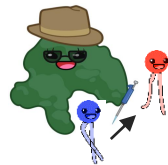


Jakob Koch – Biochemical Genetics Laboratory, MUI

MS Meets R: Unravelling Cellular Lipid Networks by
Integrative Analysis & Untangling ether lipids through in
dept utilization of LC-IM-MS/MS



2024-07-09



Salzburg

2024



Overview

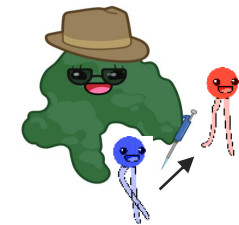
1. Cardiolipin diversity and the cellular PL pool

`library(neuralnet)`



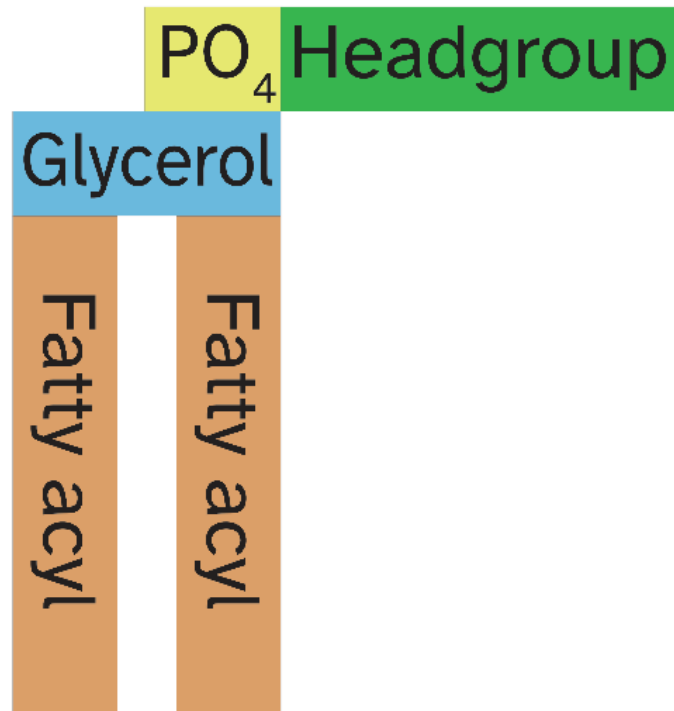
2. LC-IM-MS analysis of alkyl- and alkenyl-lipids

`library(timsr)`

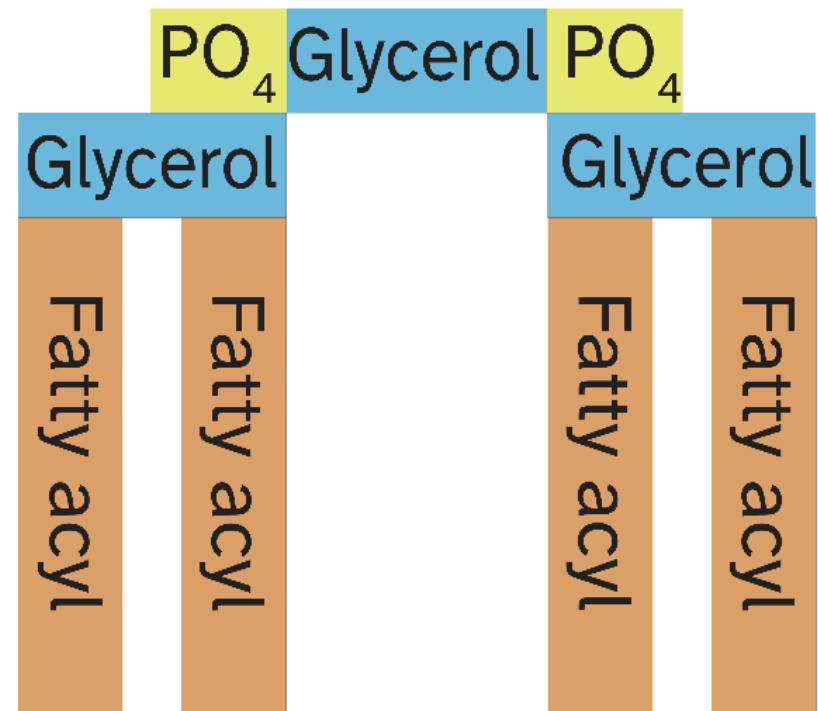


Lipids

Phospholipids (PL)



Cardiolipins (CL)



Overview

1. Cardiolipin diversity and the cellular PL pool



`library(neuralnet)`

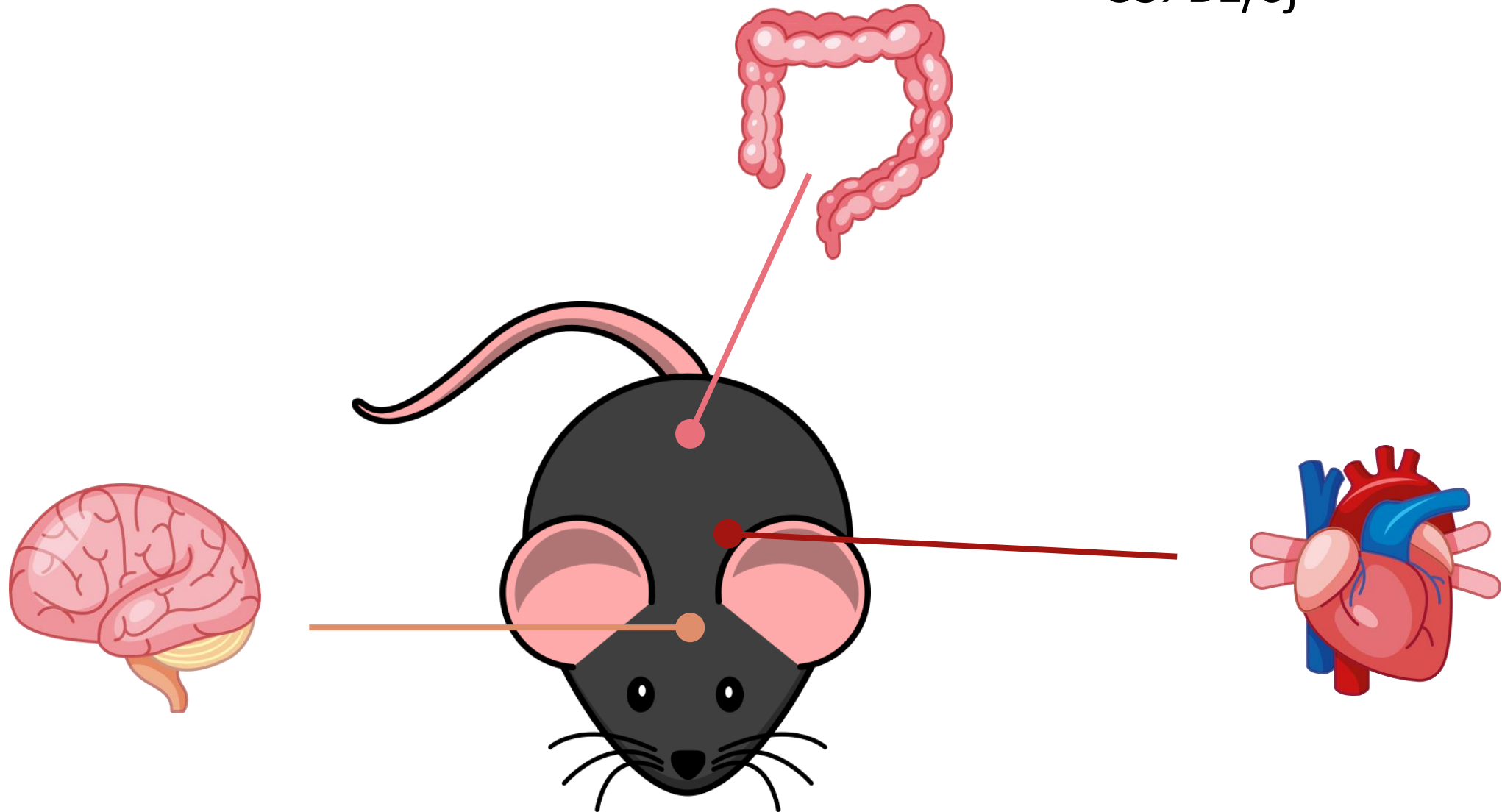
2. LC-IM-MS analysis of alkyl- and alkenyl-lipids

`library(timsr)`

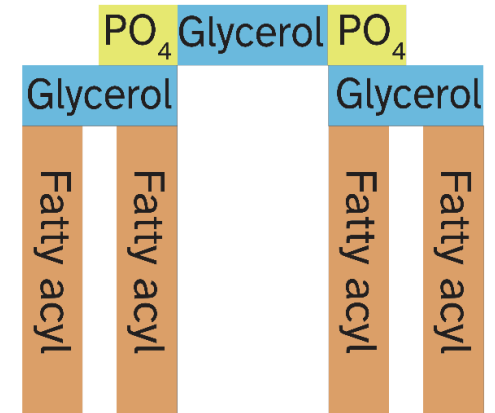
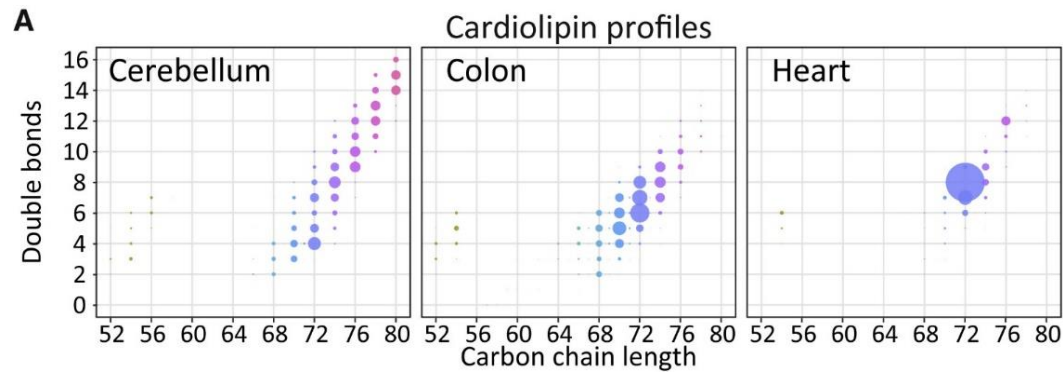
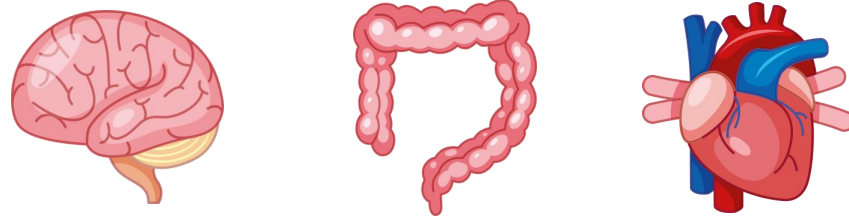


Mouse tissues

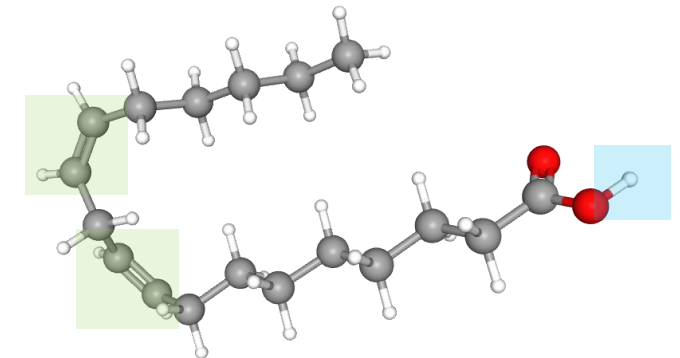
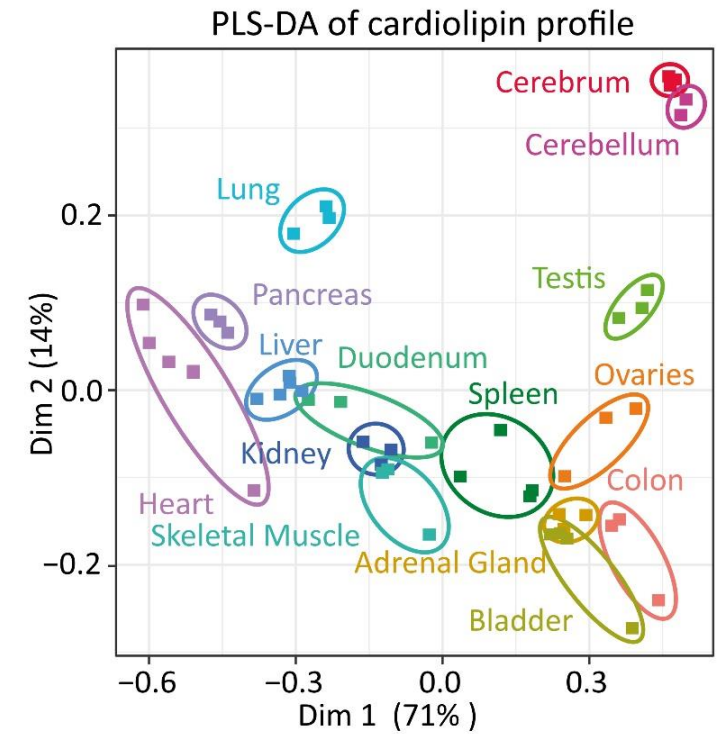
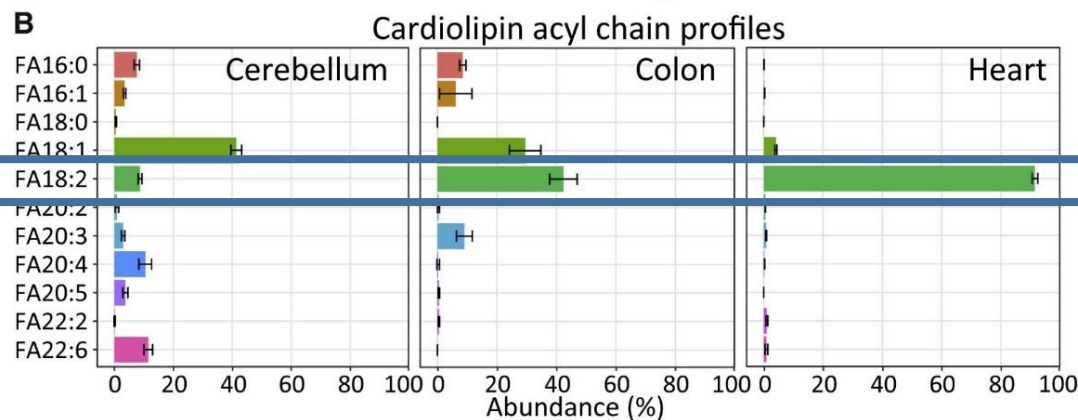
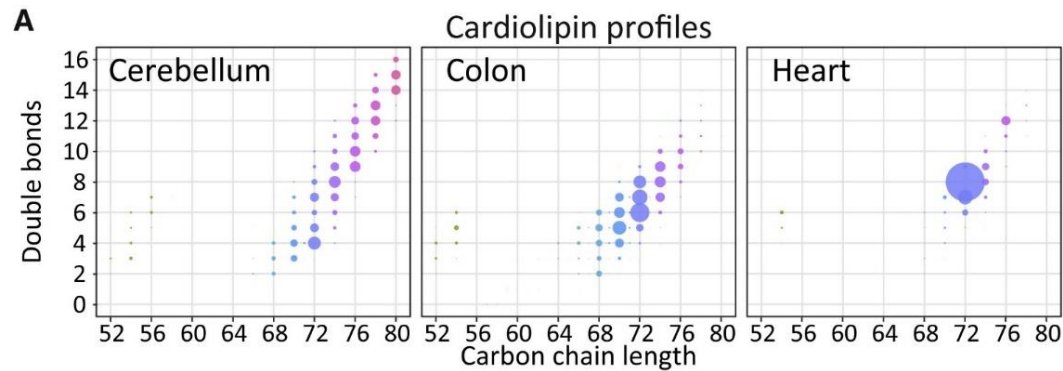
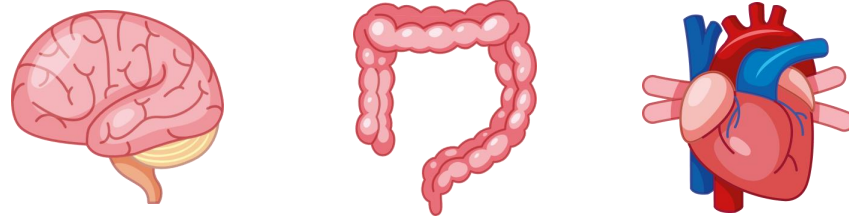
C57BL/6J



Diverse Cardiolipin profiles



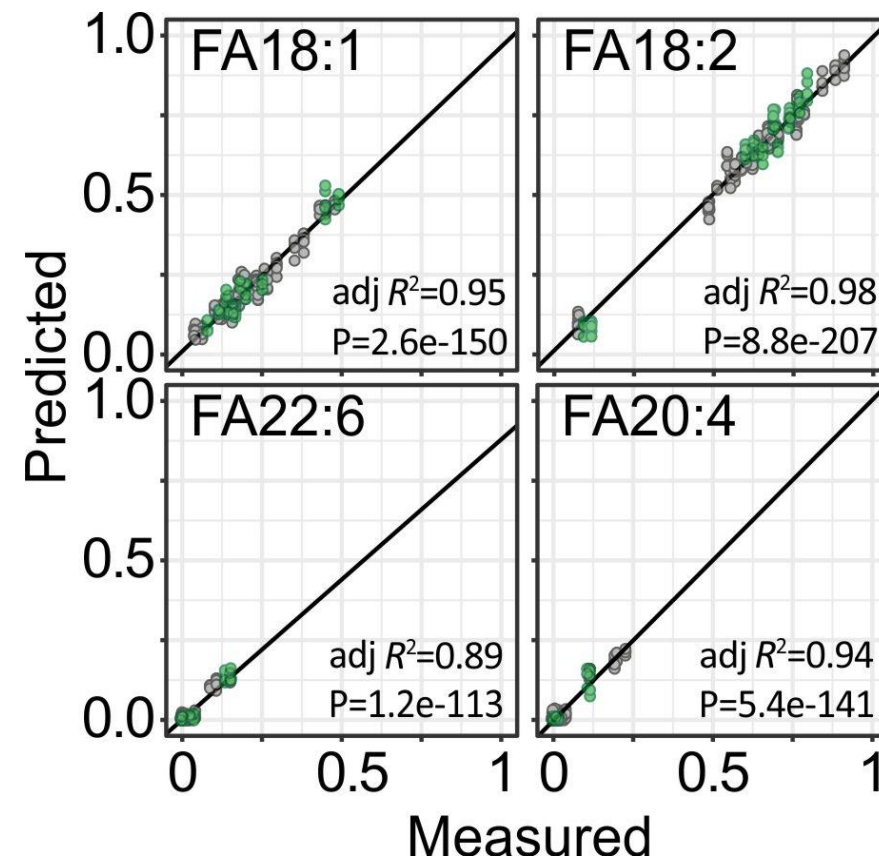
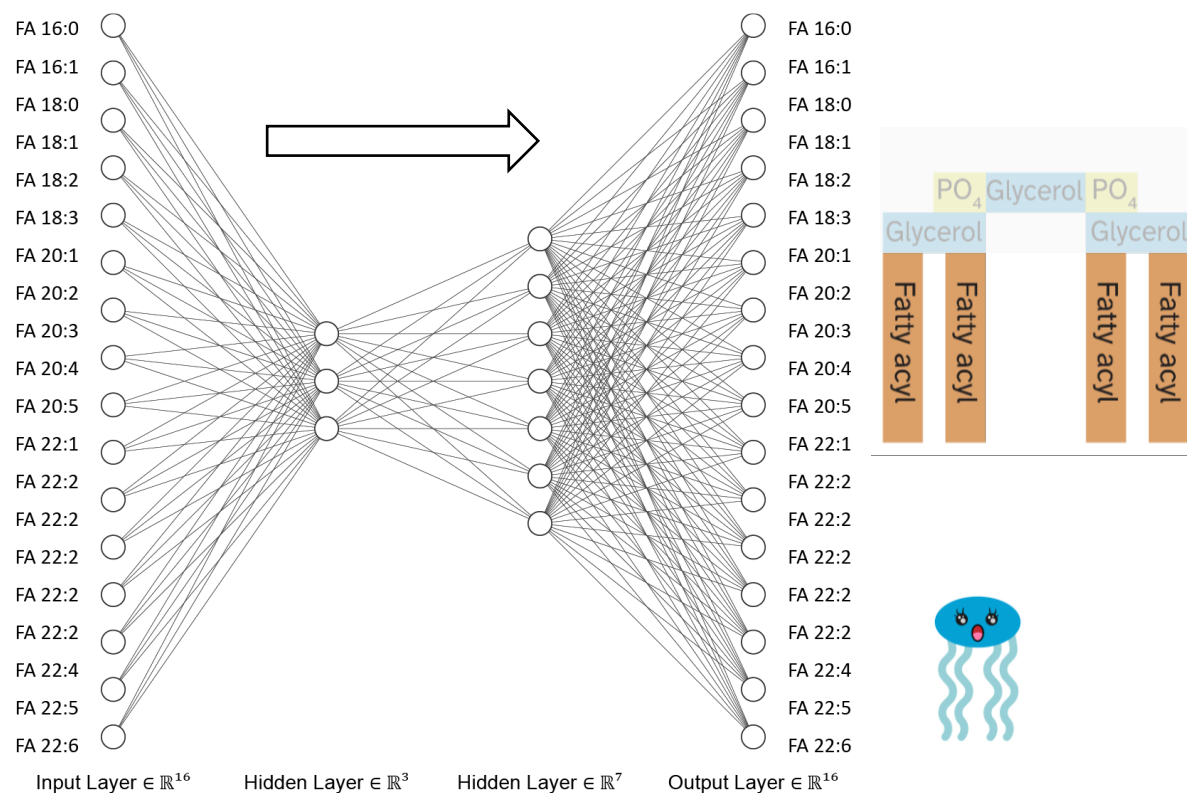
Constructed from Fatty acyls



Artificial neural network correctly predicts CL profile

Cellular
PL acyl pool

R-package *neuralnet* (version
1.44.2, [Günther and Fritsch, 2010](#))



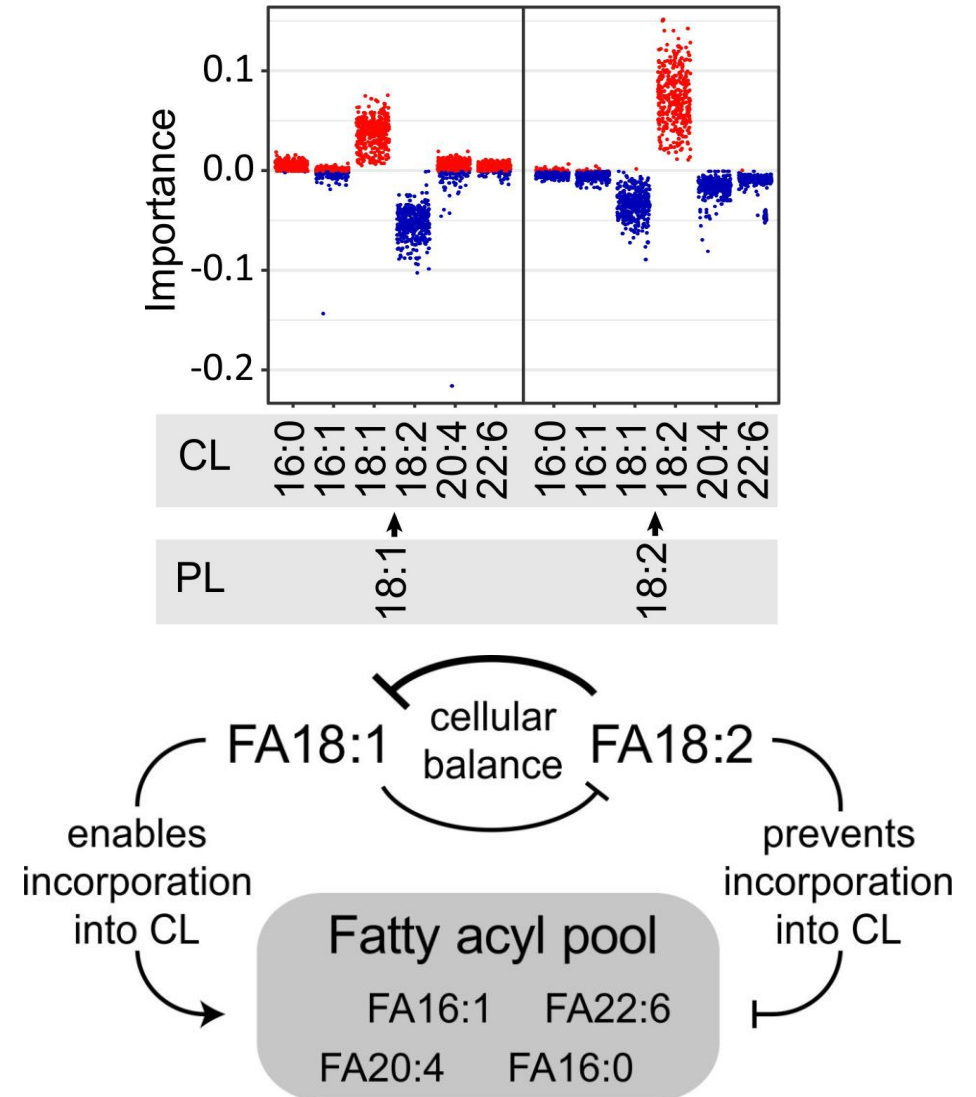
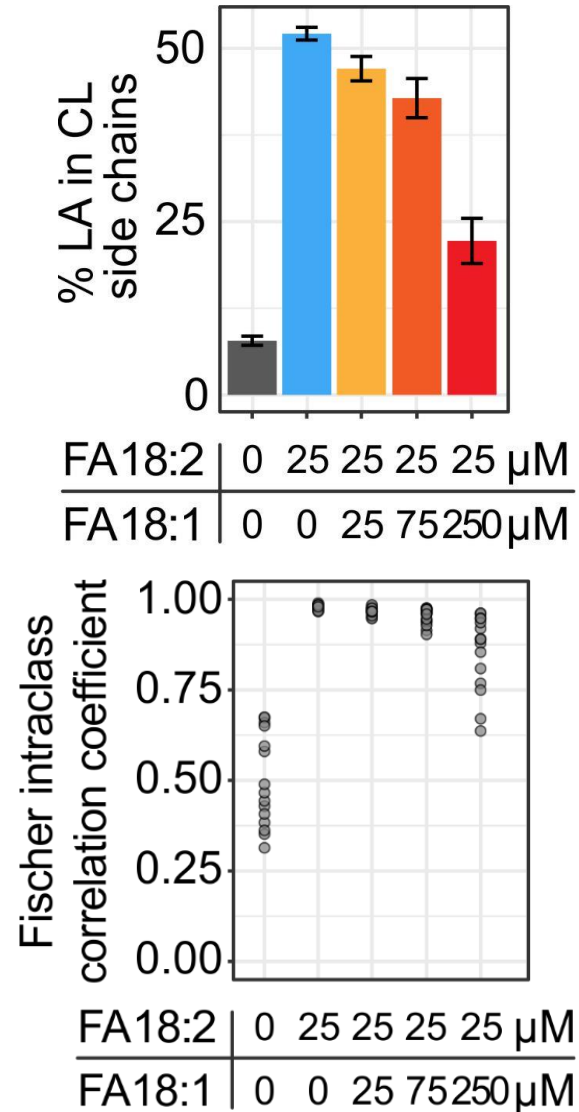
neuralnet() modelling call

```
132 plan(multiprocess(workers=(detectCores()*0.6)))
133 lusz <- foreach(i = seq_len(ncol(network_setup)), .combine="comb_cust_3",
134   .init=list(list(), list(),list(),list())) %dopar% {
135   node_definition <- unlist(network_setup[,i])[unlist(network_setup[,i])!=0]
136   rm(nn)
137   nn <- neuralnet(formula = Y16.0 + Y16.1 + Y18.0 + Y18.1 + Y18.2+ Y18.3+
Y20.1+ Y20.2+
138   Y20.3+ Y20.4 + Y20.5 + Y22.1 + Y22.2 + Y22.4 + Y22.5 +
Y22.6 ~
139   X16.0 + X16.1 + X18.0 + X18.1 + X18.2 + X18.3 + X20.1 +
X20.2 + X20.3 + X20.4 +
140   X20.5 + X22.1 + X22.2 + X22.4 + X22.5 + X22.6,data =
training,hidden = node_definition,rep=30
141   #',lifesign = "minimal"
142   #',learningrate = 0.01
143   #',algorithm = "slr"|
144   ,algorithm = "rprop+"
145   ,linear.output = FALSE
146   #',err.fct = "sse"
147   #',act.fct = "tanh",
148   ,likelihood = TRUE
149   ,act.fct = softplus
150   ,threshold = 0.008
151   ,stepmax = 200000
152   )
153   rm(test,pred,test_real,act,distance,cumulative)
```

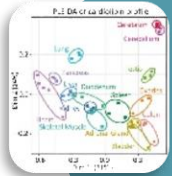


Rational of cardiolipin regulation

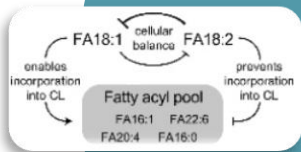
Ex vivo
validation of
in vivo
ANN model



FACTS I



Mouse tissue mitochondria are defined by characteristic cardiolipin architectures



ANN reveals PL pools are in control of cardiolipin specificity



neuralnet

“Simple” r package enabled insights about complex biochemical processes



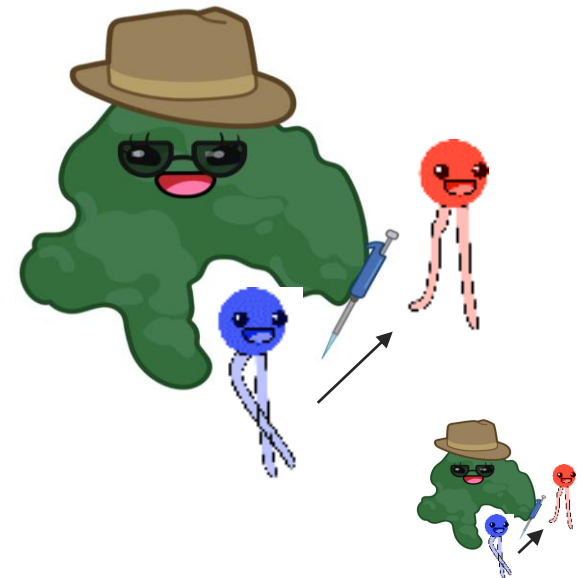
Overview

1. Cardiolipin diversity and the cellular PL pool

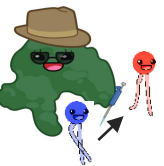
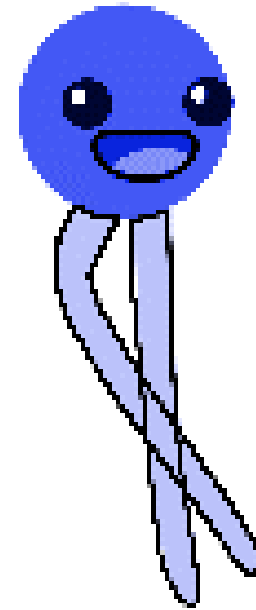
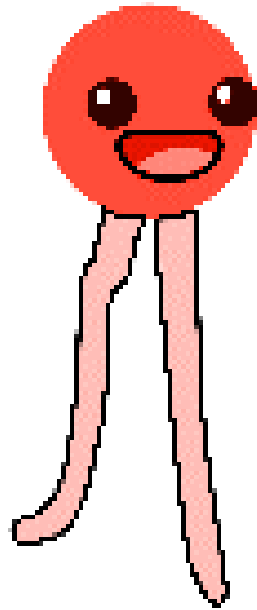
`library(neuralnet)`

2. LC-IM-MS analysis of alkyl- and alkenyl-lipids

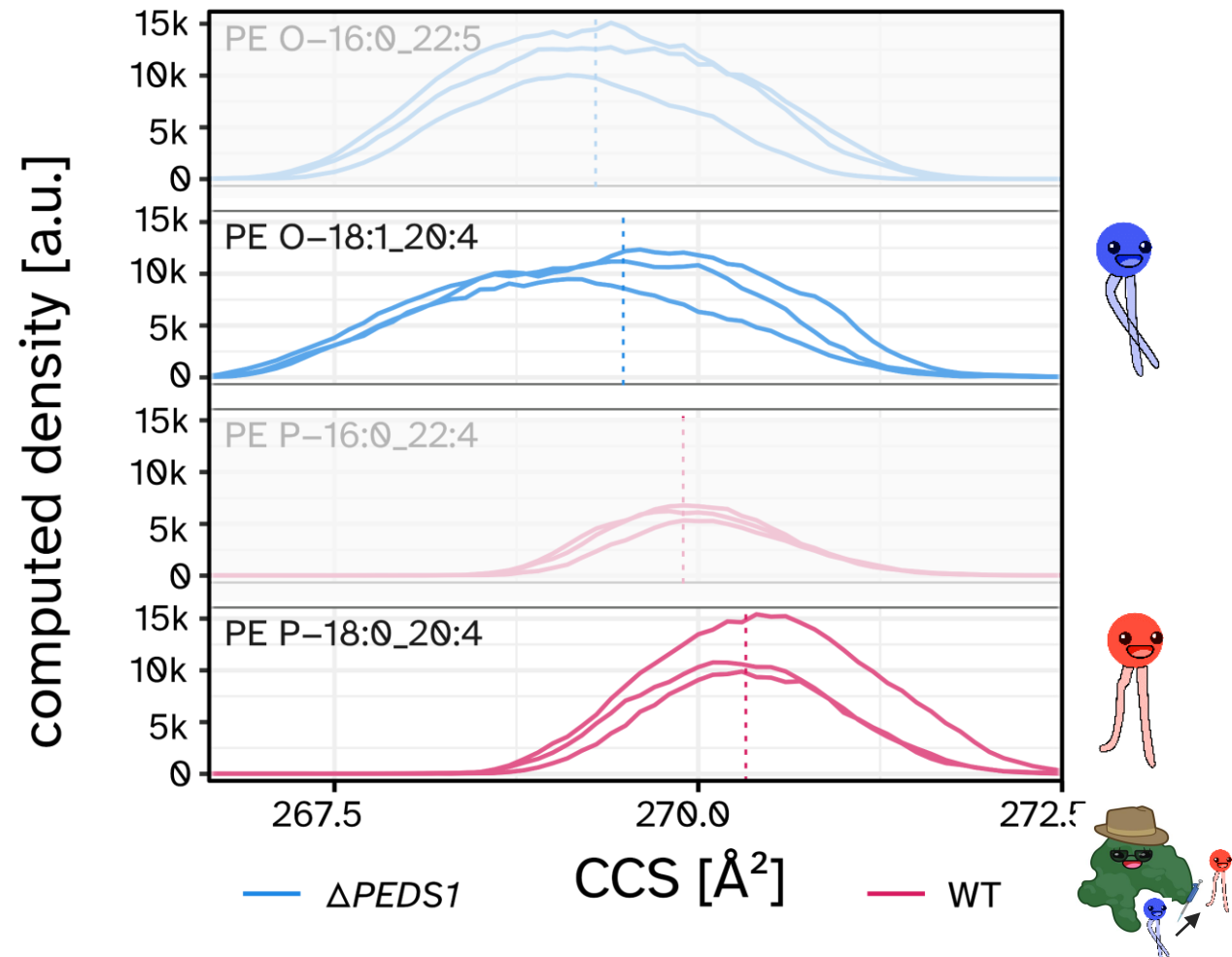
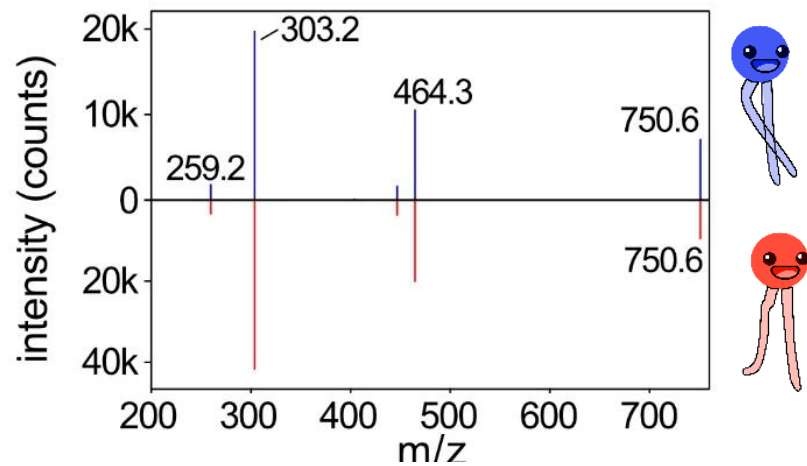
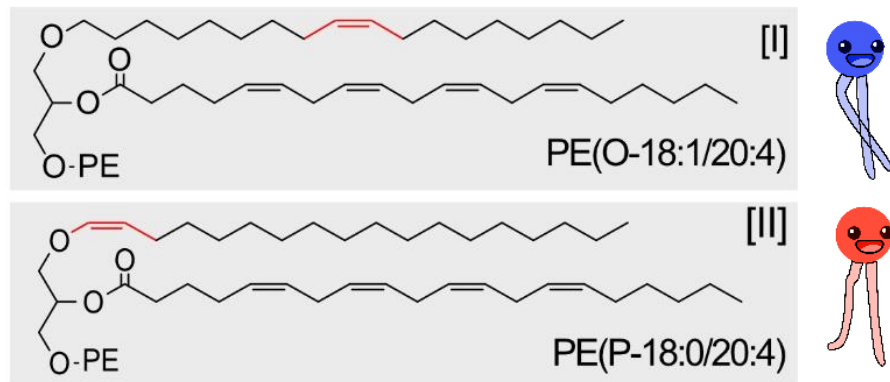
`library(timsr)`



Plasmalogens & other ether lipids



Isomeric ether lipid characteristics



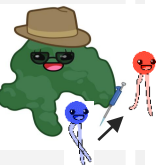
New MS new data format and Mobility

Github: OpenTIMS

[TimsR](#)

TimsPy

```
24 ▾ ``{r download_brukerddl,include=FALSE,eval=FALSE}
25 accept_Bruker_EULA_and_on_Windows_or_Linux = TRUE
26 #
27 ▾ if(accept_Bruker_EULA_and_on_Windows_or_Linux){
28     folder_to_store_proprietary_code = here()
29     path_to_bruker_dll =
30     download_bruker_proprietary_code(folder_to_store_proprietary_code)
31     setup_bruker_so(path_to_bruker_dll)
32     all_columns = c('frame','scan','tof','intensity','mz','inv_ion_mobility',
33     'retention_time')
34 ▾ } else {
35     all_columns = c('frame','scan','tof','intensity','retention_time')
36     }
```

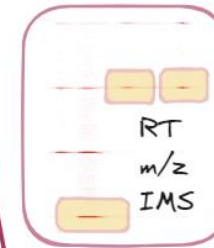


Further data processing

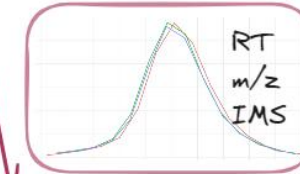
```
pprint(query(D, frames=c(1,5,67), col
#   frame scan   tof intensity
# 1      1    33 312260          9 1174.
# 2      1    34 220720          9  733.
# 3      1    34 261438          9  916.
# 4      1    36  33072          9  152.
# 5      1    36 242110          9  827.
# 6      1    38 204868         62  667.
#
```

Data preparation

1. Readout of initial data - 05



1.1. QC plotting of initial data - 04



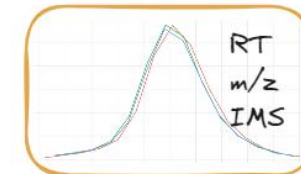
1.2. create sample wise lipid list - 36

	A	B	C	D	E	F
1	lipid	mz	rt_min	rt_max	Formula	ESneg
2	late_peak_1	750.58	8.55	9.4		750.58
3	late_peak_2	762.58	9.625	10.4		760.59
4	late_peak_3	762.61	10.8	11.55		762.51

2. Readout of final data - 05



2.1 QC plotting of final data - 04

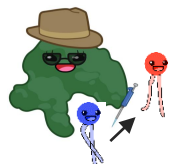


2.2 Calibrate IMS (\bar{x}) to CCS - 13

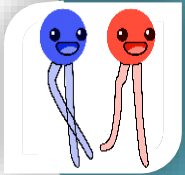
inv_ion_mobility -> CCS

Export

3. Filter \bar{x} data and export - 17



FACTS II



Distinguishing plasmany and plasmenyl lipids by LC-MS
is not trivial



Ion mobility offers additional information to aid distinction

	frame	scan	tof
1	1	33	312260
2	1	34	220720
3	1	34	261438
4	1	36	33072
5	1	36	242110
6	1	38	204868

`library(timsr)` offers direct access to bruker .d raw data in R



Institute of Human Genetics – Biochemical Genetics Laboratory



Markus A. Keller
Yvonne Wohlfarter
Sabrina Sailer
Tereza Vlasáková
Vanessa Rechner
Jan C. Cremer
Nina Weidacher
Viktorija Juric
Jessica Popottnigg



In collaboration with:



Institute of
Biological Chemistry

Katrin Watschinger
Ernst R. Werner
Katharina Lackner



Department of Basic Sciences
in Engineering Science

Lukas Neumann



Bruker
Daltonics
Bremen

Funded by:





Phospholipid Acyl Chain Diversity Controls the Tissue-Specific Assembly of Mitochondrial Cardiolipins

Gregor Oemer,^{1,7} Jakob Koch,^{1,7} Yvonne Wohlfarter,¹ Mohammad T. Alam,² Katharina Lackner,³ Sabrina Sailer,³ Lukas Neumann,⁴ Herbert H. Lindner,⁵ Katrin Watschinger,³ Markus Haltmeier,⁶ Ernst R. Werner,³ Johannes Zschocke,¹ and Markus A. Keller^{1,8,*}

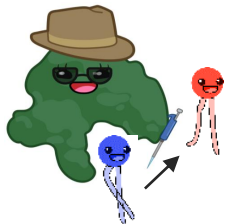
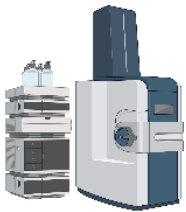
ANN Dataset 10.17632/w4vcz2434r.1 /Supplemental Dataset 8

Manuscript under preparation for submission to

bioRxiv

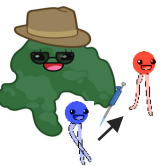
Koch et. al. 2024

Supplemental Dataset incl. RAW data & code
DOI: 10.5281/zenodo.11143478 under preparation



Readout() 1

```
9 #' Readout function
10 #'
11 #' This function provides the code basis to query for MS1 features in timSTOF Pro .tdf and .tdf_bin
12 #' The samples of interest are passed via the processing_L parameter, and in the end it should return
13 #'
14 #' @param processing_L This should be an `.xlsx` loaded in the previous chunk (originating from a
15 #' @param tol The tolerance +- used for feature extraction in m/z dimension.
16 #' @param CAL_run Boolean indicating if this is a calibration readout (m/z), in order to not
17 #' @param identifier A character input which defines the unique output filename
18 #' @param feature_list The name of an `.xlsx` file located in your current project directory, containing
19 #' @param mz_mean A boolean value defining if a column named `.mz_mean` should be used as `.mz`
20 #'
21 #' @return A list containing `.keep_data` and `.keep_data_mz` objects which ii) define the integration
22 readout <- function(processing_L, tol, CAL_run, feature_list, identifier, mz_mean = FALSE) {
23
24   ## setup packages etc. ----
25
26   # library(opentimsr)
27   #
28   library(timsr)
29   library(tidyverse)
30   library(here)
31
32   ## linux download code:
33   # # Do you want to have access only to 'Sample' 'Sample' 'Time of Flight' and 'Identified'?
```



Readout() 2

```
88 #' in a lipid wise manner iterate through the different lipids (at the moment sec
89 for (a in seq_len(nrow(sample_lipidlist))) {
90     lipid = sample_lipidlist$lipid[a]
91     # print(lipid)
92     mz <- ifelse(test = mz_mean == TRUE, yes = sample_lipidlist$mz_mean[a], no = san
93
94     rt_min = sample_lipidlist$rt_min[a]
95     rt_max = sample_lipidlist$rt_max[a]
96     mz_min = mz-tol
97     mz_max = mz+tol
98     lipid
99     mz_min;mz_max
100    rt_min;rt_max
101
102    keep_data <- RT_mz_filter_static2(D = D, rt_min = rt_min, rt_max = rt_max, mz_min
103    keep_data_mz <- integration_window_saver(sample = sample, lipid = lipid, mz = mz,
104    # summary(keep_data$retention_time)
105    # # print(a)
106 }
107 print(str_c(i, "/", nrow(processing_L), " filename: ", sample))
108 # print(nrow(keep_data))
109 # print(nrow(keep_data_mz))
110 }
111 tictoc::toc()
112
```

