NMR-based metabolomic analysis of the dataset MTBLS242: serum samples

Institute for Bioengineering of Catalonia

September 4, 2019

This is an example of using AlpsNMR package on the MTBLS242 dataset structured as a pipeline so that inputs are needed and outputs are obtained in the selected folders. Edit "inputs" to match your parameters and just run the "code to run" for the pipeline execution. However, you can see the vignettes and use the functions as you wish. You can download the MTBLS242 dataset from MetaboLights database: https://www.ebi.ac.uk/metabolights/MTBLS242

```
library(AlpsNMR)
```

```
## Warning: replacing previous import 'ggplot2::last_plot' by
## 'plotly::last_plot' when loading 'rDolphin'
```

```
##
## Attaching package: 'AlpsNMR'
```

```
## The following object is masked from 'package:stats':
##
##     filter
```

Pipeline preparation

To work as in a pipeline manner we need to set an output directory. We can set the number of cores of your computer for parallelization.

```
# Set a folder to keep results
output_dir <- "C:/Users/fmadrid/MTBLS242/results"

# How many cores to use for parallelization
num_workers <- 12
```

Node 1: Load samples

Loads samples from a specified directory into a `nmr_dataset` object. Then we can save the loaded data into the output directory.

Input parameters

```
# Path of NMR samples downloaded from https://www.ebi.ac.uk/metabolights/MTBLS242:
dataset_path_nmr <- "C:/Users/fmadrid/MTBLS242/MTBLS242"

# Files/directories ending in "s" corresponding to the spectra in the dataset:
filename_glob <- "*s"
```

Code to run

```
NMRExperiments <- as.character(fs::dir_ls(dataset_path_nmr, glob = filename_glob))
plan(multiprocess, workers = num_workers)
nmr_dataset <- nmr_read_samples(NMRExperiments)
plan(sequential)
```

Node 2: Append metadata

We now merge the metadata. To do that, you need an Excel file containing a first column called "NMRExperiments" with the name of the imported spectra (it does not have to be the name of the individuals).

Input parameters

```
# Path where metada is contained
excel_file <- "C:/Users/fmadrid/MTBLS242/nmr_dataset_metadata_tidy.xlsx"
```

Code to run

```
nmr_dataset <- nmr_meta_add_tidy_excel(nmr_dataset, excel_file)
```

Node 3: Interpolation

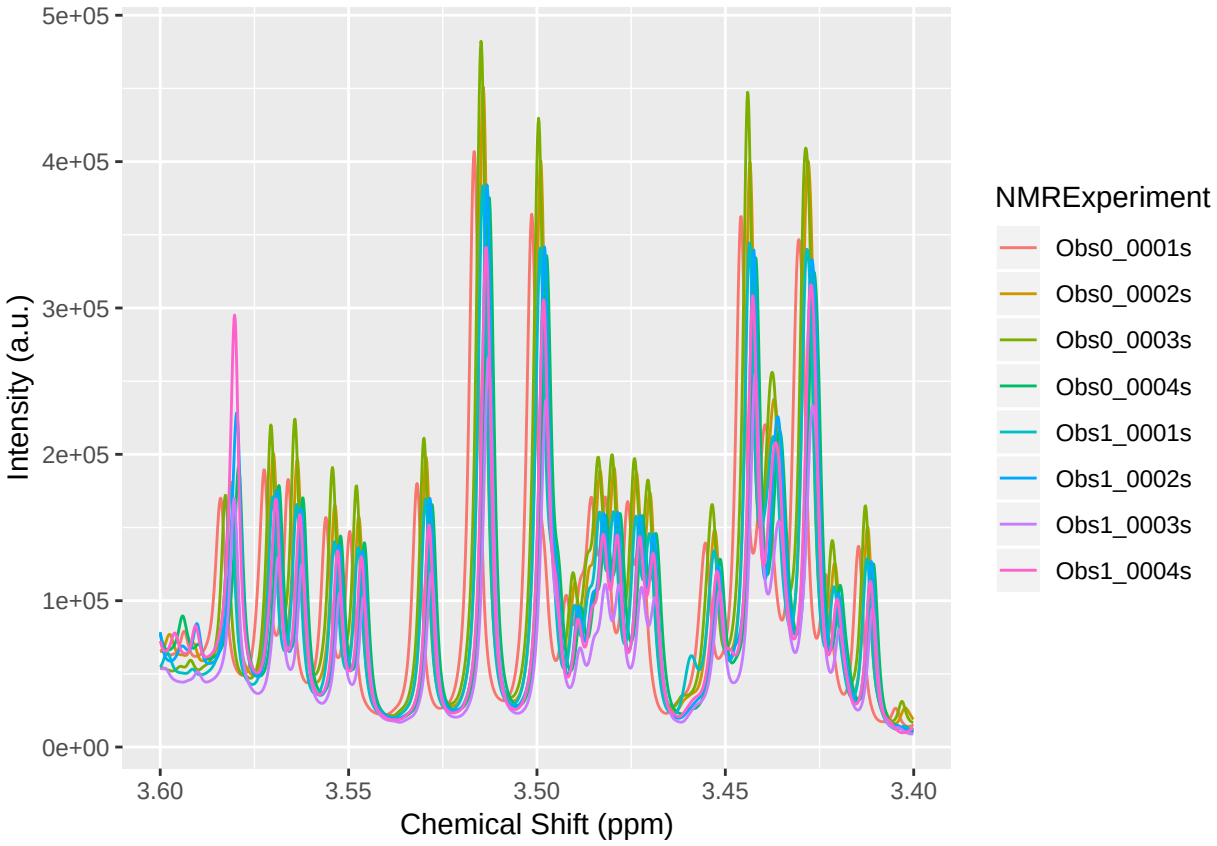
Interpolation is used to unify the ppm axis from all spectra. However, you also can set a range for next steps avoiding noise regions from here. Note that ppm resolution is automatically calculated with the function `nmr_ppm_resolution` in "Code to run".

Input parameters

```
ppm_range_start <- 0.7
ppm_range_end <- 9.5
```

Code to run

```
ppm_resolution <- unlist(nmr_ppm_resolution(nmr_dataset[1]))
axis <- c(min = ppm_range_start, max = ppm_range_end, by = ppm_resolution)
nmr_dataset <- nmr_interpolate_1D(nmr_dataset, axis = axis)

plot(nmr_dataset,
  NMRExperiment = c(
  "Obs0_0001s",
  "Obs0_0002s",
  "Obs0_0003s",
  "Obs0_0004s",
  "Obs1_0001s",
  "Obs1_0002s",
  "Obs1_0003s",
  "Obs1_0004s"),
  chemshift_range = c(3.40, 3.60))
```



Node 4: Region Exclusion

Here it is important to know what type of signals can mask the results due to their intensity or what type of solvent has been used in sample processing since this can create artifacts in the spectra and should be removed. In this case, the biological samples correspond to serum, which contains a lot of water and its signal should be removed from further steps. To do this, we define a vector containing the range (min ppm

value, max ppm value) of the water signal, but other signals can be eliminated, for example: exclude_regions <- list(water = c(4.5, 5.1), methanol = c(3.33, 3.34))

Input parameters

```
exclude_regions <-  list(water = c(4.5, 5.1))
```

Code to run

```
nmr_dataset <- nmr_exclude_region(nmr_dataset, exclude = exclude_regions)
```

Node 5: Initial Outlier Rejection

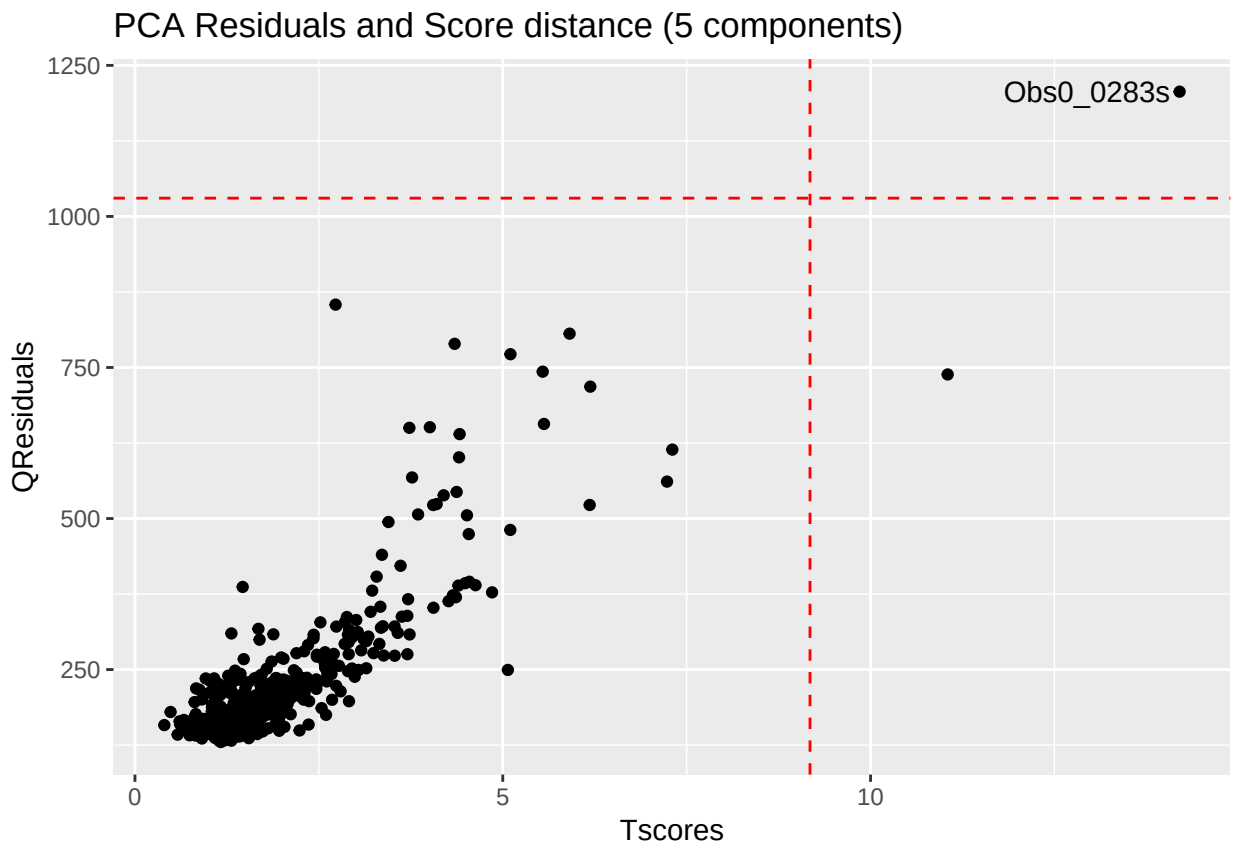
The robust principal component analysis (rPCA) for outlier detection gives an idea of potential outliers. A proposed threshold, based on quantiles, for Q residual and T2 score values, results less sensitive to extreme intensities. Then you choose if any sample should be excluded. The plot below indicated that a sample "Obs0_0283s" is extremely different than the other samples. The function is prepared to annotated samples that are in the top-right corner, exhibiting high differences.

Input parameters

```
# Nothing
```

Code to run

```
pca_outliers <- nmr_pca_outliers_robust(nmr_dataset)
nmr_pca_outliers_plot(nmr_dataset, pca_outliers)
```



Then, if we decide to discard this sample, we just run the function below. Otherwise, just ignore this:

```
nmr_dataset_with_outliers <- nmr_dataset
nmr_dataset <- nmr_pca_outliers_filter(nmr_dataset, pca_outliers)
```

Node 6: Filter samples

Input parameters

The filter node takes care of keeping only some samples. In this case, we want to compare two time points of the MTBLS242 dataset to compare them: "preop" and "3 months after surgery". However, you can filter to keep other conditions kept in the metadata. Some examples: - `Cohort == "A"`: Keeps the A cohort - `TimePoint %in% c("preop", "3 months after surgery")`: Keeps timepoints "preop" and "3 months after surgery" - `Gender == "Female"`: Keeps Female samples - others

```
samples_to_keep_conditions <- 'Timepoint %in% c("preop", "3 months after surgery")'
```

Code to run

```
conditions_expr <- rlang::parse_exprs(samples_to_keep_conditions)
nmr_dataset <- AlpsNMR::filter(nmr_dataset, !!!conditions_expr)
```

Node 7: Peak detection and Alignment

Peak detection is based on a combination of an automated baseline threshold, signal to noise ratio and maximum tolerance. Alignment is based on hierarchical cluster-based peak alignment (CluPA) (Vu et al., 2011).

Input parameters

```
# Leave those as default/recommended for serum.
# Size of peak detection segments
nDivRange_ppm <- 0.1

# Baseline threshold
baselineThresh <- NULL

# Signal to noise ratio
SNR.Th <- 3

# Maximum alignment shift
maxShift_ppm <- 0.0015
```

Code to run

```
scales <- seq(1, 16, 2)
acceptLostPeak <- FALSE

# For parallelization
plan(multiprocess, workers = num_workers)

# Step 1: Peak detection
message("Detecting peaks...")
```

```
## Detecting peaks...
```

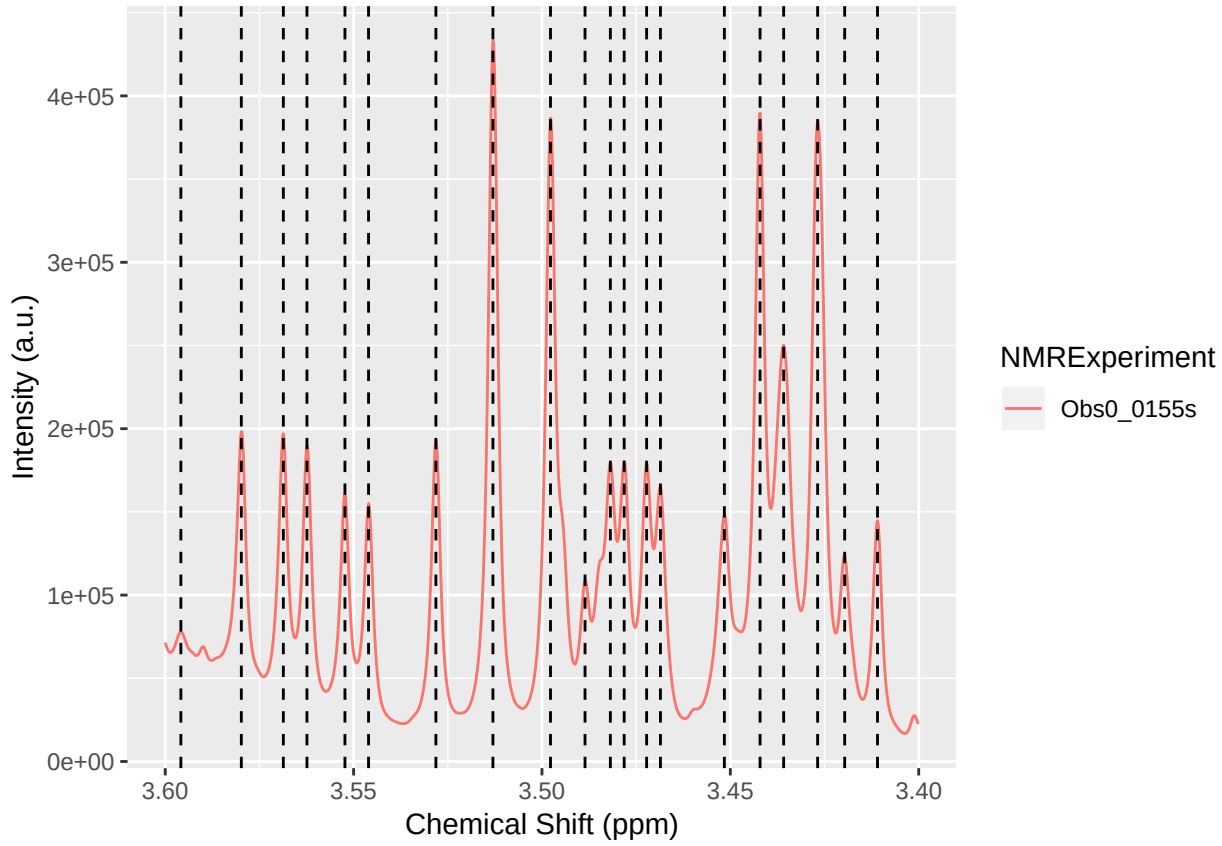
```
peak_data <- nmr_detect_peaks(nmr_dataset,
                              nDivRange_ppm = nDivRange_ppm,
                              scales = scales,
                              baselineThresh = baselineThresh,
                              SNR.Th = SNR.Th)

# Step 2: Finding the reference spectrum for alignment
message("Choosing alignment reference...")
```

```
## Choosing alignment reference...
```

```
NMRExp_ref <- nmr_align_find_ref(nmr_dataset, peak_data)

# Step 3: Alignment
message("Starting alignment...")
```

```
## Starting alignment...
```

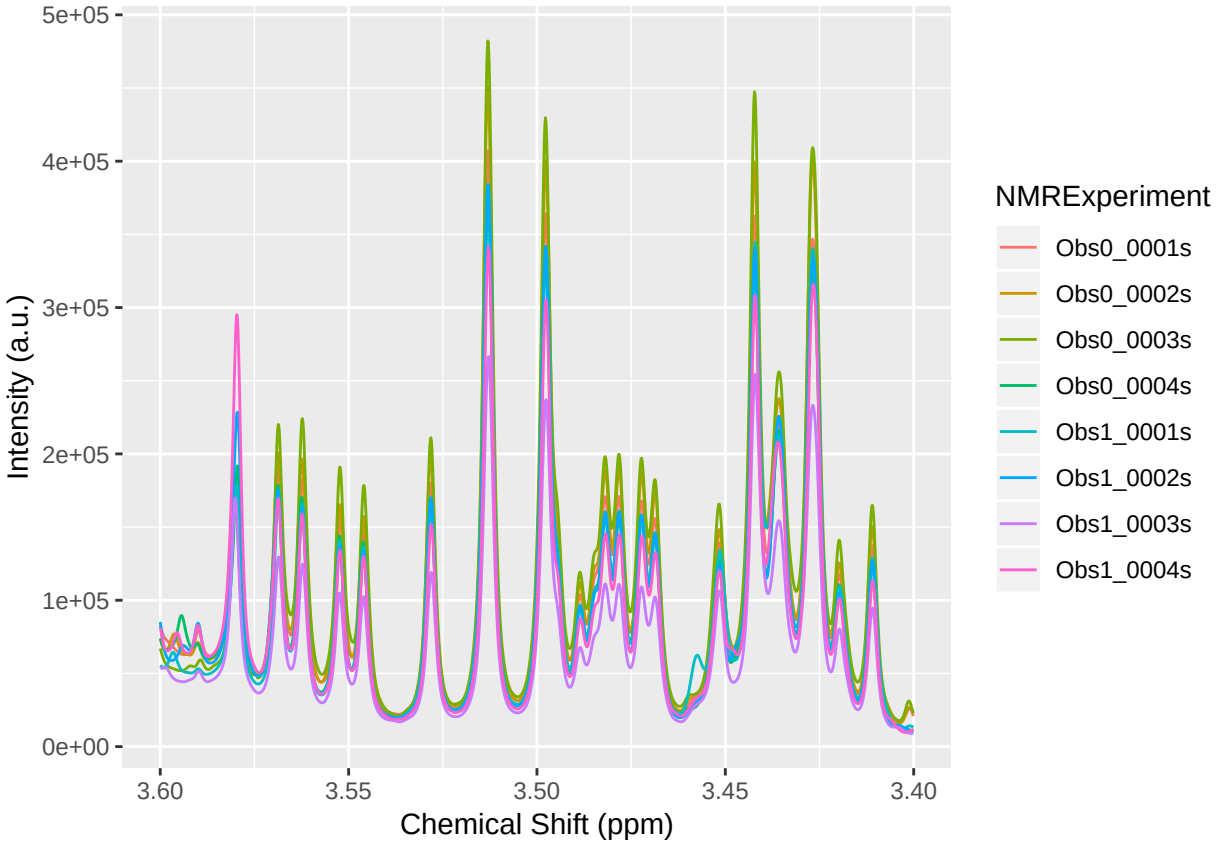
```
nmr_dataset <- nmr_align(nmr_dataset, peak_data,
                         NMRExp_ref = NMRExp_ref,
                         maxShift_ppm = maxShift_ppm,
                         acceptLostPeak = acceptLostPeak)

plan(sequential)

# Plotting results
nmr_detect_peaks_plot(
  nmr_dataset,
  peak_data,
  NMRExperiment = NMRExp_ref,
  chemshift_range = c(3.40, 3.60)
  )
```



we can take a look into the detected peaks. The interactive plot allows you to zoom in in HTML files.

```
plot(nmr_dataset,
  NMRExperiment = c(
  "Obs0_0001s",
  "Obs0_0002s",
  "Obs0_0003s",
  "Obs0_0004s",
  "Obs1_0001s",
```

```
  "Obs1_0002s",
  "Obs1_0003s",
  "Obs1_0004s"),
  chemshift_range = c(3.40, 3.60))
```



Node 8: Normalization

We can normalize the dataset. This is recommended for biosamples, controlling for dilution factors, irregular pipetting, etc. Probabilistic quotient normalization is one of the most used model-based techniques NMR-based metabolomics.

Input parameters

```
# nothing
```

Code to run

```
nmr_dataset <- nmr_normalize(nmr_dataset, method = "pqn")
norm_pqn_diagnostic <- nmr_diagnose(nmr_dataset)
gplt_norm_factor_pqn <- norm_pqn_diagnostic$plot
```

Node 9: Integration

For peak integration, calculation of peak width may be performed automatically (set peak_width_ppm = NULL), from the detected peaks in the reference spectrum (if you wish, you can combine detected peaks other than the reference spectrum, see help), or manually, in which users can select a specific peak width for integrating the detected peaks. This differs than the bucketing approach in which spectra are equally divided into buckets (for example of 0.01 ppm) and this normally leads to a higher number of total variables. this has the inconvenient that several peaks might be split into several parts, lowering the statistical power, and vice-versa, certain overlapping tails might result in false positives because of this noisy parts. However, a good match between them is expected.

Input parameters

```
peak_width_ppm <- NULL
```

Code to run

```
# be carefull, you integrate based on peaks from a unique ref sample
peak_data_integ <- dplyr::filter(peak_data, NMRExperiment == !!NMRExp_ref)

nmr_peak_table <- nmr_integrate_peak_positions(
  samples = nmr_dataset,
  peak_pos_ppm = peak_data_integ$ppm,
  peak_width_ppm = peak_width_ppm)
```

```
## calculated width for integration is 0.00427905515639537 ppm
```

```
nmr_peak_table_completed <- get_integration_with_metadata(nmr_peak_table)
```

Node 10: Machine learning

Pairwise multilevel approach takes into consideration variability within the same individual. Therefore, we use the function "rdCV_PLS_RF_ML" to perform a multilevel repeated double cross-validation optimized for unbiased variable selection (MUVR algorithm, see Shi et al., 2018). The double cross-validation procedure comprises an inner "tuning" loop nested within an outer loop aimed at reducing bias resulting from overfitting models to experimental data. Then, autoselected variables are ranked according to their VIP value.

Modeling through the multivariate modelling with minimally biased variable selection (MUVR) algorithm

```
model <- rdCV_PLS_RF(nmr_data(nmr_peak_table),
                     Y = nmr_peak_table_completed$Timepoint)
```

```
## Warning: package 'pROC' was built under R version 3.5.3
```

```
## Type 'citation("pROC")' for a citation.

##
## Attaching package: 'pROC'

## The following objects are masked from 'package:stats':
##
##     cov, smooth, var

##
## Missing ID -> Assume all unique (i.e. sample independence)
## Y is factor -> Classification (2 classes)
##  Elapsed time 0.077 mins
```

```
message("AUC model is ", model$auc[[2]])
```

```
## AUC model is 0.960264237814033
```

This is the misclassification plot from multilevel PLS-DA model between "preop" and "3 months after surgery". Misclassification plot shows each sample predicted as "preop" and "3 months after surgery". The first half of the horizontal axis represents "preop" actual samples, while the second half represents actual "3 months after surgery" samples. Bold points are the average predicted samples, while grey points display predictions from each iteration. To see the misclassification plot, which gives the information about the actual class and predicted class, run:

Model performance

```
MUVR_model_plot(model)
```

Permutation test

The function permutation_test_model performs the permutation test. Set the number of permutation at "nPerm".

```
permutations = permutation_test_model(model, nPerm = 20)
```

```
##
## "MVObj" permutation 1 of 20
##
## Y is factor -> Classification (2 classes)
##  Elapsed time 0.076 mins
##
## Estimated time left: 1.444 mins
##
##
## "MVObj" permutation 2 of 20
##
## Y is factor -> Classification (2 classes)
##  Elapsed time 0.063 mins
##
## Estimated time left: 1.251 mins
##
##
```

```
## "MVObj" permutation 3 of 20
##
## Y is factor -> Classification (2 classes)
##  Elapsed time 0.06383333 mins
##
## Estimated time left: 1.149389 mins
##
##
## "MVObj" permutation 4 of 20
##
## Y is factor -> Classification (2 classes)
##  Elapsed time 0.06133333 mins
##
## Estimated time left: 1.056667 mins
##
##
## "MVObj" permutation 5 of 20
##
## Y is factor -> Classification (2 classes)
##  Elapsed time 0.06233333 mins
##
## Estimated time left: 0.9795 mins
##
##
## "MVObj" permutation 6 of 20
##
## Y is factor -> Classification (2 classes)
##  Elapsed time 0.06333333 mins
##
## Estimated time left: 0.9096111 mins
##
##
## "MVObj" permutation 7 of 20
##
## Y is factor -> Classification (2 classes)
##  Elapsed time 0.06483333 mins
##
## Estimated time left: 0.844381 mins
##
##
## "MVObj" permutation 8 of 20
##
## Y is factor -> Classification (2 classes)
##  Elapsed time 0.06116667 mins
##
## Estimated time left: 0.774 mins
##
##
## "MVObj" permutation 9 of 20
##
## Y is factor -> Classification (2 classes)
##  Elapsed time 0.0655 mins
##
## Estimated time left: 0.7107222 mins
```

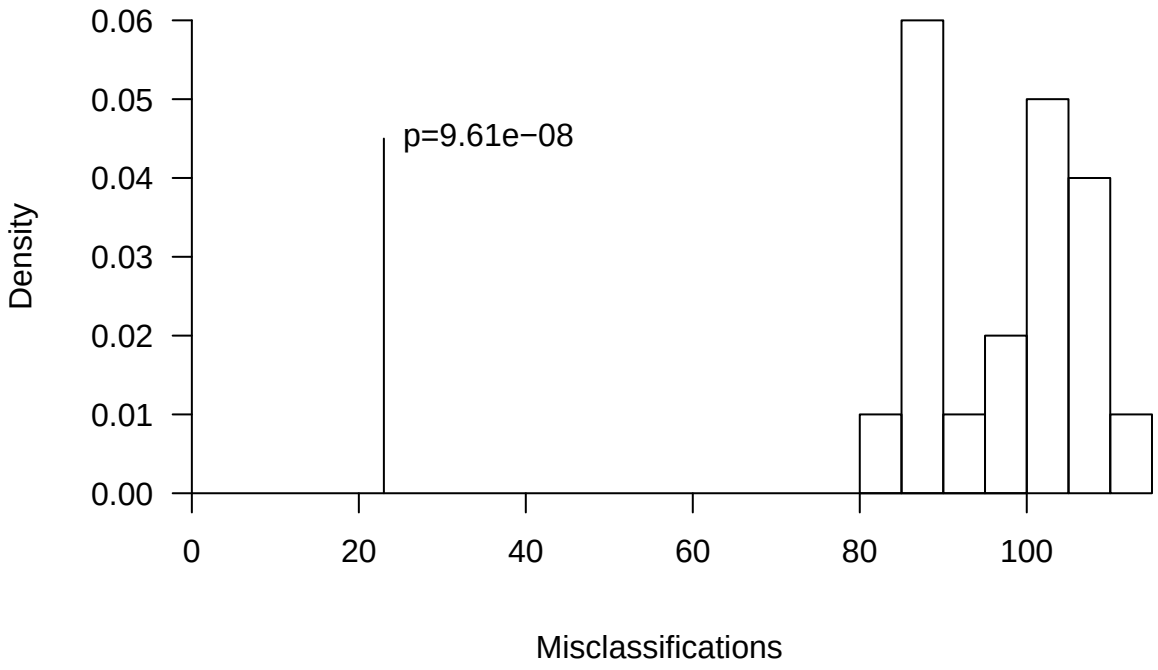
```
##
##
## "MVObj" permutation 10 of 20
##
## Y is factor -> Classification (2 classes)
##  Elapsed time 0.0595 mins
##
## Estimated time left: 0.641 mins
##
##
## "MVObj" permutation 11 of 20
##
## Y is factor -> Classification (2 classes)
##  Elapsed time 0.06333333 mins
##
## Estimated time left: 0.5762727 mins
##
##
## "MVObj" permutation 12 of 20
##
## Y is factor -> Classification (2 classes)
##  Elapsed time 0.06483333 mins
##
## Estimated time left: 0.5127778 mins
##
##
## "MVObj" permutation 13 of 20
##
## Y is factor -> Classification (2 classes)
##  Elapsed time 0.062 mins
##
## Estimated time left: 0.4475513 mins
##
##
## "MVObj" permutation 14 of 20
##
## Y is factor -> Classification (2 classes)
##  Elapsed time 0.06166667 mins
##
## Estimated time left: 0.3826429 mins
##
##
## "MVObj" permutation 15 of 20
##
## Y is factor -> Classification (2 classes)
##  Elapsed time 0.06383333 mins
##
## Estimated time left: 0.3188889 mins
##
##
## "MVObj" permutation 16 of 20
##
## Y is factor -> Classification (2 classes)
##  Elapsed time 0.061 mins
```

```
##
## Estimated time left: 0.2544167 mins
##
##
## "MVObj" permutation 17 of 20
##
## Y is factor -> Classification (2 classes)
##  Elapsed time 0.06433333 mins
##
## Estimated time left: 0.1909412 mins
##
##
## "MVObj" permutation 18 of 20
##
## Y is factor -> Classification (2 classes)
##  Elapsed time 0.0635 mins
##
## Estimated time left: 0.1272778 mins
##
##
## "MVObj" permutation 19 of 20
##
## Y is factor -> Classification (2 classes)
##  Elapsed time 0.06333333 mins
##
## Estimated time left: 0.06362281 mins
##
##
## "MVObj" permutation 20 of 20
##
## Y is factor -> Classification (2 classes)
##  Elapsed time 0.06583333 mins
##
## Estimated time left: 0 mins
```

Permutation test plot

Permutation test plot for multilevel PLS-DA model gives a p-value for the model performance. The actual misclassification departs from the null hypothesis distribution.

```
permutation_test_plot(model, permutations)
```



Autoselected features

Then, we can extract VIP values from the autoselected features.

```
VIPs= model_VIP(model)
VIPs
```

```
##            order       name    rank
## ppm_3.5798     1 ppm_3.5798   1.515
## ppm_1.2120     2 ppm_1.2120   4.125
## ppm_1.3515     3 ppm_1.3515   5.740
## ppm_1.0136     4 ppm_1.0136   5.905
## ppm_1.0639     5 ppm_1.0639   6.410
## ppm_4.1352     6 ppm_4.1352   6.465
## ppm_2.2463     7 ppm_2.2463   7.855
## ppm_4.1235     8 ppm_4.1235   9.780
## ppm_1.0521     9 ppm_1.0521  10.155
## ppm_1.0018    10 ppm_1.0018  11.850
## ppm_4.1119    11 ppm_4.1119  14.605
## ppm_1.2217    12 ppm_1.2217  15.185
## ppm_1.3399    13 ppm_1.3399  15.715
## ppm_1.1835    14 ppm_1.1835  17.175
## ppm_1.1938    15 ppm_1.1938  18.250
## ppm_3.0602    16 ppm_3.0602  19.810
## ppm_4.1468    17 ppm_4.1468  20.075
```

```
## ppm_1.2002    18 ppm_1.2002  24.385
## ppm_0.9881    19 ppm_0.9881  27.215
## ppm_3.6628    20 ppm_3.6628  27.565
## ppm_0.8461    21 ppm_0.8461  28.165
## ppm_1.2926    22 ppm_1.2926  31.175
## ppm_2.1518    23 ppm_2.1518  35.625
## ppm_2.2501    24 ppm_2.2501  36.240
## ppm_3.2311    25 ppm_3.2311  36.515
## ppm_5.2826    26 ppm_5.2826  36.690
## ppm_0.9775    27 ppm_0.9775  41.075
## ppm_1.2653    28 ppm_1.2653  43.485
## ppm_3.8940    29 ppm_3.8940  43.510
## ppm_1.5908    30 ppm_1.5908  45.005
## ppm_0.9670    31 ppm_0.9670  45.310
## ppm_1.5995    32 ppm_1.5995  46.975
## ppm_1.3199    33 ppm_1.3199  48.280
## ppm_2.0632    34 ppm_2.0632  50.705
## ppm_5.3228    35 ppm_5.3228  54.490
## ppm_3.3799    36 ppm_3.3799  54.795
## ppm_0.9422    37 ppm_0.9422  55.490
## ppm_1.5039    38 ppm_1.5039  56.265
## ppm_2.2611    39 ppm_2.2611  56.405
## ppm_2.2353    40 ppm_2.2353  57.250
## ppm_2.2324    41 ppm_2.2324  57.365
## ppm_1.5792    42 ppm_1.5792  58.470
## ppm_2.0569    43 ppm_2.0569  59.215
## ppm_2.0970    44 ppm_2.0970  63.265
## ppm_3.4782    45 ppm_3.4782  65.565
## ppm_0.9546    46 ppm_0.9546  65.665
## ppm_3.2807    47 ppm_3.2807  73.330
## ppm_2.2350    48 ppm_2.2350  74.530
## ppm_3.8118    49 ppm_3.8118  74.625
## ppm_3.2673    50 ppm_3.2673  79.705
## ppm_3.4722    51 ppm_3.4722  82.270
## ppm_0.8955    52 ppm_0.8955  87.950
## ppm_1.9937    53 ppm_1.9937  88.220
## ppm_3.6894    54 ppm_3.6894  89.375
## ppm_2.7630    55 ppm_2.7630  89.515
## ppm_1.4916    56 ppm_1.4916  89.545
## ppm_2.0172    57 ppm_2.0172  90.060
## ppm_3.9276    58 ppm_3.9276  90.490
## ppm_2.0045    59 ppm_2.0045  92.685
## ppm_3.4109    60 ppm_3.4109  93.815
## ppm_2.0858    61 ppm_2.0858  95.940
## ppm_3.4977    62 ppm_3.4977  96.325
## ppm_0.9071    63 ppm_0.9071  96.795
## ppm_3.4268    64 ppm_3.4268  97.070
## ppm_3.2520    65 ppm_3.2520  97.885
## ppm_0.8621    66 ppm_0.8621  98.745
## ppm_3.9072    67 ppm_3.9072  98.900
## ppm_3.5130    68 ppm_3.5130 102.005
## ppm_3.7528    69 ppm_3.7528 102.595
```

Node 11: Identification

Finally, AlpsNMR allows an identification step, in which we can select between plasma/serum, urine and cell functions giving a ranked dataframe with ppm and proposed candidates by the Human Metabolome Database (http://www.hmdb.ca). However, this needs to be double-checked, as there is overlap between several potential compounds for a given ppm value and different specific metabolite-shifts, this should be carefully taken as a first step in the identification. First, we extract a vector with significant ppm values. Then, we run "nmr_identify_regions_blood". You can set the number of proposed candidates, but in this particular case, we set to 3. Even though, several NAs in the identification corresponded to Supplemental Table 1 in Supplemental Material.

Autoselected features (blood samples)

```
ppm_to_assign <- VIPs %>%
  tidyr::separate(col = name,
                  into = c("x1", "ppms"),
                  sep = "_")

ppm_to_assign <- as.numeric(ppm_to_assign$ppms)
assignation <-
  dplyr::select(nmr_identify_regions_blood(ppm_to_assign,
                                           num_proposed_compounds = 3),
                -Height,
                -n_reported_in_Blood,
                -J_Hz)
```

```
## [1] 3.5798

## your peak at 3.5798 probably corresponds to D-Mannose, Glycerol or L-Threonine

##

## [1] 1.212

## your peak at 1.212 probably corresponds to 3-Hydroxybutyric acid, Isobutyric acid or NA

## [1] 1.3515

## your peak at 1.3515 probably corresponds to NA, NA or NA

## [1] 1.0136

## your peak at 1.0136 probably corresponds to L-Isoleucine, NA or NA

## [1] 1.0639

## your peak at 1.0639 probably corresponds to NA, NA or NA

## [1] 4.1352
```

```
## your peak at 4.1352 probably corresponds to L-Proline, Uridine or NA

## [1] 2.2463

## your peak at 2.2463 probably corresponds to Acetone, NA or NA

## [1] 4.1235

## your peak at 4.1235 probably corresponds to L-Lactic acid, L-Proline or Uridine

## [1] 1.0521

## your peak at 1.0521 probably corresponds to L-Valine, NA or NA

## [1] 1.0018

## your peak at 1.0018 probably corresponds to L-Alpha-aminobutyric acid, L-Isoleucine or L-Valine

## [1] 4.1119

## your peak at 4.1119 probably corresponds to L-Lactic acid, NA or NA

## [1] 1.2217

## your peak at 1.2217 probably corresponds to 3-Hydroxybutyric acid, Isobutyric acid or NA

## [1] 1.3399

## your peak at 1.3399 probably corresponds to L-Lactic acid, L-Threonine or NA

## [1] 1.1835

## your peak at 1.1835 probably corresponds to Ethanol, Isopropyl alcohol or NA

## [1] 1.1938

## your peak at 1.1938 probably corresponds to Ethanol, NA or NA

## [1] 3.0602

## your peak at 3.0602 probably corresponds to Ornithine, NA or NA

## [1] 4.1468

## your peak at 4.1468 probably corresponds to L-Proline, Uridine or NA

## [1] 1.2002
```

```
## your peak at 1.2002 probably corresponds to NA, NA or NA

## [1] 0.9881

## your peak at 0.9881 probably corresponds to L-Alpha-aminobutyric acid, L-Valine or NA

## [1] 3.6628

## your peak at 3.6628 probably corresponds to D-Mannose, Ethanol or Glycerol

## [1] 0.8461

## your peak at 0.8461 probably corresponds to 2-Hydroxy-3-methylbutyric acid, NA or NA

## [1] 1.2926

## your peak at 1.2926 probably corresponds to NA, NA or NA

## [1] 2.1518

## your peak at 2.1518 probably corresponds to L-Acetylcarnitine, L-Glutamine or NA

## [1] 2.2501

## your peak at 2.2501 probably corresponds to NA, NA or NA

## [1] 3.2311

## your peak at 3.2311 probably corresponds to L-Arginine, L-Histidine or NA

## [1] 5.2826

## your peak at 5.2826 probably corresponds to NA, NA or NA

## [1] 0.9775

## your peak at 0.9775 probably corresponds to 2-Hydroxy-3-methylbutyric acid, L-Leucine or L-Valine

## [1] 1.2653

## your peak at 1.2653 probably corresponds to 3-Hydroxyisovaleric acid, L-Isoleucine or NA

## [1] 3.894

## your peak at 3.894 probably corresponds to Betaine, D-Glucose or D-Mannose

## [1] 1.5908
```

```
## your peak at 1.5908 probably corresponds to NA, NA or NA

## [1] 0.967

## your peak at 0.967 probably corresponds to 2-Hydroxy-3-methylbutyric acid, L-Leucine or NA

## [1] 1.5995

## your peak at 1.5995 probably corresponds to NA, NA or NA

## [1] 1.3199

## your peak at 1.3199 probably corresponds to L-Threonine, NA or NA

## [1] 2.0632

## your peak at 2.0632 probably corresponds to Acetylglycine, Isovaleric acid or L-Glutamic acid

## [1] 5.3228

## your peak at 5.3228 probably corresponds to NA, NA or NA

## [1] 3.3799

## your peak at 3.3799 probably corresponds to D-Mannose, NA or NA

## [1] 0.9422

## your peak at 0.9422 probably corresponds to Ketoleucine, L-Isoleucine or NA

## [1] 1.5039

## your peak at 1.5039 probably corresponds to NA, NA or NA

## [1] 2.2611

## your peak at 2.2611 probably corresponds to L-Valine, NA or NA

## [1] 2.2353

## your peak at 2.2353 probably corresponds to Acetone, NA or NA

## [1] 2.2324

## your peak at 2.2324 probably corresponds to Acetone, NA or NA

## [1] 1.5792
```

```
## your peak at 1.5792 probably corresponds to NA, NA or NA

## [1] 2.0569

## your peak at 2.0569 probably corresponds to Acetylglycine, Isovaleric acid or L-Glutamic acid

## [1] 2.097

## your peak at 2.097 probably corresponds to Ketoleucine, NA or NA

## [1] 3.4782

## your peak at 3.4782 probably corresponds to D-Glucose, Sucrose or NA

## [1] 0.9546

## your peak at 0.9546 probably corresponds to 2-Hydroxy-3-methylbutyric acid, Ketoleucine or L-Isoleuc

## [1] 3.2807

## your peak at 3.2807 probably corresponds to Myoinositol, NA or NA

## [1] 2.235

## your peak at 2.235 probably corresponds to Acetone, NA or NA

## [1] 3.8118

## your peak at 3.8118 probably corresponds to D-Mannose, Uridine or NA

## [1] 3.2673

## your peak at 3.2673 probably corresponds to 3-Methylhistidine, Betaine or NA

## [1] 3.4722

## your peak at 3.4722 probably corresponds to D-Glucose, Sucrose or NA

## [1] 0.8955

## your peak at 0.8955 probably corresponds to NA, NA or NA

## [1] 1.9937

## your peak at 1.9937 probably corresponds to L-Isoleucine, L-Proline or NA

## [1] 3.6894
```

```
## your peak at 3.6894 probably corresponds to 1-Methylhistidine, L-Isoleucine or Sucrose

## [1] 2.763

## your peak at 2.763 probably corresponds to NA, NA or NA

## [1] 1.4916

## your peak at 1.4916 probably corresponds to L-Alanine, NA or NA

## [1] 2.0172

## your peak at 2.0172 probably corresponds to 2-Hydroxy-3-methylbutyric acid, L-Proline or NA

## [1] 3.9276

## your peak at 3.9276 probably corresponds to Creatine, D-Mannose or Uridine

## [1] 2.0045

## your peak at 2.0045 probably corresponds to L-Proline, NA or NA

## [1] 3.4109

## your peak at 3.4109 probably corresponds to D-Glucose, L-Carnitine or L-Proline

## [1] 2.0858

## your peak at 2.0858 probably corresponds to L-Proline, NA or NA

## [1] 3.4977

## your peak at 3.4977 probably corresponds to NA, NA or NA

## [1] 0.9071

## your peak at 0.9071 probably corresponds to 3-Methyl-2-oxovaleric acid, Isovaleric acid or NA

## [1] 3.4268

## your peak at 3.4268 probably corresponds to D-Glucose, L-Carnitine or L-Proline

## [1] 3.252

## your peak at 3.252 probably corresponds to 3-Methylhistidine, Betaine or D-Glucose

## [1] 0.8621
```

```
## your peak at 0.8621 probably corresponds to NA, NA or NA

## [1] 3.9072

## your peak at 3.9072 probably corresponds to Betaine, D-Glucose or D-Mannose

## [1] 3.513

## your peak at 3.513 probably corresponds to Choline, NA or NA

## [1] 3.7528

## your peak at 3.7528 probably corresponds to D-Glucose, D-Mannose or L-Glutamic acid
```

```
assignation_NArm <- na.omit(assignation)
assignation
```

```
##                            Metabolite   HMDB_code Shift_ppm Type
## 7      2-Hydroxy-3-methylbutyric acid   HMDB00407     0.835    d
## NA16                            <NA>        <NA>        NA <NA>
## NA.19                           <NA>        <NA>        NA <NA>
## NA51                            <NA>        <NA>        NA <NA>
## NA.132                          <NA>        <NA>        NA <NA>
## NA.214                          <NA>        <NA>        NA <NA>
## NA42                            <NA>        <NA>        NA <NA>
## NA.126                          <NA>        <NA>        NA <NA>
## NA.211                          <NA>        <NA>        NA <NA>
## 17         3-Methyl-2-oxovaleric acid   HMDB00491     0.915    t
## 88                    Isovaleric acid   HMDB00718     0.915    d
## NA50                            <NA>        <NA>        NA <NA>
## 135                     L-Isoleucine   HMDB00172     0.941    t
## 92                       Ketoleucine   HMDB00695     0.945    d
## NA29                            <NA>        <NA>        NA <NA>
## 1351                    L-Isoleucine   HMDB00172     0.941    t
## 921                      Ketoleucine   HMDB00695     0.945    d
## 102    2-Hydroxy-3-methylbutyric acid   HMDB00407     0.965    d
## 1421                       L-Leucine   HMDB00687     0.963    t
## 101    2-Hydroxy-3-methylbutyric acid   HMDB00407     0.965    d
## NA24                            <NA>        <NA>        NA <NA>
## 1782                       L-Valine   HMDB00883     0.991    d
## 142                        L-Leucine   HMDB00687     0.963    t
## 10     2-Hydroxy-3-methylbutyric acid   HMDB00407     0.965    d
## 1781                       L-Valine   HMDB00883     0.991    d
## 1031       L-Alpha-aminobutyric acid   HMDB00452     0.997    t
## NA15                            <NA>        <NA>        NA <NA>
## 178                        L-Valine   HMDB00883     0.991    d
## 1371                    L-Isoleucine   HMDB00172     1.012    d
## 103        L-Alpha-aminobutyric acid   HMDB00452     0.997    t
## 137                     L-Isoleucine   HMDB00172     1.012    d
## NA2                             <NA>        <NA>        NA <NA>
## NA.11                           <NA>        <NA>        NA <NA>
## 179                        L-Valine   HMDB00883     1.044    d
```

```
## NA6                            <NA>        <NA>        NA <NA>
## NA.14                          <NA>        <NA>        NA <NA>
## NA3                            <NA>        <NA>        NA <NA>
## NA.12                          <NA>        <NA>        NA <NA>
## NA.21                          <NA>        <NA>        NA <NA>
## 86                Isopropyl alcohol   HMDB00863     1.177    d
## 70                          Ethanol   HMDB00108     1.185    t
## NA10                           <NA>        <NA>        NA <NA>
## 701                         Ethanol   HMDB00108     1.185    t
## NA11                           <NA>        <NA>        NA <NA>
## NA.16                          <NA>        <NA>        NA <NA>
## NA14                           <NA>        <NA>        NA <NA>
## NA.18                          <NA>        <NA>        NA <NA>
## NA.22                          <NA>        <NA>        NA <NA>
## 12             3-Hydroxybutyric acid   HMDB00357     1.219    d
## 84                   Isobutyric acid   HMDB01873     1.225    d
## NA                             <NA>        <NA>        NA <NA>
## 121            3-Hydroxybutyric acid   HMDB00357     1.219    d
## 841                  Isobutyric acid   HMDB01873     1.225    d
## NA8                            <NA>        <NA>        NA <NA>
## 139                   L-Isoleucine   HMDB00172     1.263    m
## 16          3-Hydroxyisovaleric acid   HMDB00754     1.275    s
## NA22                           <NA>        <NA>        NA <NA>
## NA17                           <NA>        <NA>        NA <NA>
## NA.110                         <NA>        <NA>        NA <NA>
## NA.23                          <NA>        <NA>        NA <NA>
## 1671                    L-Threonine   HMDB00167     1.331    d
## NA26                           <NA>        <NA>        NA <NA>
## NA.115                         <NA>        <NA>        NA <NA>
## 141                   L-Lactic acid   HMDB00190     1.335    d
## 167                     L-Threonine   HMDB00167     1.331    d
## NA9                            <NA>        <NA>        NA <NA>
## NA1                            <NA>        <NA>        NA <NA>
## NA.1                           <NA>        <NA>        NA <NA>
## NA.2                           <NA>        <NA>        NA <NA>
## 1011                      L-Alanine   HMDB00161     1.485    d
## NA45                           <NA>        <NA>        NA <NA>
## NA.128                         <NA>        <NA>        NA <NA>
## NA30                           <NA>        <NA>        NA <NA>
## NA.118                         <NA>        <NA>        NA <NA>
## NA.29                          <NA>        <NA>        NA <NA>
## NA34                           <NA>        <NA>        NA <NA>
## NA.122                         <NA>        <NA>        NA <NA>
## NA.210                         <NA>        <NA>        NA <NA>
## NA23                           <NA>        <NA>        NA <NA>
## NA.113                         <NA>        <NA>        NA <NA>
## NA.26                          <NA>        <NA>        NA <NA>
## NA25                           <NA>        <NA>        NA <NA>
## NA.114                         <NA>        <NA>        NA <NA>
## NA.27                          <NA>        <NA>        NA <NA>
## 161                       L-Proline   HMDB00162     2.005    m
## 134                   L-Isoleucine   HMDB00172     1.983    m
## NA43                           <NA>        <NA>        NA <NA>
## 1612                      L-Proline   HMDB00162     2.005    m
```

```
## NA47                             <NA>        <NA>        NA <NA>
## NA.129                           <NA>        <NA>        NA <NA>
## 1611                        L-Proline   HMDB00162     2.005    m
## 9      2-Hydroxy-3-methylbutyric acid   HMDB00407     2.025    m
## NA46                             <NA>        <NA>        NA <NA>
## 301                    Acetylglycine   HMDB00532     2.065    s
## 1251                 L-Glutamic acid   HMDB00148     2.055    m
## 891                  Isovaleric acid   HMDB00718     2.065    d
## 30                     Acetylglycine   HMDB00532     2.065    s
## 125                  L-Glutamic acid   HMDB00148     2.055    m
## 89                   Isovaleric acid   HMDB00718     2.065    d
## 163                       L-Proline   HMDB00162     2.075    m
## NA48                             <NA>        <NA>        NA <NA>
## NA.130                           <NA>        <NA>        NA <NA>
## 91                       Ketoleucine   HMDB00695     2.105    m
## NA35                             <NA>        <NA>        NA <NA>
## NA.123                           <NA>        <NA>        NA <NA>
## 127                     L-Glutamine   HMDB00641     2.140    m
## 97                L-Acetylcarnitine   HMDB00201     2.145    s
## NA18                             <NA>        <NA>        NA <NA>
## 282                         Acetone   HMDB01659     2.235    s
## NA33                             <NA>        <NA>        NA <NA>
## NA.121                           <NA>        <NA>        NA <NA>
## 283                         Acetone   HMDB01659     2.235    s
## NA38                             <NA>        <NA>        NA <NA>
## NA.125                           <NA>        <NA>        NA <NA>
## 281                         Acetone   HMDB01659     2.235    s
## NA32                             <NA>        <NA>        NA <NA>
## NA.120                           <NA>        <NA>        NA <NA>
## 28                          Acetone   HMDB01659     2.235    s
## NA5                              <NA>        <NA>        NA <NA>
## NA.13                            <NA>        <NA>        NA <NA>
## NA19                             <NA>        <NA>        NA <NA>
## NA.111                           <NA>        <NA>        NA <NA>
## NA.24                            <NA>        <NA>        NA <NA>
## 180                         L-Valine   HMDB00883     2.276    m
## NA31                             <NA>        <NA>        NA <NA>
## NA.119                           <NA>        <NA>        NA <NA>
## NA44                             <NA>        <NA>        NA <NA>
## NA.127                           <NA>        <NA>        NA <NA>
## NA.212                           <NA>        <NA>        NA <NA>
## 187                       Ornithine   HMDB00214     3.061    t
## NA12                             <NA>        <NA>        NA <NA>
## NA.17                            <NA>        <NA>        NA <NA>
## 107                      L-Arginine   HMDB00517     3.245    t
## 133                      L-Histidine   HMDB00177     3.245   dd
## NA20                             <NA>        <NA>        NA <NA>
## 54                        D-Glucose   HMDB00122     3.248   dd
## 381                         Betaine   HMDB00043     3.265    s
## 261               3-Methylhistidine   HMDB00479     3.255    m
## 38                          Betaine   HMDB00043     3.265    s
## 26                3-Methylhistidine   HMDB00479     3.255    m
## NA40                             <NA>        <NA>        NA <NA>
## 183                     Myoinositol   HMDB00211     3.283    t
```

```
## NA37                          <NA>        <NA>        NA <NA>
## NA.124                        <NA>        <NA>        NA <NA>
## 60                     D-Mannose   HMDB00169     3.385  ddd
## NA28                          <NA>        <NA>        NA <NA>
## NA.117                        <NA>        <NA>        NA <NA>
## 50                     D-Glucose   HMDB00122     3.413    m
## 164                    L-Proline   HMDB00162     3.425   dt
## 1211                 L-Carnitine HMDB0000062     3.419    m
## 501                    D-Glucose   HMDB00122     3.413    m
## 1641                   L-Proline   HMDB00162     3.425   dt
## 1212                 L-Carnitine HMDB0000062     3.419    m
## 491                    D-Glucose   HMDB00122     3.473    m
## 2071                     Sucrose   HMDB00258     3.475    t
## NA41                          <NA>        <NA>        NA <NA>
## 49                     D-Glucose   HMDB00122     3.473    m
## 207                      Sucrose   HMDB00258     3.475    t
## NA36                          <NA>        <NA>        NA <NA>
## NA49                          <NA>        <NA>        NA <NA>
## NA.131                        <NA>        <NA>        NA <NA>
## NA.213                        <NA>        <NA>        NA <NA>
## 41                       Choline   HMDB00097     3.522   dd
## NA52                          <NA>        <NA>        NA <NA>
## NA.133                        <NA>        <NA>        NA <NA>
## 168                  L-Threonine   HMDB00167     3.590    d
## 73                      Glycerol   HMDB00131     3.566    m
## 57                     D-Mannose   HMDB00169     3.579    t
## 75                      Glycerol   HMDB00131     3.659    m
## 61                     D-Mannose   HMDB00169     3.664    m
## 69                       Ethanol   HMDB00108     3.665    q
## 138                L-Isoleucine   HMDB00172     3.676    d
## 1              1-Methylhistidine   HMDB00001     3.695    s
## 211                      Sucrose   HMDB00258     3.685    s
## 55                     D-Glucose   HMDB00122     3.743    m
## 63                     D-Mannose   HMDB00169     3.754    m
## 122              L-Glutamic acid   HMDB00148     3.763   dd
## 59                     D-Mannose   HMDB00169     3.814    m
## 217                      Uridine   HMDB00296     3.816   dd
## NA39                          <NA>        <NA>        NA <NA>
## 51                     D-Glucose   HMDB00122     3.904   dd
## 37                       Betaine   HMDB00043     3.905    s
## 64                     D-Mannose   HMDB00169     3.898   dd
## 511                    D-Glucose   HMDB00122     3.904   dd
## 371                      Betaine   HMDB00043     3.905    s
## 641                    D-Mannose   HMDB00169     3.898   dd
## 58                     D-Mannose   HMDB00169     3.938    m
## 45                      Creatine   HMDB00064     3.935    s
## 221                      Uridine   HMDB00296     3.922   dd
## 1401               L-Lactic acid   HMDB00190     4.115    q
## NA7                           <NA>        <NA>        NA <NA>
## NA.15                         <NA>        <NA>        NA <NA>
## 140                L-Lactic acid   HMDB00190     4.115    q
## 1621                   L-Proline   HMDB00162     4.135   dd
## 2221                     Uridine   HMDB00296     4.136    m
## 162                    L-Proline   HMDB00162     4.135   dd
```

```
## 222                          Uridine   HMDB00296     4.136    m
## NA4                             <NA>        <NA>        NA <NA>
## 1622                       L-Proline   HMDB00162     4.135   dd
## 2222                         Uridine   HMDB00296     4.136    m
## NA13                            <NA>        <NA>        NA <NA>
## NA21                            <NA>        <NA>        NA <NA>
## NA.112                          <NA>        <NA>        NA <NA>
## NA.25                           <NA>        <NA>        NA <NA>
## NA27                            <NA>        <NA>        NA <NA>
## NA.116                          <NA>        <NA>        NA <NA>
## NA.28                           <NA>        <NA>        NA <NA>
##        Blood_concentration ppm_to_assign
## 7                  6.10000        0.8461
## NA16                    NA        0.8461
## NA.19                   NA        0.8461
## NA51                    NA        0.8621
## NA.132                  NA        0.8621
## NA.214                  NA        0.8621
## NA42                    NA        0.8955
## NA.126                  NA        0.8955
## NA.211                  NA        0.8955
## 17                 20.35000        0.9071
## 88                  1.60000        0.9071
## NA50                    NA        0.9071
## 135                52.44667        0.9422
## 92                 30.75000        0.9422
## NA29                    NA        0.9422
## 1351               52.44667        0.9546
## 921                30.75000        0.9546
## 102                 6.10000        0.9546
## 1421              112.74615        0.9670
## 101                 6.10000        0.9670
## NA24                    NA        0.9670
## 1782              179.04615        0.9775
## 142               112.74615        0.9775
## 10                  6.10000        0.9775
## 1781              179.04615        0.9881
## 1031               22.80000        0.9881
## NA15                    NA        0.9881
## 178               179.04615        1.0018
## 1371               52.44667        1.0018
## 103                22.80000        1.0018
## 137                52.44667        1.0136
## NA2                     NA        1.0136
## NA.11                   NA        1.0136
## 179               179.04615        1.0521
## NA6                     NA        1.0521
## NA.14                   NA        1.0521
## NA3                     NA        1.0639
## NA.12                   NA        1.0639
## NA.21                   NA        1.0639
## 86                 83.30000        1.1835
## 70                 27.50000        1.1835
## NA10                    NA        1.1835
```

```
## 701              27.50000       1.1938
## NA11                   NA       1.1938
## NA.16                  NA       1.1938
## NA14                   NA       1.2002
## NA.18                  NA       1.2002
## NA.22                  NA       1.2002
## 12              147.74000       1.2120
## 84                2.30000       1.2120
## NA                     NA       1.2120
## 121             147.74000       1.2217
## 841               2.30000       1.2217
## NA8                    NA       1.2217
## 139              52.44667       1.2653
## 16                4.00000       1.2653
## NA22                   NA       1.2653
## NA17                   NA       1.2926
## NA.110                 NA       1.2926
## NA.23                  NA       1.2926
## 1671            124.23077       1.3199
## NA26                   NA       1.3199
## NA.115                 NA       1.3199
## 141            2235.17500       1.3399
## 167             124.23077       1.3399
## NA9                    NA       1.3399
## NA1                    NA       1.3515
## NA.1                   NA       1.3515
## NA.2                   NA       1.3515
## 1011            298.12000       1.4916
## NA45                   NA       1.4916
## NA.128                 NA       1.4916
## NA30                   NA       1.5039
## NA.118                 NA       1.5039
## NA.29                  NA       1.5039
## NA34                   NA       1.5792
## NA.122                 NA       1.5792
## NA.210                 NA       1.5792
## NA23                   NA       1.5908
## NA.113                 NA       1.5908
## NA.26                  NA       1.5908
## NA25                   NA       1.5995
## NA.114                 NA       1.5995
## NA.27                  NA       1.5995
## 161             151.34545       1.9937
## 134              52.44667       1.9937
## NA43                   NA       1.9937
## 1612            151.34545       2.0045
## NA47                   NA       2.0045
## NA.129                 NA       2.0045
## 1611            151.34545       2.0172
## 9                 6.10000       2.0172
## NA46                   NA       2.0172
## 301              89.57000       2.0569
## 1251             45.47267       2.0569
## 891               1.60000       2.0569
```

```
## 30               89.57000       2.0632
## 125              45.47267       2.0632
## 89                1.60000       2.0632
## 163             151.34545       2.0858
## NA48                   NA       2.0858
## NA.130                 NA       2.0858
## 91               30.75000       2.0970
## NA35                   NA       2.0970
## NA.123                 NA       2.0970
## 127             539.32857       2.1518
## 97                5.57000       2.1518
## NA18                   NA       2.1518
## 282              57.54000       2.2324
## NA33                   NA       2.2324
## NA.121                 NA       2.2324
## 283              57.54000       2.2350
## NA38                   NA       2.2350
## NA.125                 NA       2.2350
## 281              57.54000       2.2353
## NA32                   NA       2.2353
## NA.120                 NA       2.2353
## 28               57.54000       2.2463
## NA5                    NA       2.2463
## NA.13                  NA       2.2463
## NA19                   NA       2.2501
## NA.111                 NA       2.2501
## NA.24                  NA       2.2501
## 180             179.04615       2.2611
## NA31                   NA       2.2611
## NA.119                 NA       2.2611
## NA44                   NA       2.7630
## NA.127                 NA       2.7630
## NA.212                 NA       2.7630
## 187              55.76923       3.0602
## NA12                   NA       3.0602
## NA.17                  NA       3.0602
## 107              82.44688       3.2311
## 133              74.94615       3.2311
## NA20                   NA       3.2311
## 54             4976.01765       3.2520
## 381              61.14750       3.2520
## 261               2.85000       3.2520
## 38               61.14750       3.2673
## 26                2.85000       3.2673
## NA40                   NA       3.2673
## 183              23.52500       3.2807
## NA37                   NA       3.2807
## NA.124                 NA       3.2807
## 60               51.50000       3.3799
## NA28                   NA       3.3799
## NA.117                 NA       3.3799
## 50             4976.01765       3.4109
## 164             151.34545       3.4109
## 1211             28.50000       3.4109
```

```
## 501            4976.01765       3.4268
## 1641            151.34545       3.4268
## 1212             28.50000       3.4268
## 491            4976.01765       3.4722
## 2071              1.80000       3.4722
## NA41                   NA       3.4722
## 49             4976.01765       3.4782
## 207               1.80000       3.4782
## NA36                   NA       3.4782
## NA49                   NA       3.4977
## NA.131                 NA       3.4977
## NA.213                 NA       3.4977
## 41               13.08000       3.5130
## NA52                   NA       3.5130
## NA.133                 NA       3.5130
## 168             124.23077       3.5798
## 73              103.51429       3.5798
## 57               51.50000       3.5798
## 75              103.51429       3.6628
## 61               51.50000       3.6628
## 69               27.50000       3.6628
## 138              52.44667       3.6894
## 1                12.52857       3.6894
## 211               1.80000       3.6894
## 55             4976.01765       3.7528
## 63               51.50000       3.7528
## 122              45.47267       3.7528
## 59               51.50000       3.8118
## 217               3.11000       3.8118
## NA39                   NA       3.8118
## 51             4976.01765       3.8940
## 37               61.14750       3.8940
## 64               51.50000       3.8940
## 511            4976.01765       3.9072
## 371              61.14750       3.9072
## 641              51.50000       3.9072
## 58               51.50000       3.9276
## 45               48.47000       3.9276
## 221               3.11000       3.9276
## 1401           2235.17500       4.1119
## NA7                    NA       4.1119
## NA.15                  NA       4.1119
## 140            2235.17500       4.1235
## 1621            151.34545       4.1235
## 2221              3.11000       4.1235
## 162             151.34545       4.1352
## 222               3.11000       4.1352
## NA4                    NA       4.1352
## 1622            151.34545       4.1468
## 2222              3.11000       4.1468
## NA13                   NA       4.1468
## NA21                   NA       5.2826
## NA.112                 NA       5.2826
## NA.25                  NA       5.2826
```

```
## NA27                    NA        5.3228
## NA.116                  NA        5.3228
## NA.28                   NA        5.3228
```

```
assignation_NArm
```

```
##                          Metabolite   HMDB_code Shift_ppm Type
## 7    2-Hydroxy-3-methylbutyric acid   HMDB00407     0.835    d
## 17       3-Methyl-2-oxovaleric acid   HMDB00491     0.915    t
## 88                  Isovaleric acid   HMDB00718     0.915    d
## 135                    L-Isoleucine   HMDB00172     0.941    t
## 92                     Ketoleucine   HMDB00695     0.945    d
## 1351                   L-Isoleucine   HMDB00172     0.941    t
## 921                    Ketoleucine   HMDB00695     0.945    d
## 102  2-Hydroxy-3-methylbutyric acid   HMDB00407     0.965    d
## 1421                     L-Leucine   HMDB00687     0.963    t
## 101  2-Hydroxy-3-methylbutyric acid   HMDB00407     0.965    d
## 1782                     L-Valine   HMDB00883     0.991    d
## 142                      L-Leucine   HMDB00687     0.963    t
## 10   2-Hydroxy-3-methylbutyric acid   HMDB00407     0.965    d
## 1781                     L-Valine   HMDB00883     0.991    d
## 1031      L-Alpha-aminobutyric acid   HMDB00452     0.997    t
## 178                      L-Valine   HMDB00883     0.991    d
## 1371                   L-Isoleucine   HMDB00172     1.012    d
## 103       L-Alpha-aminobutyric acid   HMDB00452     0.997    t
## 137                    L-Isoleucine   HMDB00172     1.012    d
## 179                      L-Valine   HMDB00883     1.044    d
## 86               Isopropyl alcohol   HMDB00863     1.177    d
## 70                         Ethanol   HMDB00108     1.185    t
## 701                        Ethanol   HMDB00108     1.185    t
## 12           3-Hydroxybutyric acid   HMDB00357     1.219    d
## 84                 Isobutyric acid   HMDB01873     1.225    d
## 121          3-Hydroxybutyric acid   HMDB00357     1.219    d
## 841                Isobutyric acid   HMDB01873     1.225    d
## 139                    L-Isoleucine   HMDB00172     1.263    m
## 16        3-Hydroxyisovaleric acid   HMDB00754     1.275    s
## 1671                    L-Threonine   HMDB00167     1.331    d
## 141                   L-Lactic acid   HMDB00190     1.335    d
## 167                     L-Threonine   HMDB00167     1.331    d
## 1011                      L-Alanine   HMDB00161     1.485    d
## 161                       L-Proline   HMDB00162     2.005    m
## 134                    L-Isoleucine   HMDB00172     1.983    m
## 1612                      L-Proline   HMDB00162     2.005    m
## 1611                      L-Proline   HMDB00162     2.005    m
## 9    2-Hydroxy-3-methylbutyric acid   HMDB00407     2.025    m
## 301                   Acetylglycine   HMDB00532     2.065    s
## 1251                L-Glutamic acid   HMDB00148     2.055    m
## 891                 Isovaleric acid   HMDB00718     2.065    d
## 30                    Acetylglycine   HMDB00532     2.065    s
## 125                 L-Glutamic acid   HMDB00148     2.055    m
## 89                  Isovaleric acid   HMDB00718     2.065    d
## 163                       L-Proline   HMDB00162     2.075    m
## 91                      Ketoleucine   HMDB00695     2.105    m
## 127                     L-Glutamine   HMDB00641     2.140    m
```

```
## 97               L-Acetylcarnitine   HMDB00201     2.145    s
## 282                        Acetone   HMDB01659     2.235    s
## 283                        Acetone   HMDB01659     2.235    s
## 281                        Acetone   HMDB01659     2.235    s
## 28                         Acetone   HMDB01659     2.235    s
## 180                      L-Valine    HMDB00883     2.276    m
## 187                     Ornithine    HMDB00214     3.061    t
## 107                   L-Arginine     HMDB00517     3.245    t
## 133                  L-Histidine     HMDB00177     3.245   dd
## 54                     D-Glucose     HMDB00122     3.248   dd
## 381                      Betaine     HMDB00043     3.265    s
## 261              3-Methylhistidine   HMDB00479     3.255    m
## 38                       Betaine     HMDB00043     3.265    s
## 26               3-Methylhistidine   HMDB00479     3.255    m
## 183                 Myoinositol      HMDB00211     3.283    t
## 60                     D-Mannose     HMDB00169     3.385  ddd
## 50                     D-Glucose     HMDB00122     3.413    m
## 164                    L-Proline     HMDB00162     3.425   dt
## 1211                 L-Carnitine  HMDB0000062     3.419    m
## 501                    D-Glucose     HMDB00122     3.413    m
## 1641                   L-Proline     HMDB00162     3.425   dt
## 1212                 L-Carnitine  HMDB0000062     3.419    m
## 491                    D-Glucose     HMDB00122     3.473    m
## 2071                     Sucrose     HMDB00258     3.475    t
## 49                     D-Glucose     HMDB00122     3.473    m
## 207                      Sucrose     HMDB00258     3.475    t
## 41                       Choline     HMDB00097     3.522   dd
## 168                  L-Threonine     HMDB00167     3.590    d
## 73                      Glycerol     HMDB00131     3.566    m
## 57                     D-Mannose     HMDB00169     3.579    t
## 75                      Glycerol     HMDB00131     3.659    m
## 61                     D-Mannose     HMDB00169     3.664    m
## 69                       Ethanol     HMDB00108     3.665    q
## 138                 L-Isoleucine     HMDB00172     3.676    d
## 1                3-Methylhistidine   HMDB00001     3.695    s
## 211                      Sucrose     HMDB00258     3.685    s
## 55                     D-Glucose     HMDB00122     3.743    m
## 63                     D-Mannose     HMDB00169     3.754    m
## 122              L-Glutamic acid     HMDB00148     3.763   dd
## 59                     D-Mannose     HMDB00169     3.814    m
## 217                      Uridine     HMDB00296     3.816   dd
## 51                     D-Glucose     HMDB00122     3.904   dd
## 37                       Betaine     HMDB00043     3.905    s
## 64                     D-Mannose     HMDB00169     3.898   dd
## 511                    D-Glucose     HMDB00122     3.904   dd
## 371                      Betaine     HMDB00043     3.905    s
## 641                    D-Mannose     HMDB00169     3.898   dd
## 58                     D-Mannose     HMDB00169     3.938    m
## 45                      Creatine     HMDB00064     3.935    s
## 221                      Uridine     HMDB00296     3.922   dd
## 1401                L-Lactic acid     HMDB00190     4.115    q
## 140                 L-Lactic acid     HMDB00190     4.115    q
## 1621                   L-Proline     HMDB00162     4.135   dd
## 2221                     Uridine     HMDB00296     4.136    m
```

```
## 162                     L-Proline   HMDB00162     4.135   dd
## 222                       Uridine   HMDB00296     4.136    m
## 1622                    L-Proline   HMDB00162     4.135   dd
## 2222                      Uridine   HMDB00296     4.136    m
##      Blood_concentration ppm_to_assign
## 7                6.10000        0.8461
## 17              20.35000        0.9071
## 88               1.60000        0.9071
## 135             52.44667        0.9422
## 92              30.75000        0.9422
## 1351            52.44667        0.9546
## 921             30.75000        0.9546
## 102              6.10000        0.9546
## 1421           112.74615        0.9670
## 101              6.10000        0.9670
## 1782           179.04615        0.9775
## 142            112.74615        0.9775
## 10               6.10000        0.9775
## 1781           179.04615        0.9881
## 1031            22.80000        0.9881
## 178            179.04615        1.0018
## 1371            52.44667        1.0018
## 103             22.80000        1.0018
## 137             52.44667        1.0136
## 179            179.04615        1.0521
## 86              83.30000        1.1835
## 70              27.50000        1.1835
## 701             27.50000        1.1938
## 12             147.74000        1.2120
## 84               2.30000        1.2120
## 121            147.74000        1.2217
## 841              2.30000        1.2217
## 139             52.44667        1.2653
## 16               4.00000        1.2653
## 1671           124.23077        1.3199
## 141           2235.17500        1.3399
## 167            124.23077        1.3399
## 1011           298.12000        1.4916
## 161            151.34545        1.9937
## 134             52.44667        1.9937
## 1612           151.34545        2.0045
## 1611           151.34545        2.0172
## 9                6.10000        2.0172
## 301             89.57000        2.0569
## 1251            45.47267        2.0569
## 891              1.60000        2.0569
## 30              89.57000        2.0632
## 125             45.47267        2.0632
## 89               1.60000        2.0632
## 163            151.34545        2.0858
## 91              30.75000        2.0970
## 127            539.32857        2.1518
## 97               5.57000        2.1518
## 282             57.54000        2.2324
```

```
## 283            57.54000       2.2350
## 281            57.54000       2.2353
## 28             57.54000       2.2463
## 180           179.04615       2.2611
## 187            55.76923       3.0602
## 107            82.44688       3.2311
## 133            74.94615       3.2311
## 54           4976.01765       3.2520
## 381            61.14750       3.2520
## 261             2.85000       3.2520
## 38             61.14750       3.2673
## 26              2.85000       3.2673
## 183            23.52500       3.2807
## 60             51.50000       3.3799
## 50           4976.01765       3.4109
## 164           151.34545       3.4109
## 1211           28.50000       3.4109
## 501          4976.01765       3.4268
## 1641          151.34545       3.4268
## 1212           28.50000       3.4268
## 491          4976.01765       3.4722
## 2071            1.80000       3.4722
## 49           4976.01765       3.4782
## 207             1.80000       3.4782
## 41             13.08000       3.5130
## 168           124.23077       3.5798
## 73            103.51429       3.5798
## 57             51.50000       3.5798
## 75            103.51429       3.6628
## 61             51.50000       3.6628
## 69             27.50000       3.6628
## 138            52.44667       3.6894
## 1              12.52857       3.6894
## 211             1.80000       3.6894
## 55           4976.01765       3.7528
## 63             51.50000       3.7528
## 122            45.47267       3.7528
## 59             51.50000       3.8118
## 217             3.11000       3.8118
## 51           4976.01765       3.8940
## 37             61.14750       3.8940
## 64             51.50000       3.8940
## 511          4976.01765       3.9072
## 371            61.14750       3.9072
## 641            51.50000       3.9072
## 58             51.50000       3.9276
## 45             48.47000       3.9276
## 221             3.11000       3.9276
## 1401         2235.17500       4.1119
## 140          2235.17500       4.1235
## 1621          151.34545       4.1235
## 2221            3.11000       4.1235
## 162           151.34545       4.1352
## 222             3.11000       4.1352
```

```
## 1622          151.34545       4.1468
## 2222            3.11000       4.1468
```