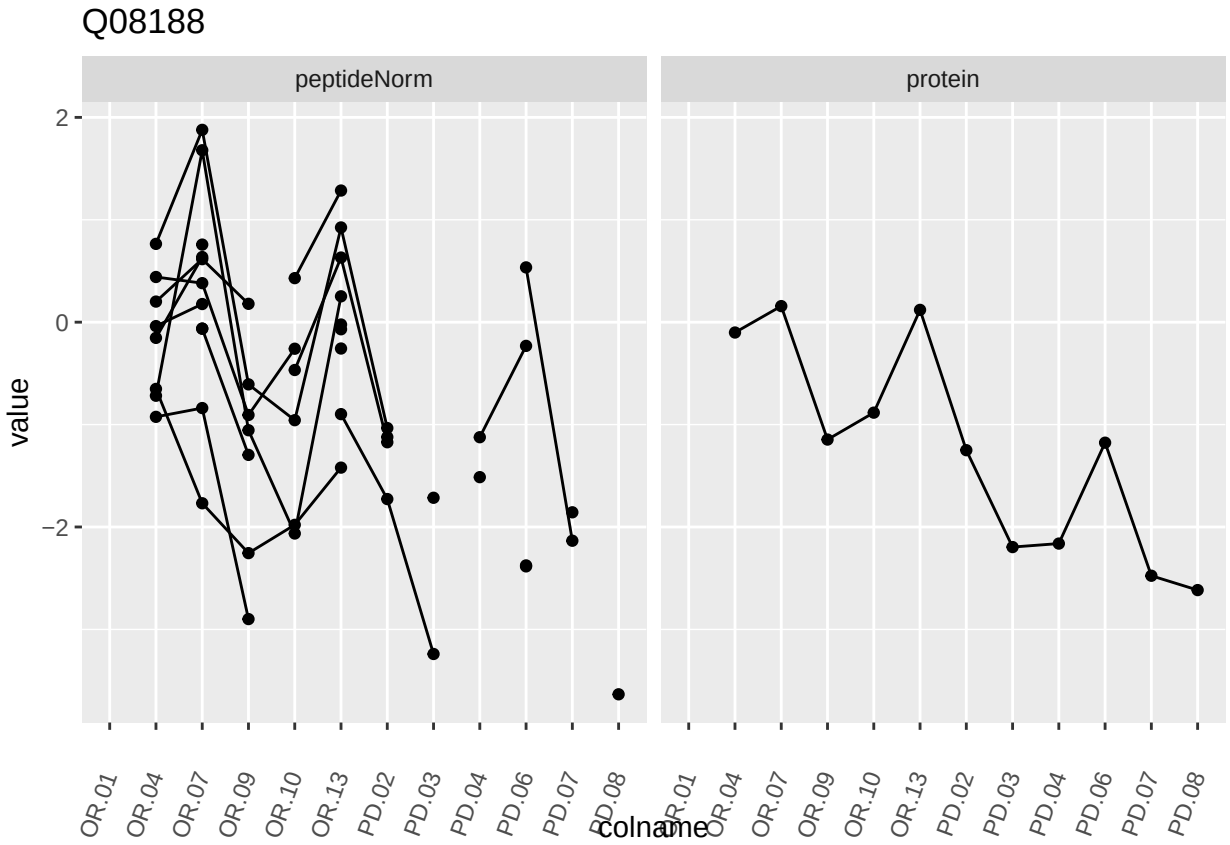

P52569
peptideNorm
protein
peptide intensity (log2)
1
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
rowname
DENNEEDAYPDNVHAAAEEK
P52569
SAIQANDHHPR
prognosis
OR
PD

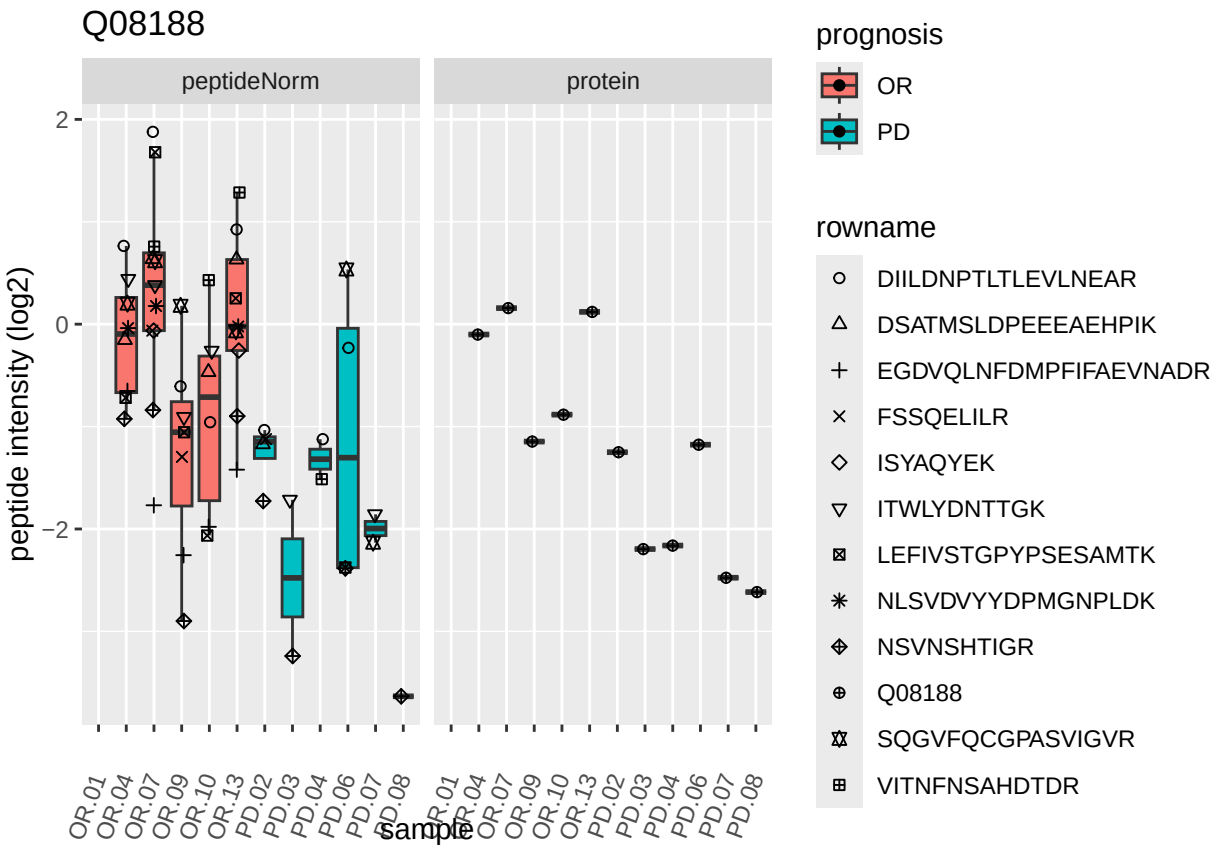

P20042
peptideNorm
protein
value
2
1
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

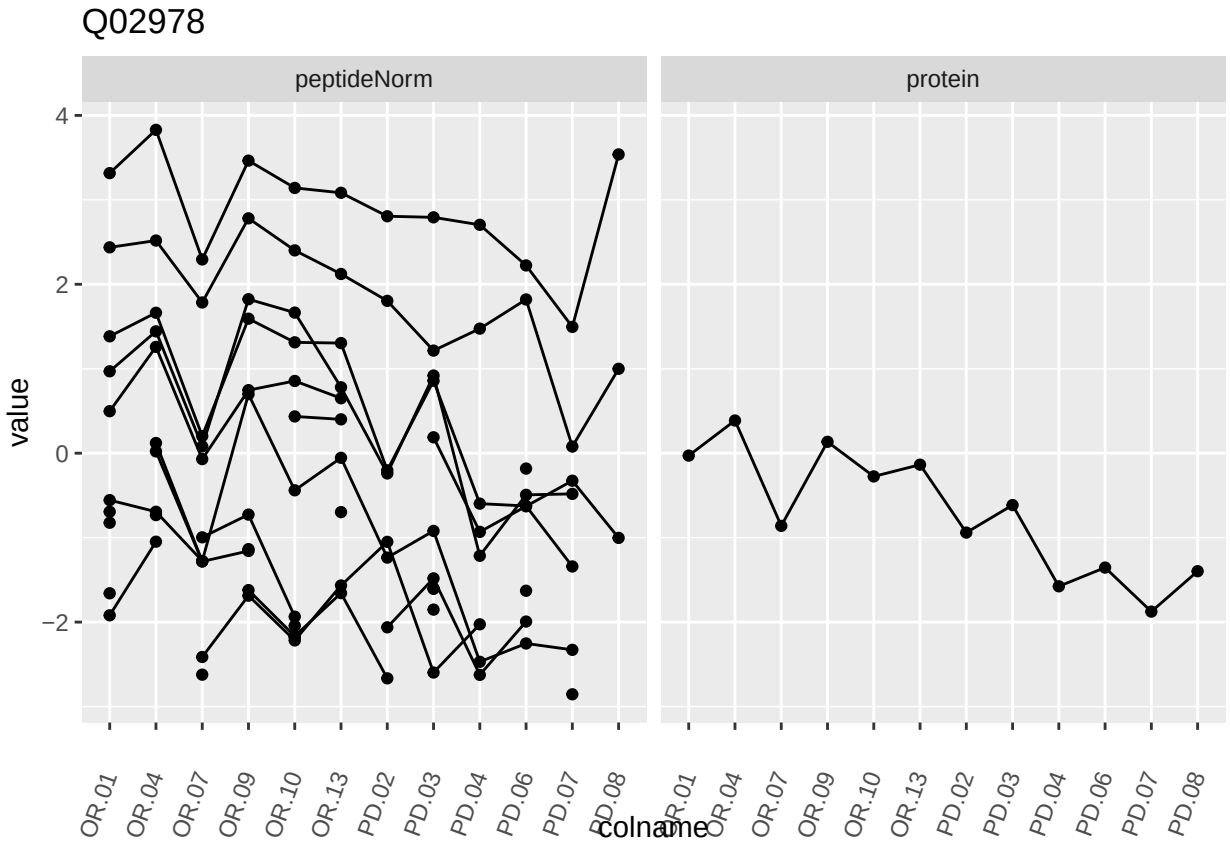

P20042
peptideNorm
protein
peptide intensity (log2)
2
1
0
−1
−2
−3
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
rowname
DASDDLDDLNFFNQK
DYTYEELLNR
EKNPDMVAGEK
EVEPEPTEDKDLEADEEDTR
IESDVQEPTEPEDDLDIMLGNK
IFDIDEAEEGVK
P20042
TGFQAVTGK
prognosis
OR
PD

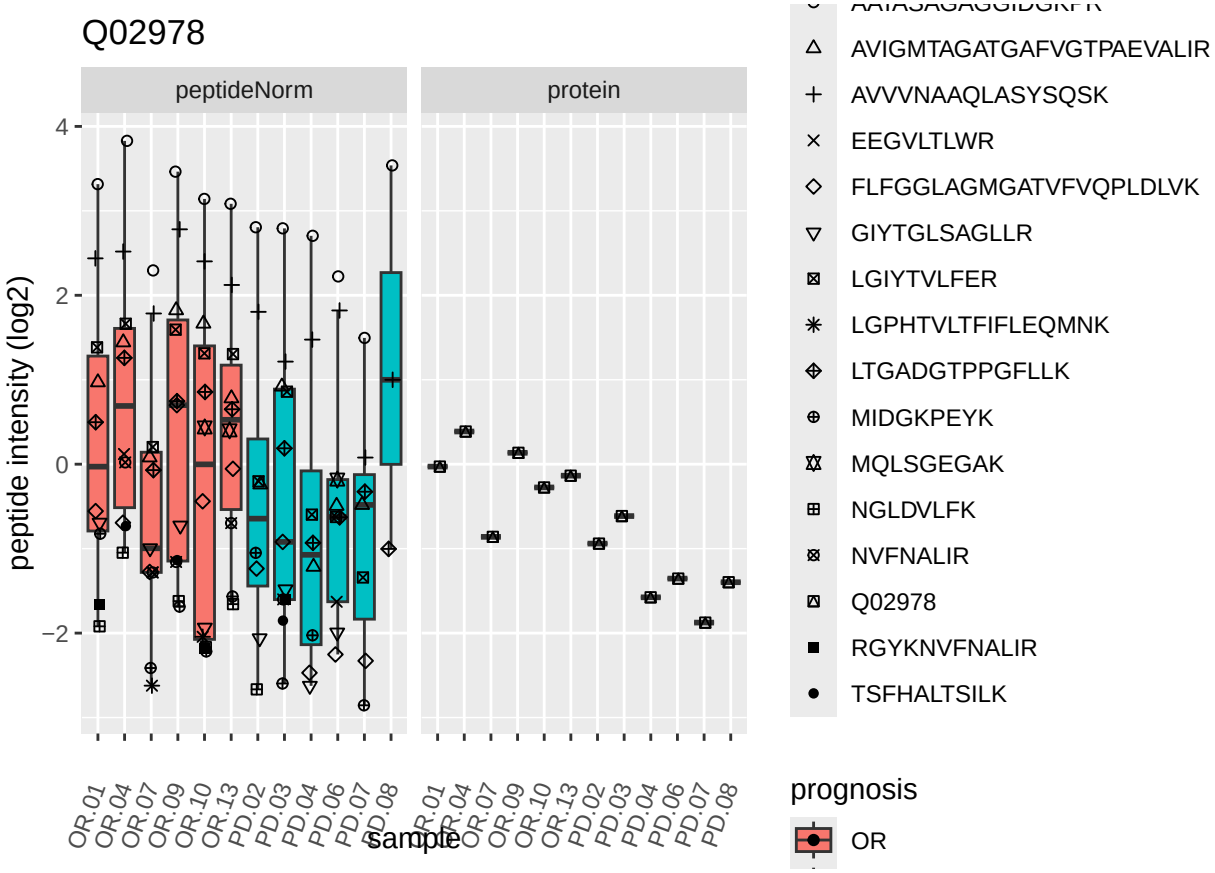

P09012
peptideNorm
protein
value
1
0
−1
−2
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

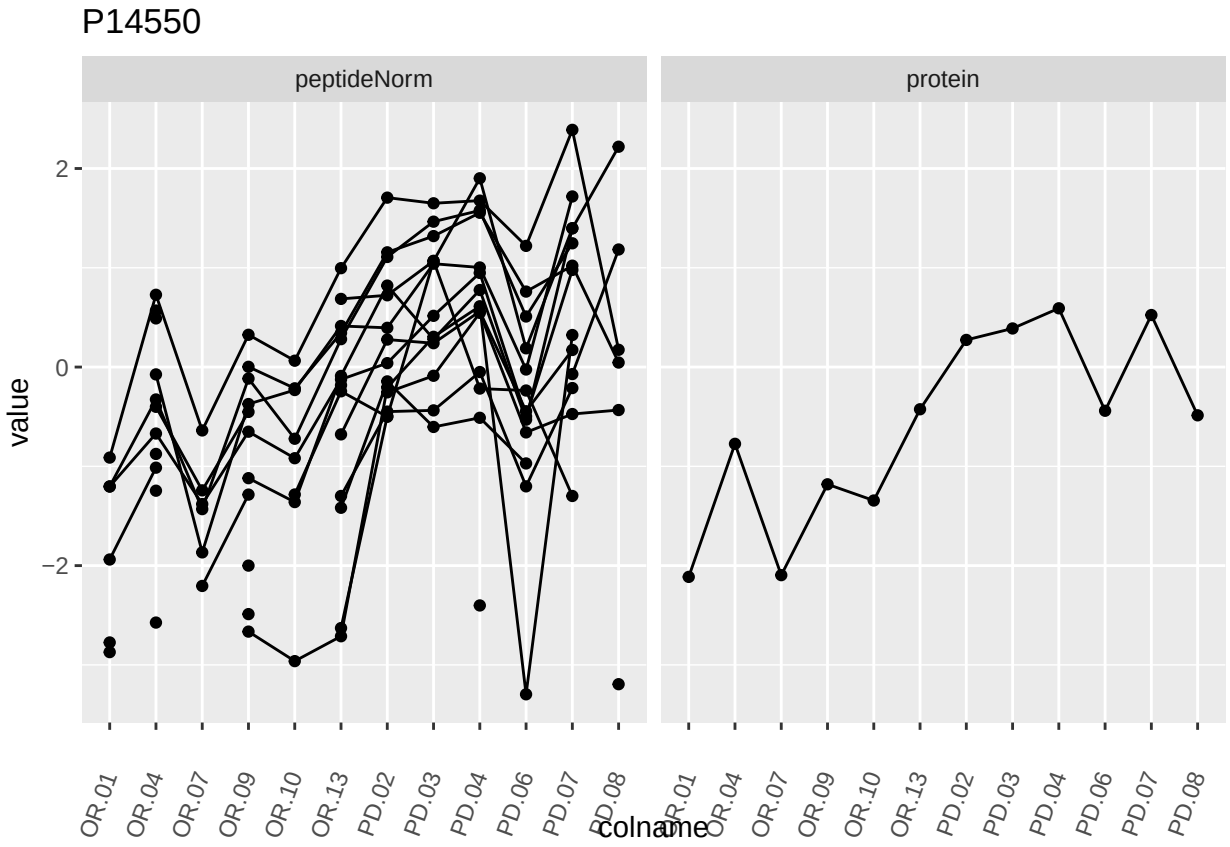

P09012
peptideNorm
protein
peptide intensity (log2)
1
0
−1
−2
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
rowname
AVPETRPNHTIYINNLNEK
AVQGGGATPVVGAVQGPVPGMPPMTQAPR
EVSSATNALR
HDIAFVEFDNEVQAGAAR
P09012
prognosis
OR
PD

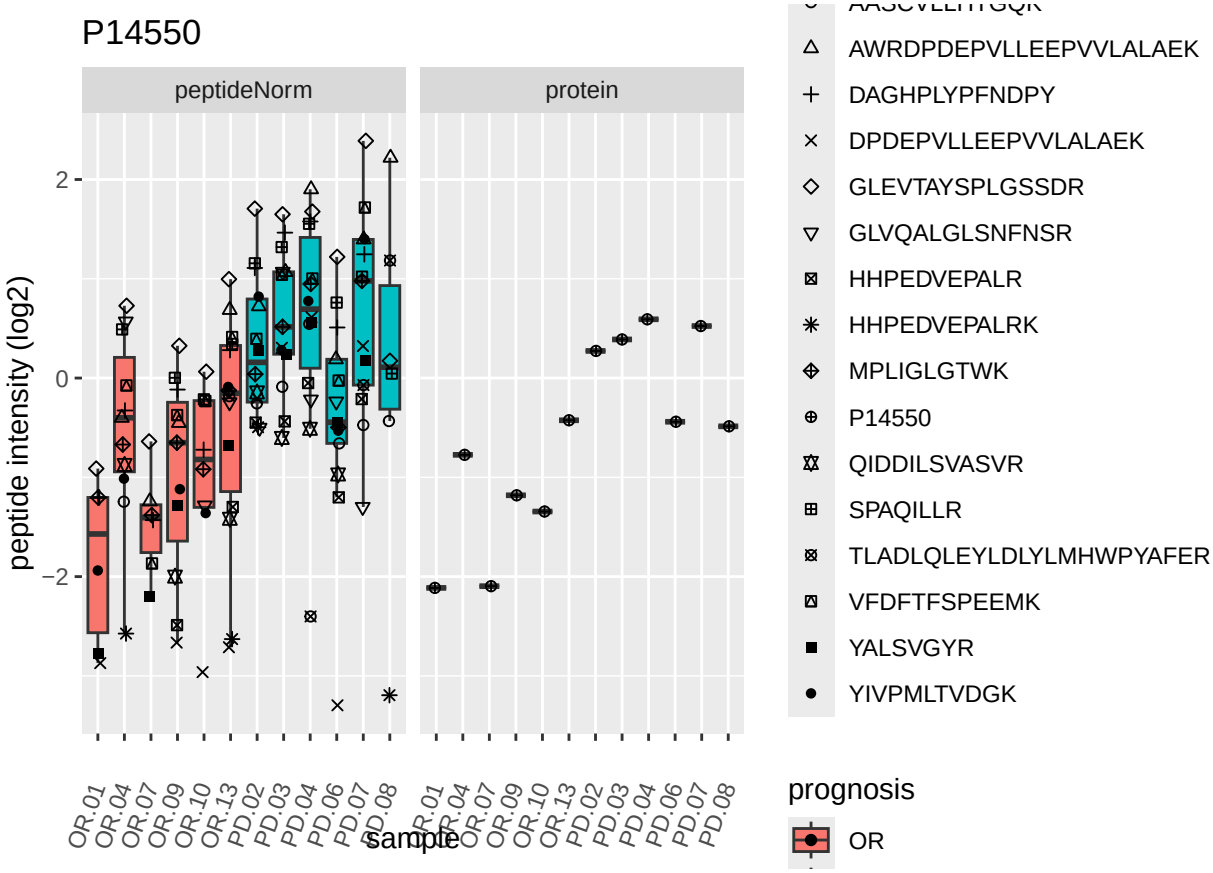

Q08188
peptideNorm
protein
value
2
0
−2
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname

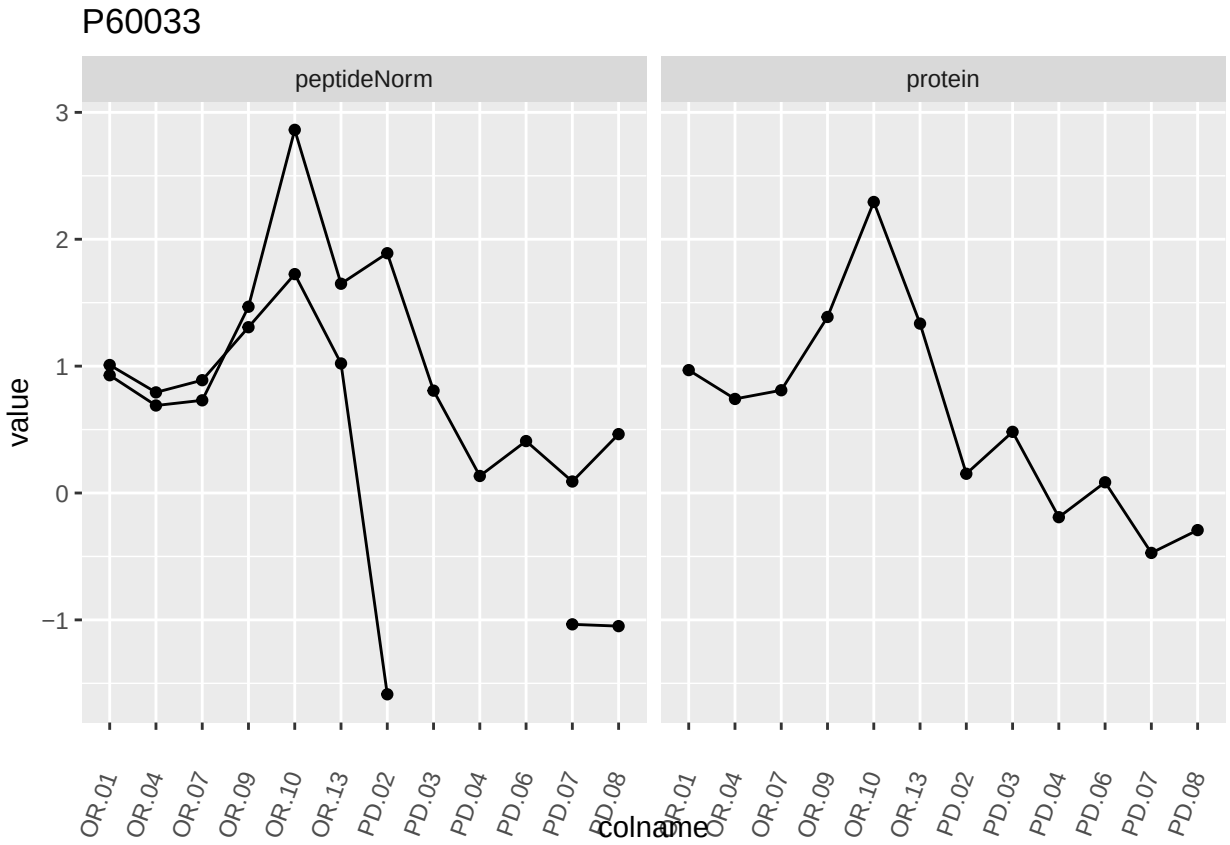

Q08188
peptideNorm
protein
peptide intensity (log2)
2
0
−2
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
prognosis
OR
PD
rowname
DIILDNPTLTLEVLNEAR
DSATMSLDPEEEAEHPIK
EGDVQLNFDMPFIFAEVNADR
FSSQELILR
ISYAQYEK
ITWLYDNTTGK
LEFIVSTGPYPSESAMTK
NLSVDVYYDPMGNPLDK
NSVNSHTIGR
Q08188
SQGVFQCGPASVIGVR
VITNFNSAHDTDR

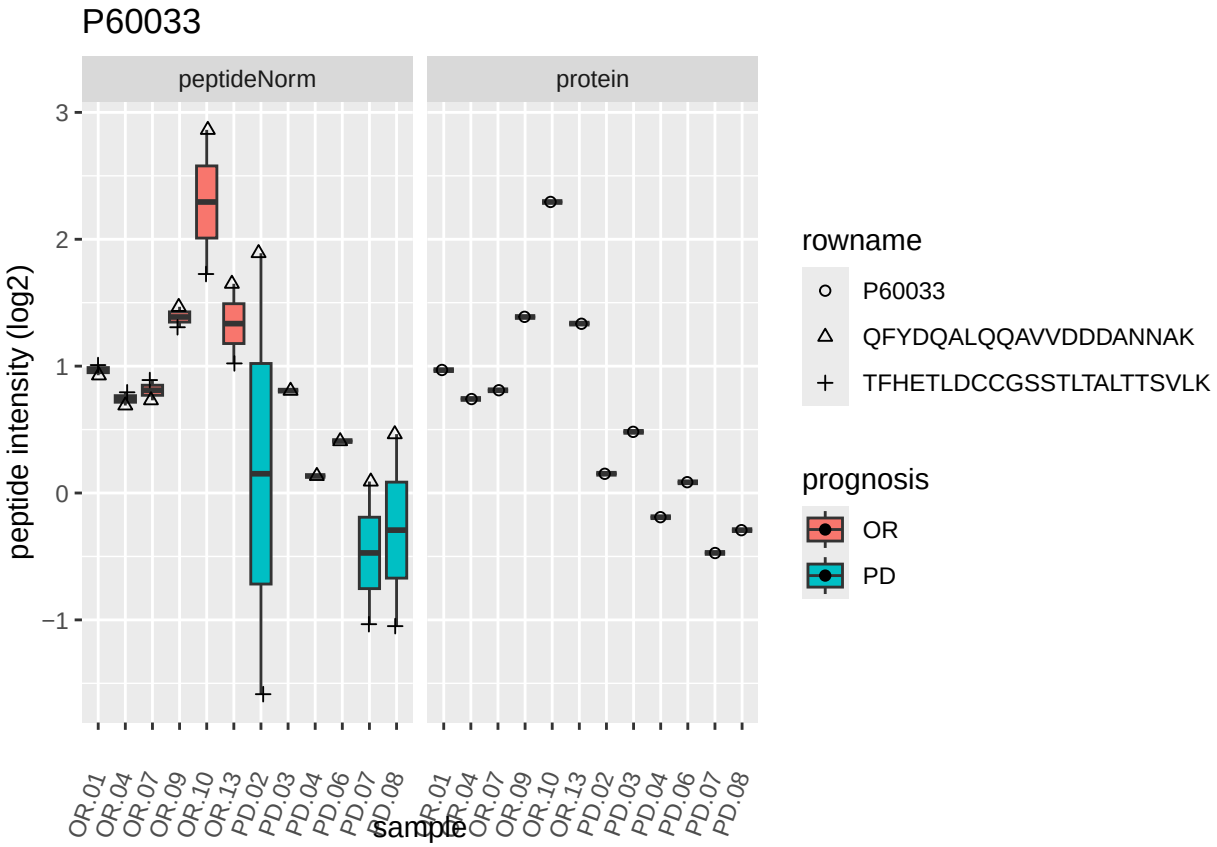

Q02978
peptideNorm
protein
value
4
2
0
−2
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname


Q02978
peptideNorm
protein
peptide intensity (log2)
4
2
0
−2
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
AAIASAGAGGIDGKPK
AVIGMTAGATGAFVGTPAEVALIR
AVVVNAAQLASYSQSK
EEGVLTLWR
FLFGGLAGMGATVFVQPLDLVK
GIYTGLSAGLLR
LGIYTVLFER
LGPHTVLTFIFLEQMNK
LTGADGTPPGFLLK
MIDGKPEYK
MQLSGEGAK
NGLDVLFK
NVFNALIR
Q02978
RGYKNVFNALIR
TSFHALTSILK
prognosis
OR


P14550
peptideNorm
protein
value
2
0
−2
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname


P14550
peptideNorm
protein
peptide intensity (log2)
2
0
−2
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
sample
AWRDPDEPVLLEEPVVLALAEK
DAGHPLYPFNDPY
DPDEPVLLEEPVVLALAEK
GLEVTAYSPLGSSDR
GLVQALGLSNFNSR
HHPEDVEPALR
HHPEDVEPALRK
MPLIGLGTWK
P14550
QIDDILSVASVR
SPAQILLR
TLADLQLEYLDLYLMHWPYAFER
VFDFTFSPEEMK
YALSVGYR
YIVPMLTVDGK
prognosis
OR


P60033
peptideNorm
protein
value
3
2
1
0
−1
OR.01
OR.04
OR.07
OR.09
OR.10
OR.13
PD.02
PD.03
PD.04
PD.06
PD.07
PD.08
colname



5 Session Info

With respect to reproducibility, it is highly recommended to include a session info in your script so that readers of your output can see your particular setup of R.

```
sessionInfo()
```

```
## R version 4.4.0 RC (2024-04-16 r86468)
## Platform: aarch64-apple-darwin20
## Running under: macOS Big Sur 11.6
##
## Matrix products: default
## BLAS:   /Library/Frameworks/R.framework/Versions/4.4-arm64/Resources/lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/4.4-arm64/Resources/lib/libRlapack.dylib;  LAPACK v
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## time zone: Europe/Brussels
## tzcode source: internal
##
## attached base packages:
## [1] stats4    stats     graphics  grDevices utils     datasets  methods
## [8] base
##
## other attached packages:
```

```
##  [1] ExploreModelMatrix_1.16.0   plotly_4.10.4
##  [3] msqrob2_1.12.0              QFeatures_1.14.2
##  [5] MultiAssayExperiment_1.30.3 SummarizedExperiment_1.34.0
##  [7] Biobase_2.64.0              GenomicRanges_1.56.1
##  [9] GenomeInfoDb_1.40.1         IRanges_2.38.1
## [11] S4Vectors_0.42.1            BiocGenerics_0.50.0
## [13] MatrixGenerics_1.16.0       matrixStats_1.4.1
## [15] limma_3.60.5                lubridate_1.9.3
## [17] forcats_1.0.0               stringr_1.5.1
## [19] dplyr_1.1.4                 purrr_1.0.2
## [21] readr_2.1.5                 tidyr_1.3.1
## [23] tibble_3.2.1                ggplot2_3.5.1
## [25] tidyverse_2.0.0
##
## loaded via a namespace (and not attached):
##  [1] rlang_1.1.4             magrittr_2.0.3          shinydashboard_0.7.2
##  [4] clue_0.3-65             compiler_4.4.0          vctrs_0.6.5
##  [7] reshape2_1.4.4          ProtGenerics_1.36.0     pkgconfig_2.0.3
## [10] crayon_1.5.3            fastmap_1.2.0           XVector_0.44.0
## [13] fontawesome_0.5.2       labeling_0.4.3          utf8_1.2.4
## [16] promises_1.3.0          rmarkdown_2.28          tzdb_0.4.0
## [19] UCSC.utils_1.0.0        nloptr_2.1.1            xfun_0.47
## [22] zlibbioc_1.50.0         cachem_1.1.0            jsonlite_1.8.9
## [25] later_1.3.2             highr_0.11              DelayedArray_0.30.1
## [28] BiocParallel_1.38.0     parallel_4.4.0          cluster_2.1.6
## [31] R6_2.5.1                bslib_0.8.0             stringi_1.8.4
## [34] boot_1.3-31             jquerylib_0.1.4         Rcpp_1.0.13
## [37] knitr_1.48              BiocBaseUtils_1.6.0     httpuv_1.6.15
## [40] Matrix_1.7-0            splines_4.4.0           igraph_2.0.3
## [43] timechange_0.3.0        tidyselect_1.2.1        rstudioapi_0.16.0
## [46] abind_1.4-8             yaml_2.3.10             codetools_0.2-20
## [49] lattice_0.22-6          plyr_1.8.9              shiny_1.9.1
## [52] withr_3.0.1             evaluate_1.0.0          pillar_1.9.0
## [55] DT_0.33                 shinyjs_2.1.0           generics_0.1.3
## [58] hms_1.1.3               munsell_0.5.1           scales_1.3.0
## [61] minqa_1.2.8             xtable_1.8-4            glue_1.8.0
## [64] lazyeval_0.2.2          tools_4.4.0             data.table_1.16.0
## [67] lme4_1.1-35.5           cowplot_1.1.3           grid_4.4.0
## [70] MsCoreUtils_1.16.1      colorspace_2.1-1        nlme_3.1-166
## [73] GenomeInfoDbData_1.2.12 cli_3.6.3               fansi_1.0.6
## [76] S4Arrays_1.4.1          viridisLite_0.4.2       AnnotationFilter_1.28.0
## [79] gtable_0.3.5            rintrojs_0.3.4          sass_0.4.9
## [82] digest_0.6.37           SparseArray_1.4.8       htmlwidgets_1.6.4
## [85] farver_2.1.2            htmltools_0.5.8.1       lifecycle_1.0.4
## [88] httr_1.4.7              mime_0.12               statmod_1.5.0
## [91] MASS_7.3-61
```