Lipidomics Data Analysis Workshop

September 23, 2021

UNT BIODISCOVERY INSTITUTE

Outline

- Overview of lipidomics methods
  - Extraction considerations & internal standards
  - Two different HPLC-MS/MS methods
    - Polar lipids method
    - Neutral lipids method
- Data analysis workflow
  - MRM generation (before analysis)
  - View chromatograms in Analyst
    - Identifying correct TAG-FA MRM transitions
  - Integrating HPLC peaks in Multiquant
  - Formatting data table for R
  - Formatting R script for analysis
  - Troubleshooting

Lipidomics Workflow — from extraction to analysis

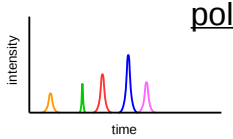
MTBE Extraction

OR

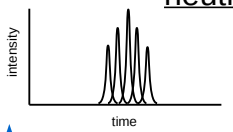
+ internal stds

HPLC-MS/MS

polar lipids



neutral lipids



Data Analysis

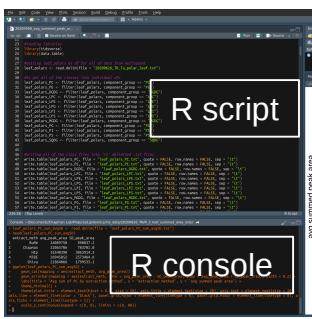
R + RStudio

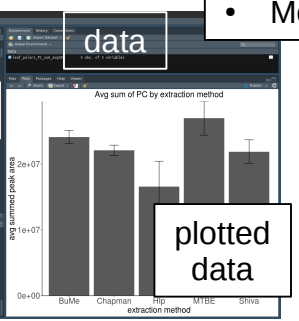
Data table → Class-specific data tables → detected molecular species; internal standards

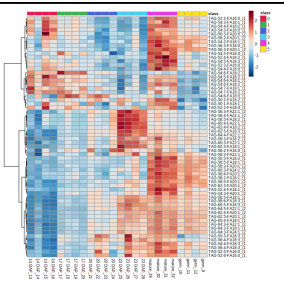
normalization by IS & tissue wt → average & SD of **molecular species**

sum of molecular species → average & SD of **summed species**

- processed data table
- plotted data
- Metaboanalyst-ready table







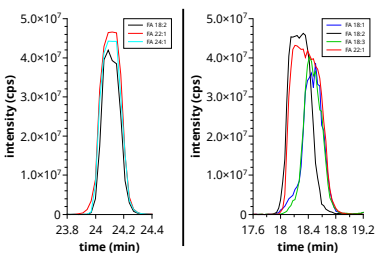
TAG MRMs

python

FA1 / FA1 / FA1
FA1 / FA1 / FA2
FA1 / FA2 / FA2
… … …

TAG specific steps

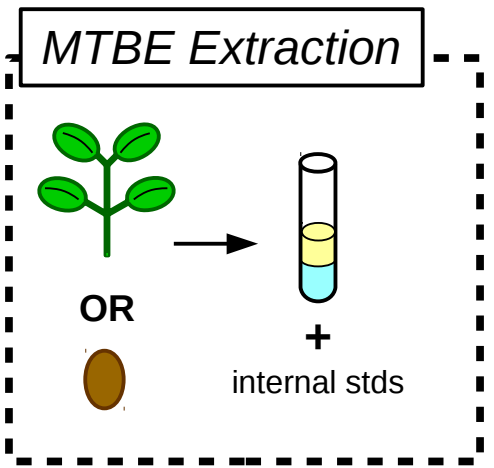
TAG-FA review for quantification



selected TAG MRMs

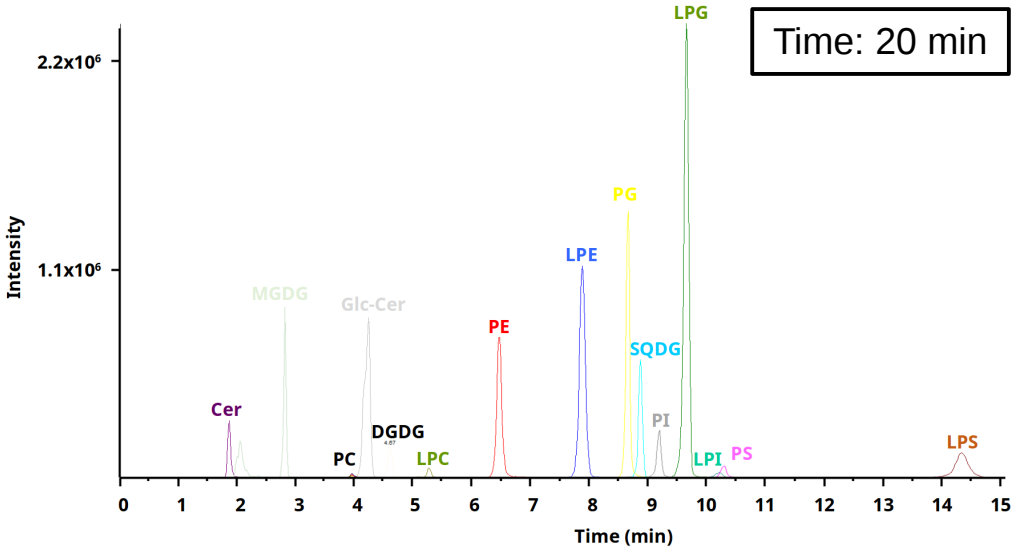
Extraction considerations & internal standards

- Which extraction protocol to use?
  - Either the MTBE or IPA/$CHCl_3$ (i.e. “Chapman method”)
  - Some considerations:
    - Resuspend in $CHCl_3$/MeOH (1:1) solvent for Chapman method
    - Extract at room temperature for MTBE method
    - Possibility to normalize by protein concentration with MTBE method(?)
- Which internal standards and how much?
  - UltimateSPLASH ONE (Avanti Polar Lipids, ca. no. 330820)
    - 10 µL per ~10-50 mg tissue extracted
    - Conc. in mix = 25 – 150 µg/mL
    - Amt. added = 0.3 – 1.5 ng (between 0.2 – 2 nmol)
  - If using other internal standards
    - Aim for similar amounts as above for the specific lipid class
    - Use at least 1 per lipid class of interest (ideally 2-3 covering different FA lengths and degrees of unsaturation)



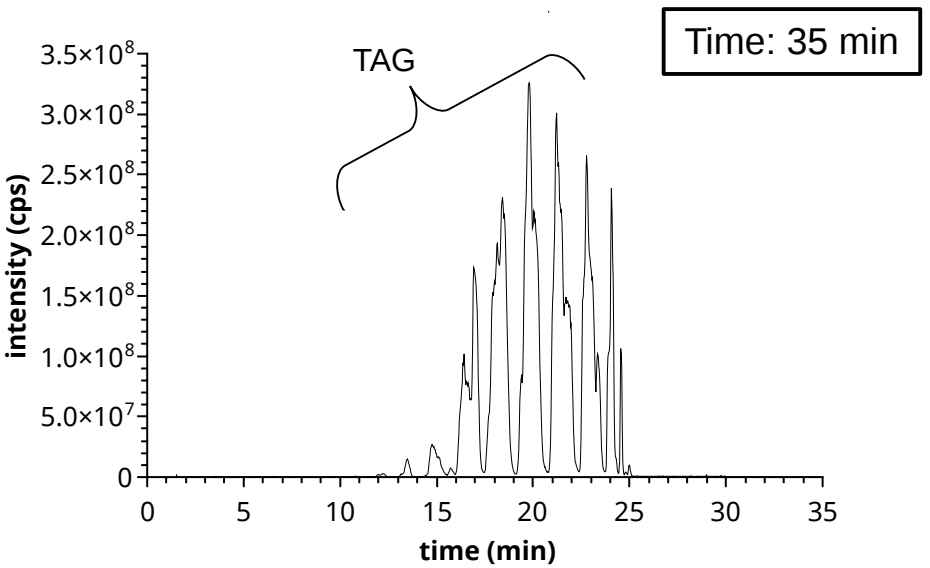
Two different HPLC-MS/MS methods

Polar lipids



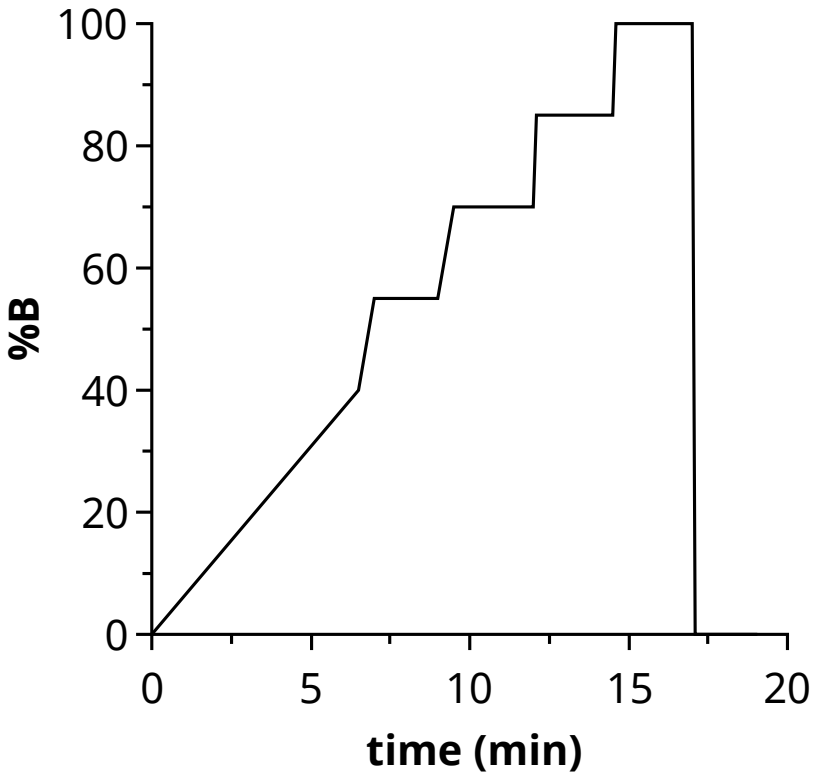
- Solvents:
  - A) AcN / H_2O / Hex (92:6:2) + 2 mM NH_4Ac
  - B) AcN / H_2O (50:50) + 2 mM NH_4Ac
- Column: Phenomenex NH_2 Luna Column (3 µm, 100 Å, 150 x 4.6 mm)
- Lipid species: PC, PE, PI, PG, PS, lyso species, MGDG, DGDG, SQDG

Neutral lipids



- Solvents:
  - A) ACN / H_2O (60:40) + 10 mM NH_4COOH + 0.1% CHOOH
  - B) IPA / ACN / H_2O (90:10:1) + 10 mM NH_4COOH + 0.1% CHOOH
- Column: ThermoFisher Accucore Core-shell C30 silica (2.6 µm, 150 Å, 150 x 2.1 mm)
- Lipid species: MAG, DAG, TAG

Polar lipids method and parameters



Solvents: A) AcN / H_2O / Hex (92:6:2) + 2mM NH_4Ac

B) AcN / H_2O (50:50) + 2mM NH_4Ac

Column: Phenomenex NH_2 Luna Column

Temp: 25°C

Flow rate: 1 mL/min

Inj. vol.: 5 µL

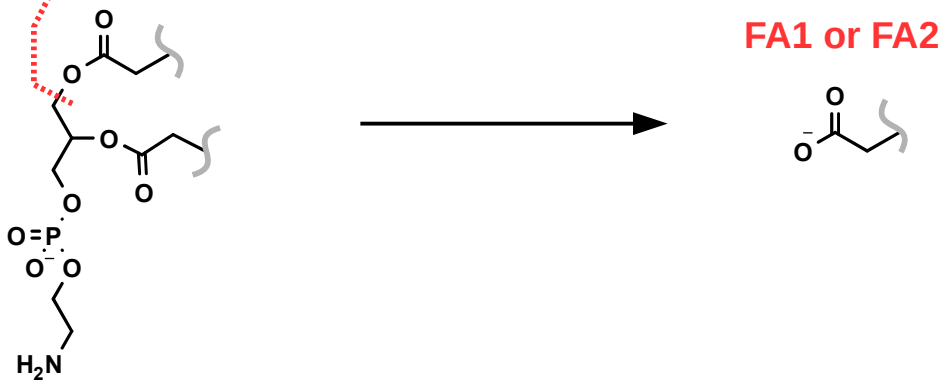
Polarity: negative

Scan: *scheduled* MRM

Polar lipids general fragmentation and MRM transition

[M – H]$^-$

Q3 = [FA – H]$^-$

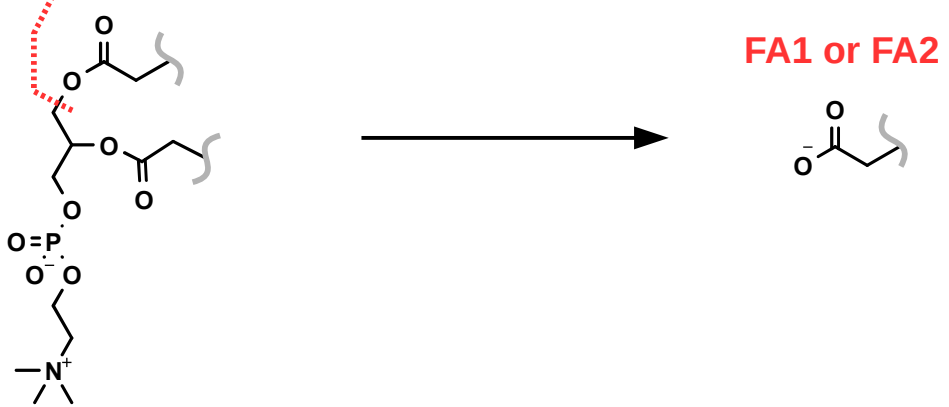


FA1 or FA2

PE

[M + Ac]$^-$

Q3 = [FA – H]$^-$



FA1 or FA2

Ac$^-$ +

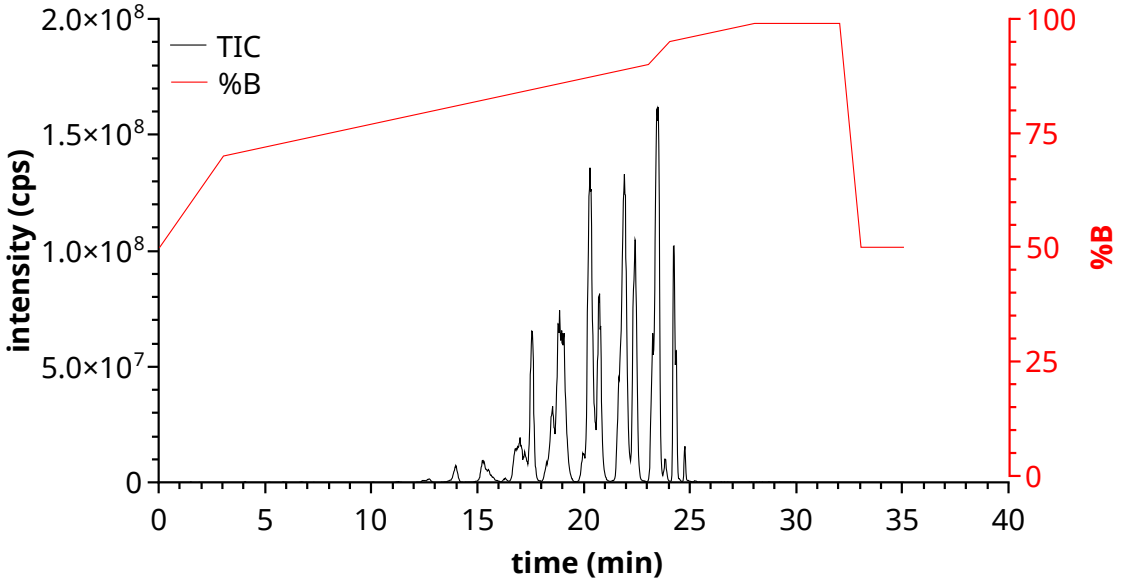
PC

<u>Parent ion (Q1):</u> [M – H]$^-$
(exception PC and LPC as [M + Ac]$^-$)

<u>Fragment ion (Q3):</u> [FA – H]$^-$

Note: Only one of the FAs is detected

Neutral lipids method and parameters



Solvents: A) AcN / H_2O (60:40)
+ 10mM NH_4CHOOH + 0.1% CHOOH

B) IPA / AcN / H_2O (90:10:1)
+ 10mM NH_4CHOOH + 0.1% CHOOH

Column: ThermoFisher Accucore C30 Core Shell

Temp: 40°C

Flow rate: 0.35 mL/min

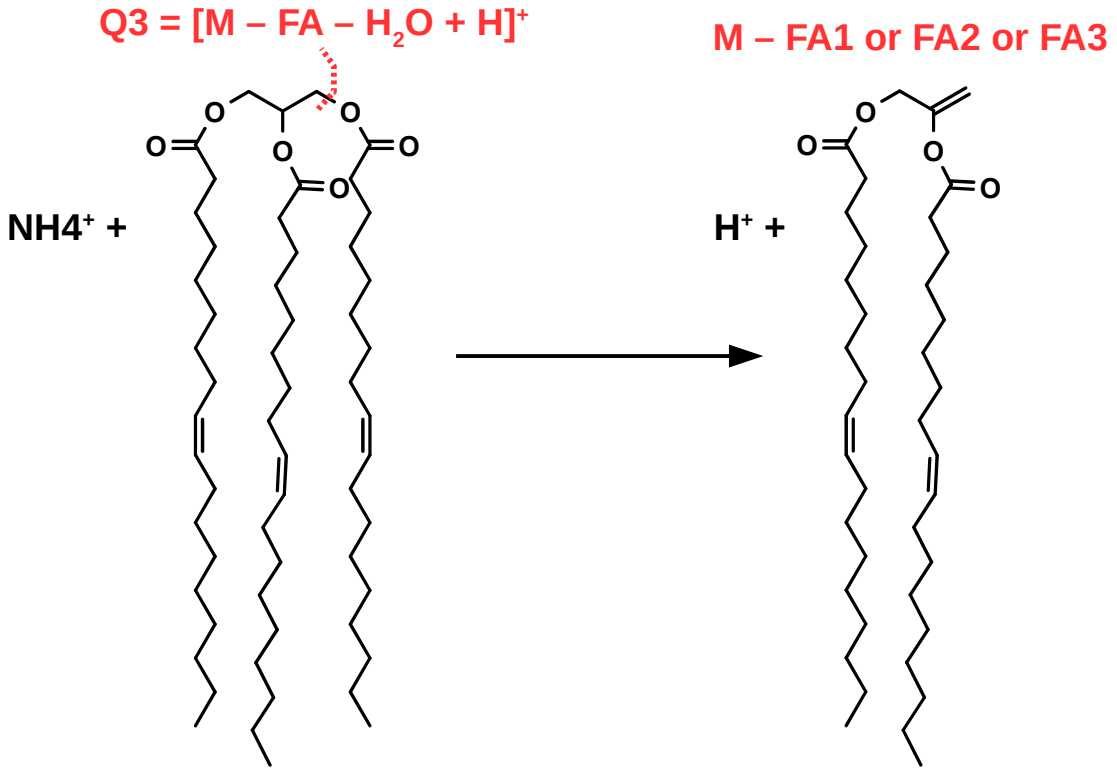
Inj. vol.: 5 µL

Polarity: positive

Scan: MRM

Neutral lipids general fragmentation and MRM transition

$[M + NH_4]^+$



Parent ion (Q1): $[M + NH_4]^+$

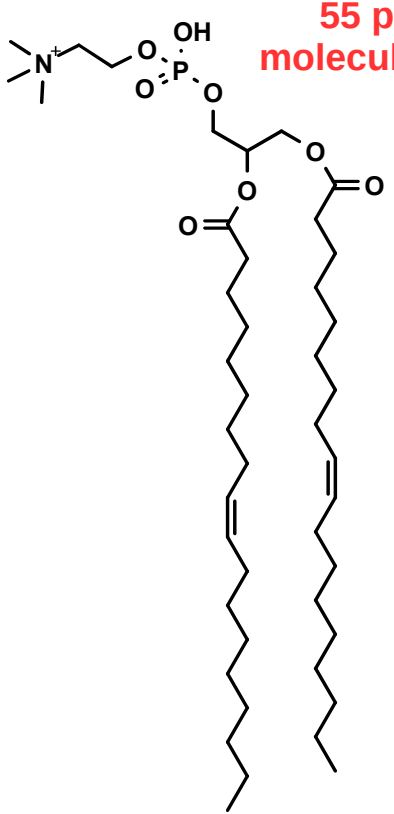
Fragment ion (Q3): $[M – FA – H_2O + H]^+$
(exception MAG as $[FA – H_2O + H]^+$)

Note: The loss of each *potential* FA is detected

MRM generation and the loss of each *potential* FA in TAG

10 fatty acids = [16:0, 18:3, 18:2, 18:1, 18:0, 20:2, 20:1, 20:0, 22:1, 22:0] *Arabidopsis thaliana*

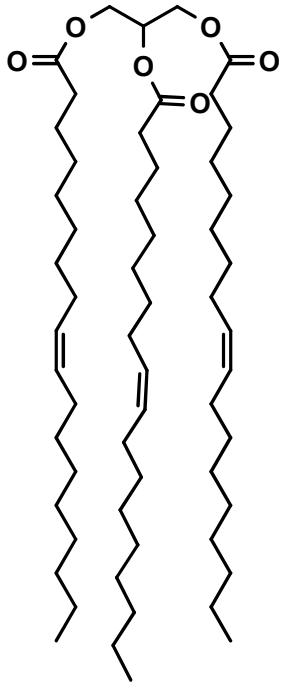
55 possible molecular species



$$P = \frac{(n+r-1)!}{r!(n-1)!}$$

P = possible combinations
n = number of FAs
r = number of FAs per molecule

220 possible molecular species



MRM generation and the loss of each *potential* FA in TAG

My first attempt…

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	AM	AN	AO	AP
1	Q1/Q3	TAG		Q3 FA		remaining			FA2							des 1										
2		C	des	C	des	C	des		16	18	20	22	24	26		0	1	2	3			FA1	TAG species	overlap 1	overlap 2	overlap 3	ove
3	TAG-50:3-FA16:3	50	3	16	3	34	0		18	16	14	12	10	8		0	-1	-2	-3		TAG-50:3	16:3	16:3/16:0/18:0				
4	TAG-50:3-FA16:2	50	3	16	2	34	1		18	16	14	12	10	8		1	0	-1	-2		TAG-50:3	16:2	16:2/16:0/18:1	16:2/16:1/18:0			
5	TAG-50:3-FA16:1	50	3	16	1	34	2		18	16	14	12	10	8		2	1	0	-1		TAG-50:3	16:1	16:1/16:0/18:2	16:1/16:1/18:1	16:1/16:2/18:0		
6	TAG-50:3-FA16:0	50	3	16	0	34	3		18	16	14	12	10	8		3	2	1	0		TAG-50:3	16:0	16:0/16:0/18:3	16:0/16:1/18:2	16:0/16:2/18:1	16:0/16:3/18:0	
7	TAG-50:3-FA18:3	50	3	18	3	32	0		16	14	12	10	8	6		0	-1	-2	-3		TAG-50:3	18:3	18:3/16:0/16:0				
8	TAG-50:3-FA18:2	50	3	18	2	32	1		16	14	12	10	8	6		1	0	-1	-2		TAG-50:3	18:2	18:2/16:0/16:1				
9	TAG-50:3-FA18:1	50	3	18	1	32	2		16	14	12	10	8	6		2	1	0	-1		TAG-50:3	18:1	18:1/16:0/16:2	18:1/16:1/16:1			
10	TAG-50:3-FA18:0	50	3	18	0	32	3		16	14	12	10	8	6		3	2	1	0		TAG-50:3	18:0	18:0/16:0/16:3	18:0/16:1/16:2			
11	TAG-50:2-FA16:2	50	2	16	2	34	0		18	16	14	12	10	8		0	-1	-2	-3		TAG-50:2	16:2	16:2/16:0/18:0				
12	TAG-50:2-FA16:1	50	2	16	1	34	1		18	16	14	12	10	8		1	0	-1	-2		TAG-50:2	16:1	16:1/16:0/18:1	16:1/16:1/18:0			
13	TAG-50:2-FA16:0	50	2	16	0	34	2		18	16	14	12	10	8		2	1	0	-1		TAG-50:2	16:0	16:0/16:0/18:2	16:0/16:1/18:1	16:0/16:2/18:0		
14	TAG-50:2-FA18:2	50	2	18	2	32	0		16	14	12	10	8	6		0	-1	-2	-3		TAG-50:2	18:2	18:2/16:0/16:0				
15	TAG-50:2-FA18:1	50	2	18	1	32	1		16	14	12	10	8	6		1	0	-1	-2		TAG-50:2	18:1	18:1/16:0/16:1				
16	TAG-50:2-FA18:0	50	2	18	0	32	2		16	14	12	10	8	6		2	1	0	-1		TAG-50:2	18:0	18:0/16:0/16:2	18:0/16:1/16:1			
17	TAG-50:1-FA16:1	50	1	16	1	34	0		18	16	14	12	10	8		0	-1	-2	-3		TAG-50:1	16:1	16:1/16:0/18:0				
18	TAG-50:1-FA16:0	50	1	16	0	34	1		18	16	14	12	10	8		1	0	-1	-2		TAG-50:1	16:0	16:0/16:0/18:1	16:0/16:1/18:0			
19	TAG-50:1-FA18:1	50	1	18	1	32	0		16	14	12	10	8	6		0	-1	-2	-3		TAG-50:1	18:1	18:1/16:0/16:0				
20	TAG-50:1-FA18:0	50	1	18	0	32	1		16	14	12	10	8	6		1	0	-1	-2		TAG-50:1	18:0	18:0/16:0/16:1				
21	TAG-52:6-FA16:3	52	6	16	3	36	3		20	18	16	14	12	10		3	2	1	0		TAG-52:6	16:3	16:3/16:2/20:1	16:3/16:3/20:0	16:3/18:0/18:3	16:3/18:1/18:2	
22	TAG-52:6-FA16:2	52	6	16	2	36	4		20	18	16	14	12	10		4	3	2	1		TAG-52:6	16:2	16:2/16:3/20:1	16:2/18:1/18:3	16:2/18:2/18:2		
23	TAG-52:6-FA16:1	52	6	16	1	36	5		20	18	16	14	12	10		5	4	3	2		TAG-52:6	16:1	16:1/18:2/18:3				
24	TAG-52:6-FA16:0	52	6	16	0	36	6		20	18	16	14	12	10		6	5	4	3		TAG-52:6	16:0	16:0/18:3/18:3				
25	TAG-52:6-FA18:3	52	6	18	3	34	3		18	16	14	12	10	8		3	2	1	0		TAG-52:6	18:3	18:3/16:0/18:3	18:3/16:1/18:2	18:3/16:2/18:1	18:3/16:3/18:0	
26	TAG-52:6-FA18:2	52	6	18	2	34	4		18	16	14	12	10	8		4	3	2	1		TAG-52:6	18:2	18:2/16:1/18:3	18:2/16:2/18:2	18:2/16:3/18:1		
27	TAG-52:6-FA18:1	52	6	18	1	34	5		18	16	14	12	10	8		5	4	3	2		TAG-52:6	18:1	18:1/16:2/18:3	18:1/16:3/18:2			
28	TAG-52:6-FA18:0	52	6	18	0	34	6		18	16	14	12	10	8		6	5	4	3		TAG-52:6	18:0	18:0/16:3/18:3				

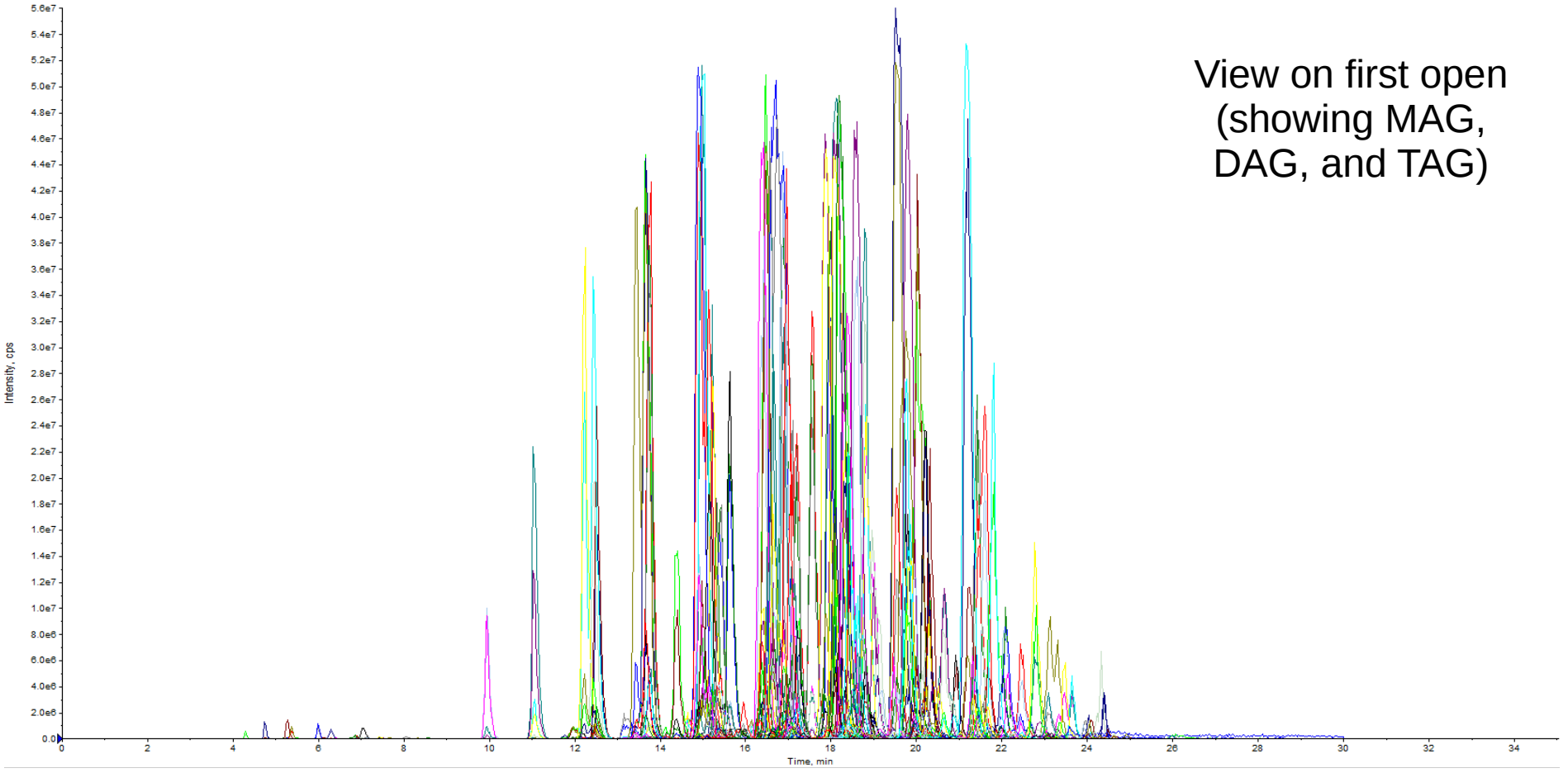
…

MRM generation and the loss of each *potential* FA in TAG

A better approach…

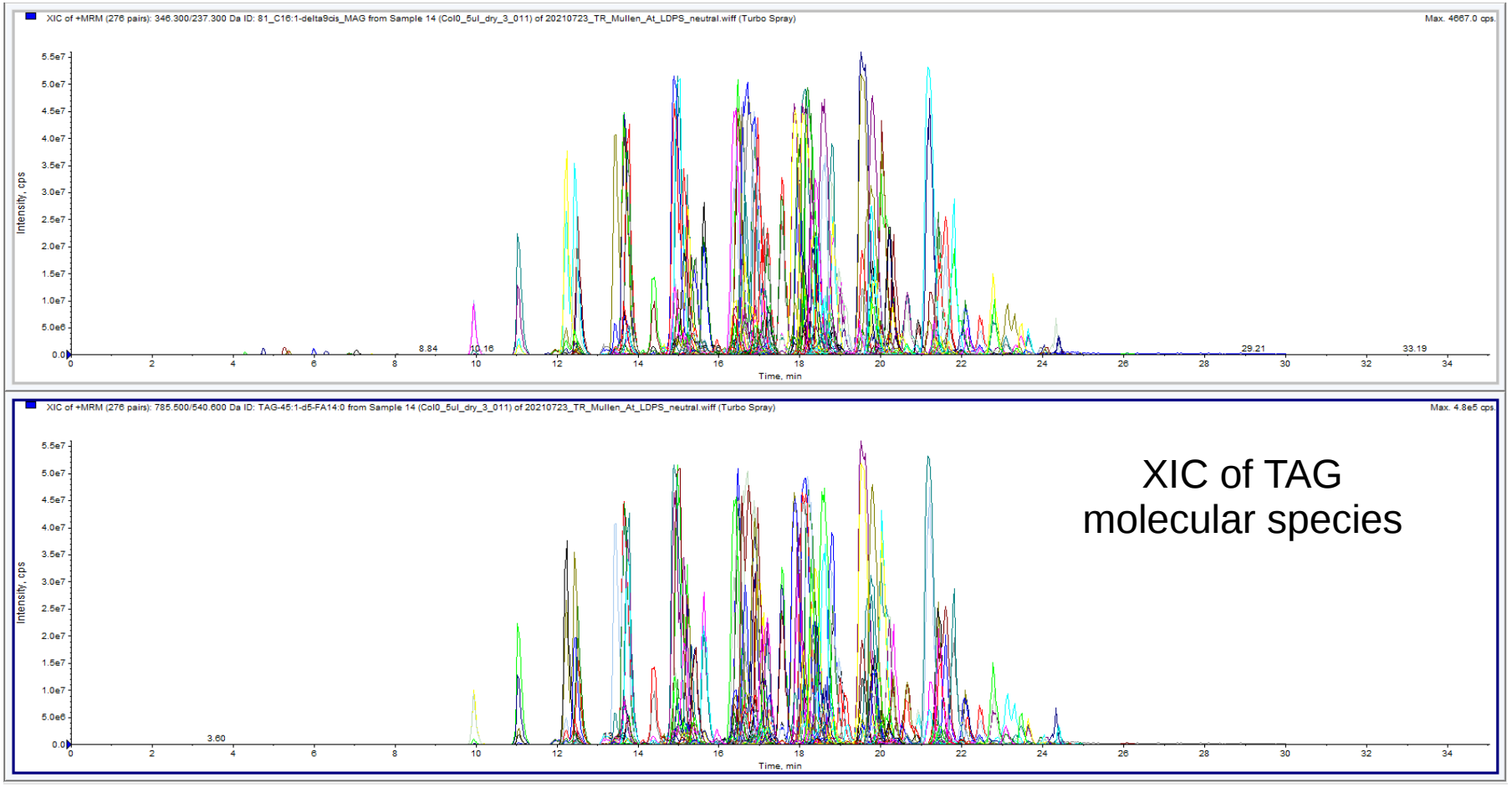
```python
# TAG-FA MRM generator 2.0
import pandas as pd

fatty_acids = ['16:0', '18:3', '18:2', '18:1', '18:0', '20:2', '20:1', '20:0', '22:1', '22:0']      # Arabidopsis https://plantfadb.org/datasets?plant_id=17354


def mrm_output(fa1):

    carbon = 12.000000
    hydrogen = 1.007825
    oxygen = 15.994915
    ammonium = 18.03382555

    fa_carbon_no = []
    fa_desat_no = []

    # converting string to int for calculations below
    for i in fa1:
        fa_carbon_no.append(int(i[:2]))
        fa_desat_no.append(int(i[-1]))



    tag_carbon_mass = []
    tag_desat_mass = []
    final_tag_carbon = []
    final_tag_desat = []

    fa_carbon_no_2 = []
    fa_desat_no_2 = []

    m = len(fa_carbon_no)
    n = len(fa_desat_no)

    # to get the total fa carbon number
    while fa_carbon_no != []:
        for i in fa_carbon_no:
            tag_carbon_mass += [fa_carbon_no[0] + i]

        fa_carbon_no_2.append(fa_carbon_no[0])
        fa_carbon_no.remove(fa_carbon_no[0])


```

1st step in TAG analysis: ID'ing which TAG-FA MRM transitions to use for quantification in Analyst

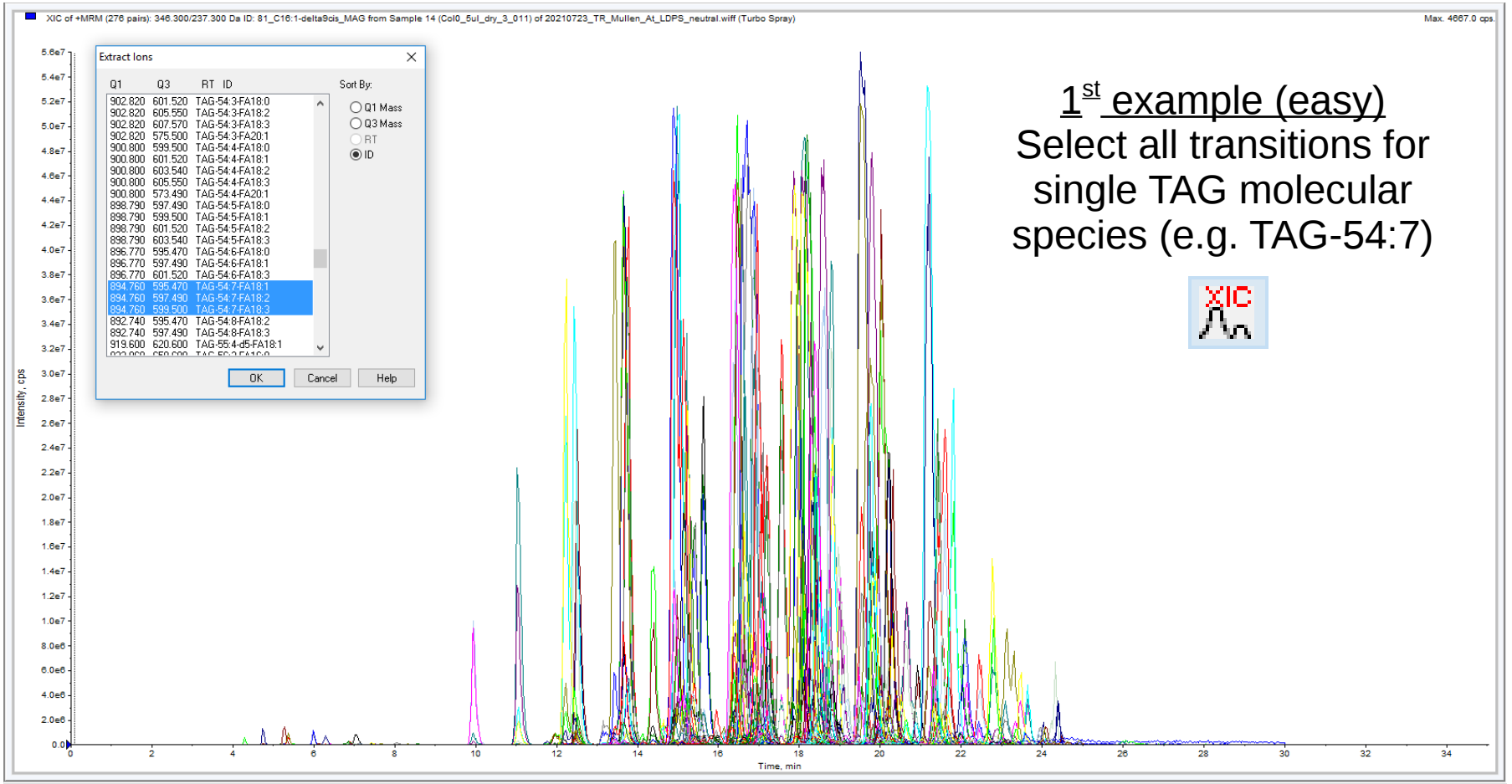


View on first open (showing MAG, DAG, and TAG)

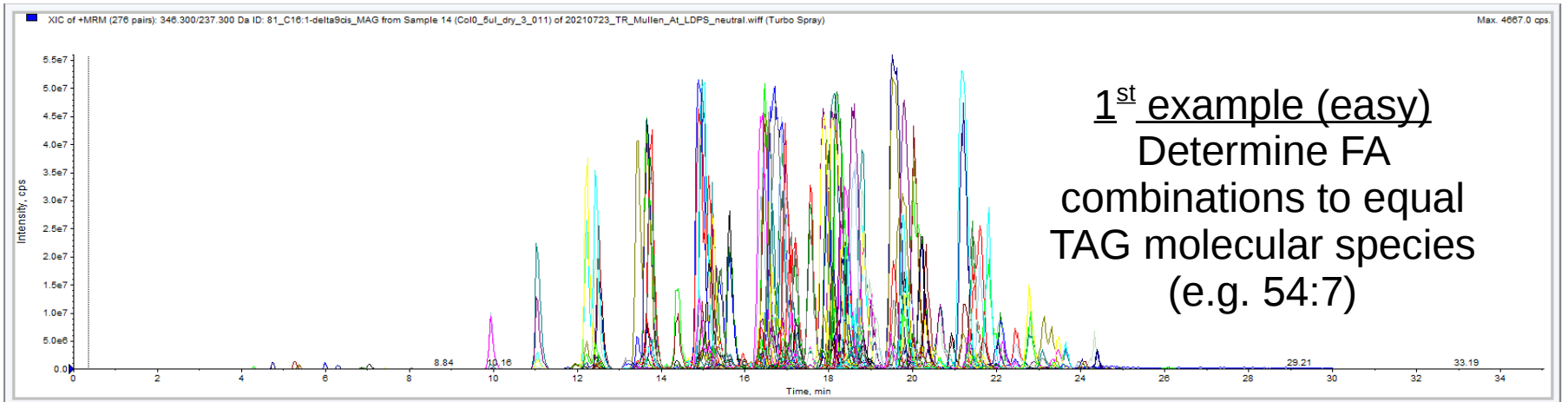
1st step in TAG analysis: ID'ing which TAG-FA MRM transitions to use for quantification in Analyst



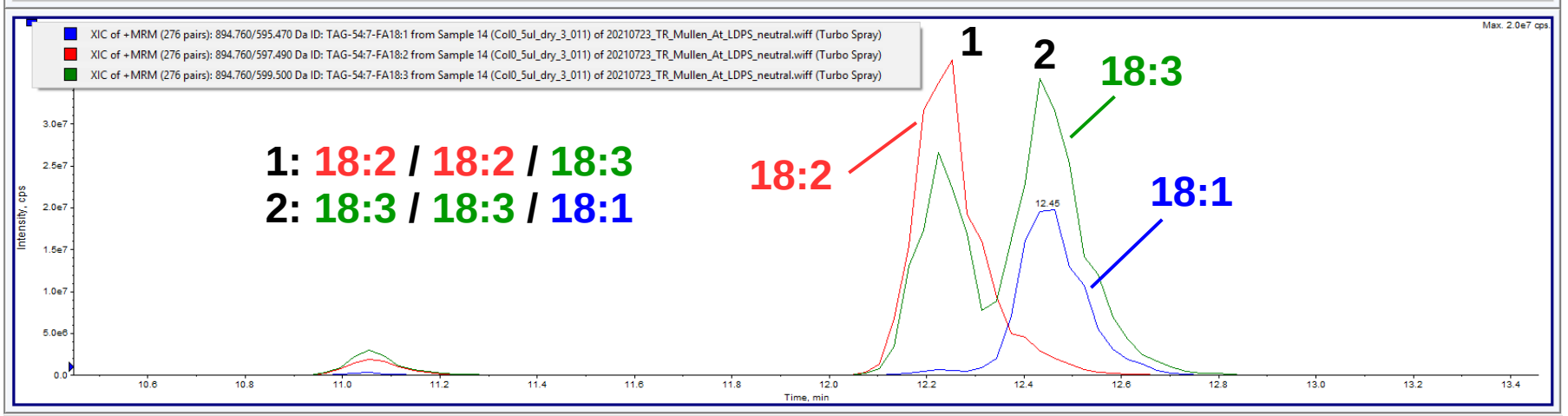
1st step in TAG analysis: ID'ing which TAG-FA MRM transitions to use for quantification in Analyst



1st step in TAG analysis: ID'ing which TAG-FA MRM transitions to use for quantification in Analyst



1st example (easy)
Determine FA combinations to equal TAG molecular species (e.g. 54:7)



1st step in TAG analysis: ID'ing which TAG-FA MRM transitions to use for quantification in Analyst

TAG 54:7
1: 18:2 / 18:2 / 18:3
2: 18:3 / 18:3 / 18:1

1st example (easy)
Record these TAG molecular species and FA combinations (RT is also good to take down)

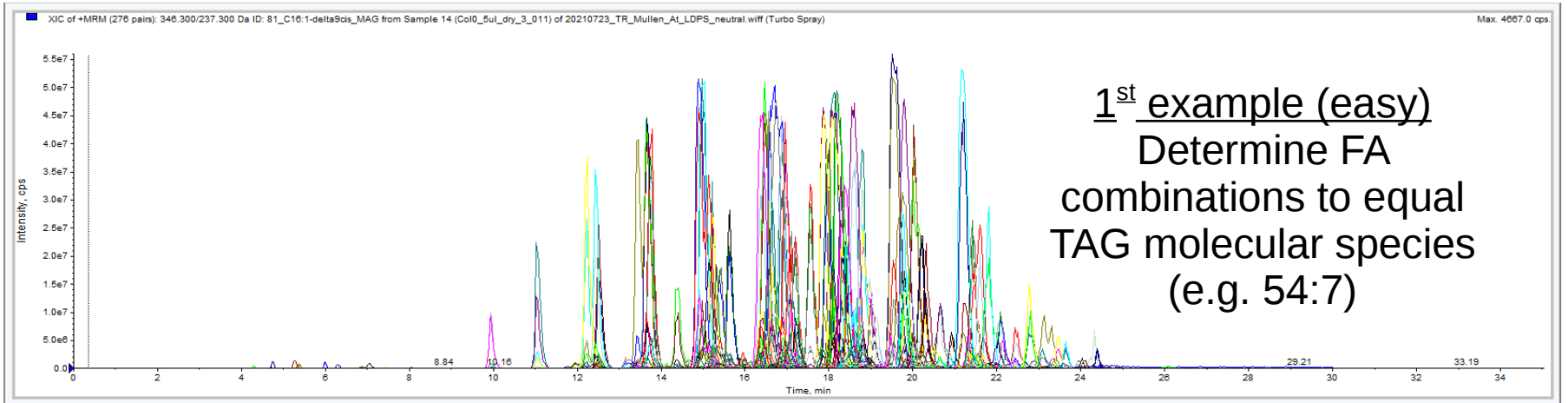
	A	B	C	D	E
1	TG molec sp	FA comp	FA for quant	RT	notes
31	tg-54:7	18:2/18:2/18:3	fa-18:2	12.25	/2
32	tg-54:7	18:3/18:3/18:1	fa-18:1	12.44	

Note this value will be divided by 2

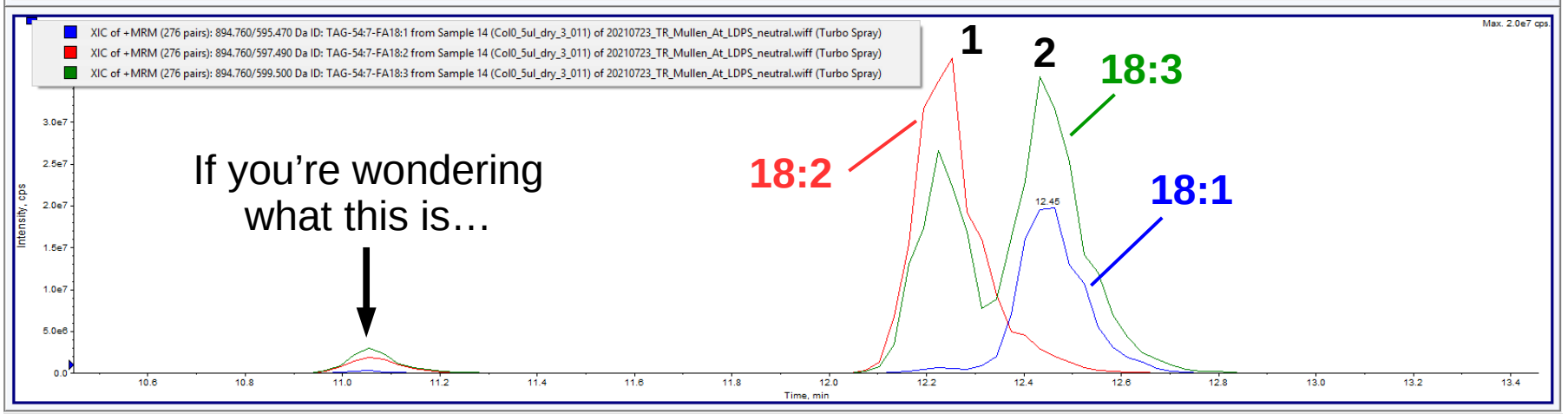
These indicate which MRM transitions to use for quantification:
TAG-54:7-FA18:2
TAG-54:7-FA18:1

TAG-54:7-FA18:3 is not used

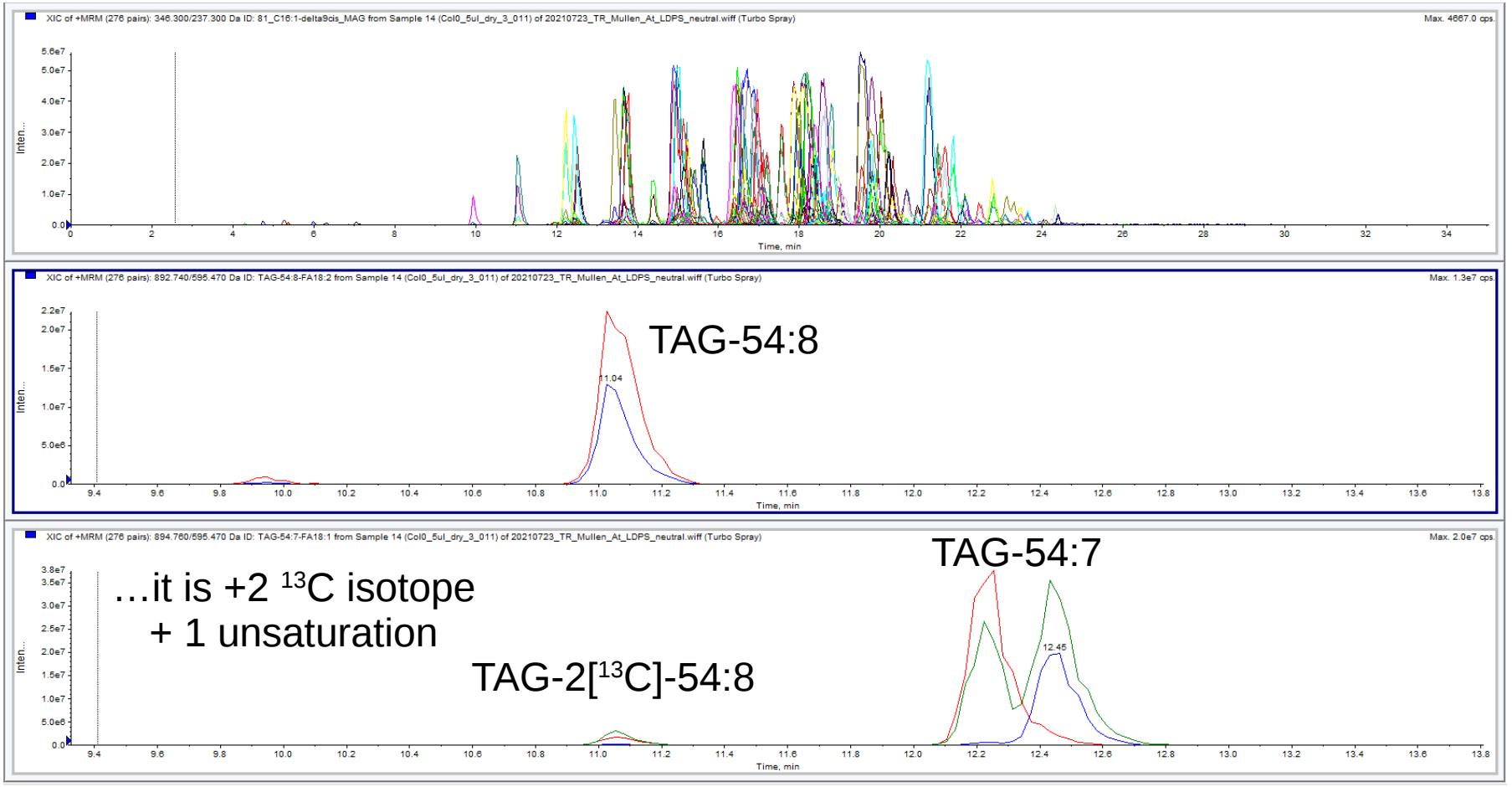
1st step in TAG analysis: ID'ing which TAG-FA MRM transitions to use for quantification in Analyst



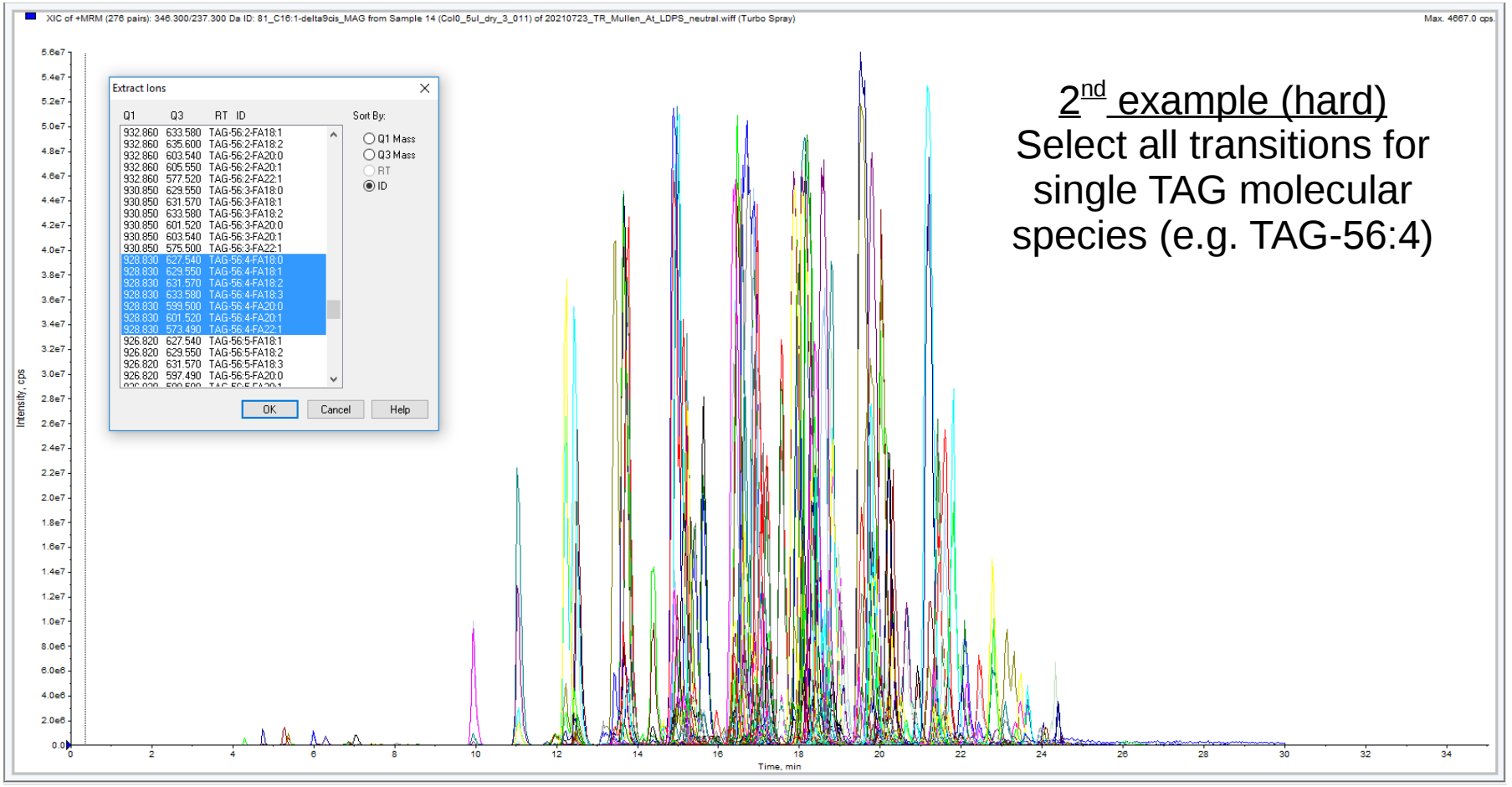
1st example (easy)
Determine FA combinations to equal TAG molecular species (e.g. 54:7)



1st step in TAG analysis: ID'ing which TAG-FA MRM transitions to use for quantification in Analyst

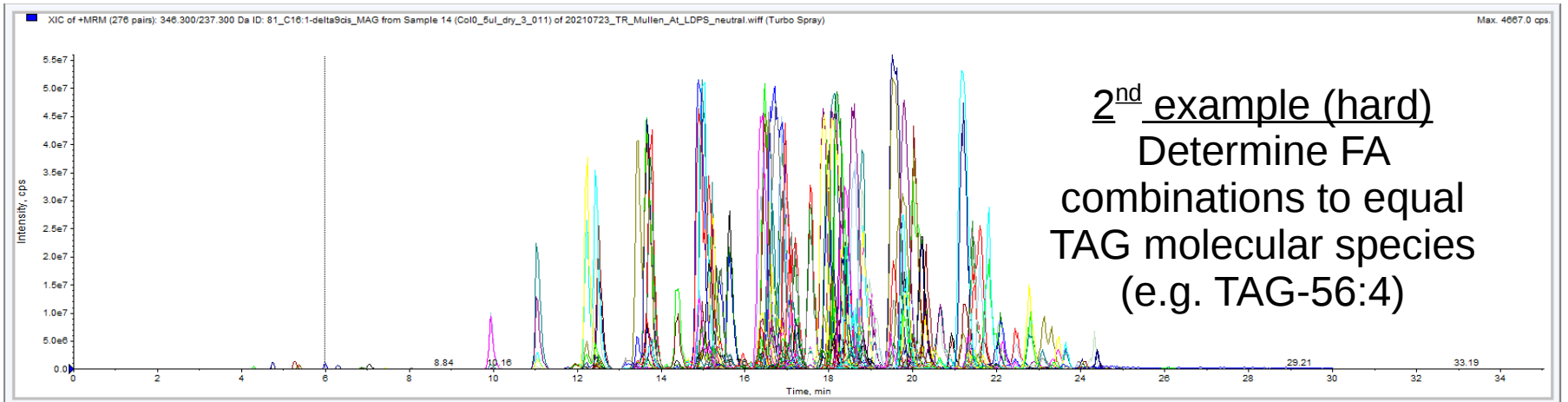


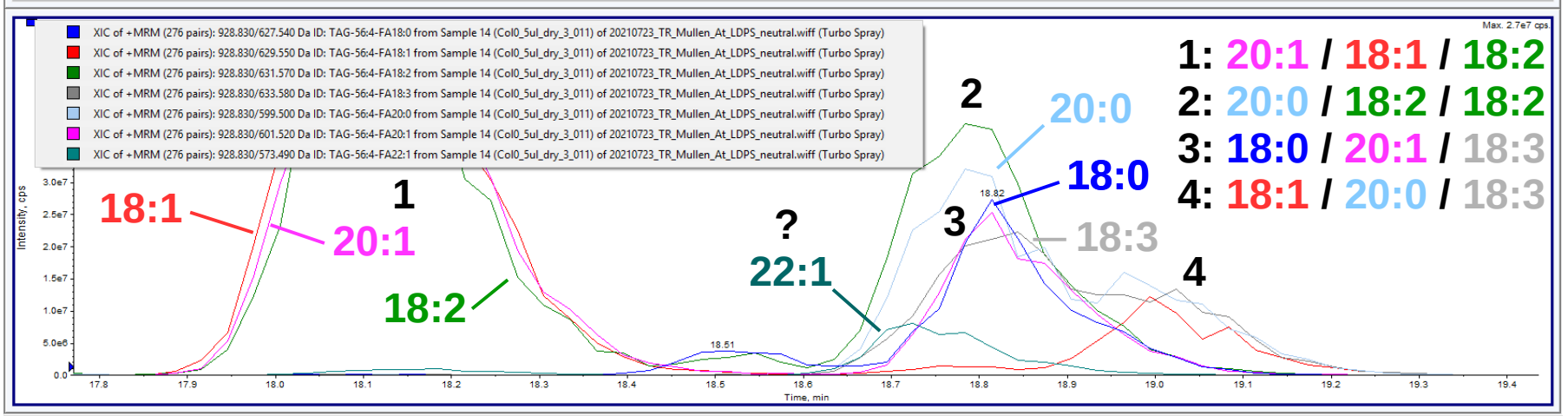
1st step in TAG analysis: ID'ing which TAG-FA MRM transitions to use for quantification in Analyst



2nd example (hard)
Select all transitions for single TAG molecular species (e.g. TAG-56:4)

1st step in TAG analysis: ID'ing which TAG-FA MRM transitions to use for quantification in Analyst





1st step in TAG analysis: ID'ing which TAG-FA MRM transitions to use for quantification in Analyst

TAG 56:4

1: 20:1 / 18:1 / 18:2
2: 20:0 / 18:2 / 18:2
3: 18:0 / 20:1 / 18:3
4: 18:1 / 20:0 / 18:3

2nd example (hard)
Record these TAG molecular species and FA combinations (RT is also good to take down)

	A	B	C	D	E	F
1	TG molec sp	FA comp	FA for quant	RT	notes	
41	tg-56:4	20:1/18:1/18:2	fa-18:1	18.19	1st peak	
42	tg-56:4	20:0/18:2/18:2	fa-18:2	18.79	/2; 2nd peak	
43	tg-56:4	18:0/20:1/18:3	fa-18:0	18.84		
44	tg-56:4	18:1/20:0/18:3	fa-20:0	18.99	/2 (second half of peak)	

MRM transitions used for quantification:

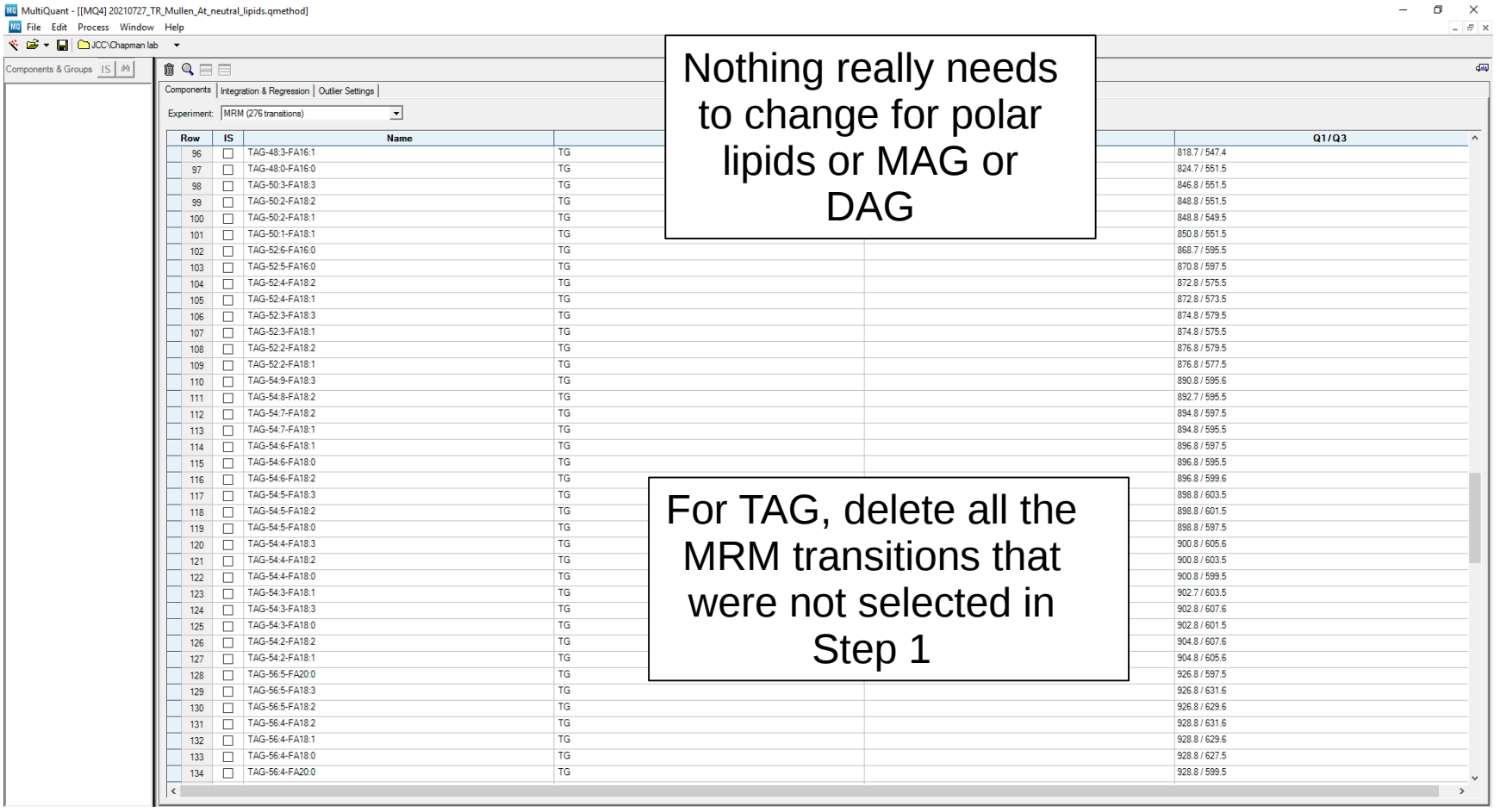
TAG-56:4-FA18:1
TAG-56:4-FA18:2
TAG-56:4-FA18:0
TAG-56:4-FA20:0

MRM transitions **NOT** used for quantification:

TAG-56:4-FA18:3
TAG-56:4-FA20:1
TAG-56:4-FA22:1

2nd step in TAG analysis: Creating Multiquant method to integrate HPLC peaks

(Also 1st step for all the other lipids)

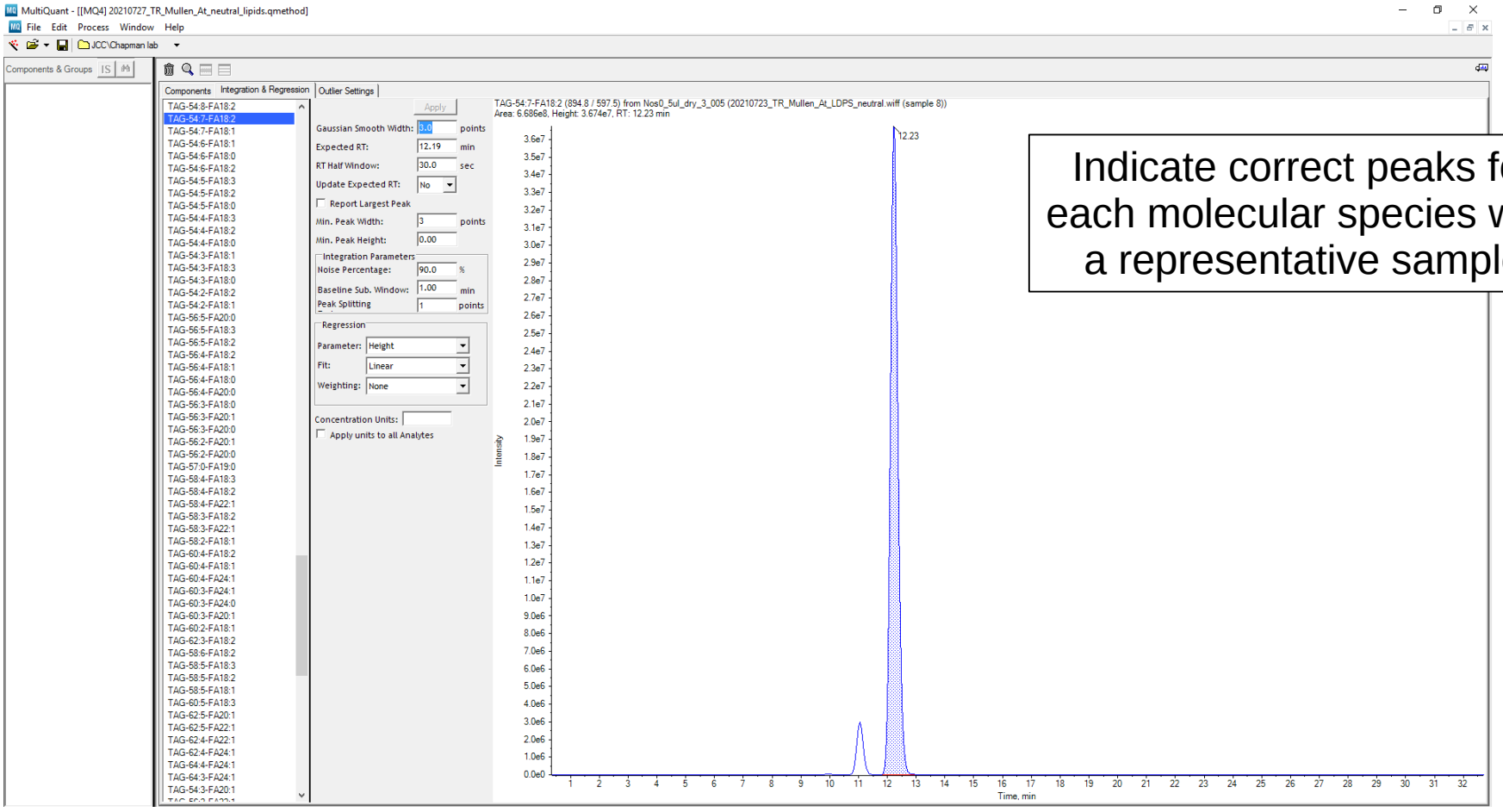


Nothing really needs to change for polar lipids or MAG or DAG

For TAG, delete all the MRM transitions that were not selected in Step 1

2nd step in TAG analysis: Creating Multiquant method to integrate HPLC peaks

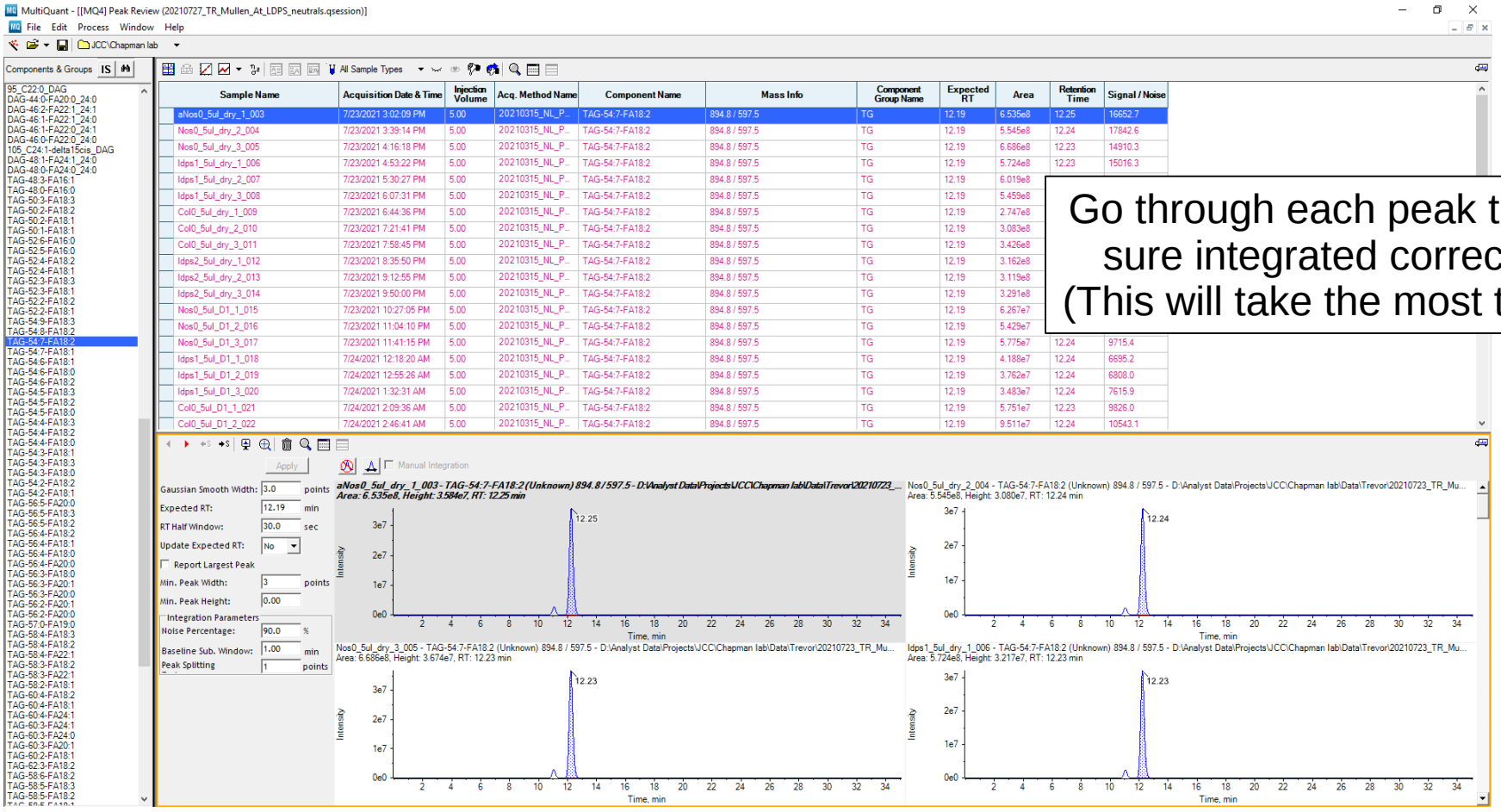
(Also 1st step for all the other lipids)



Indicate correct peaks for each molecular species with a representative sample

2nd step in TAG analysis: Creating Multiquant method to integrate HPLC peaks

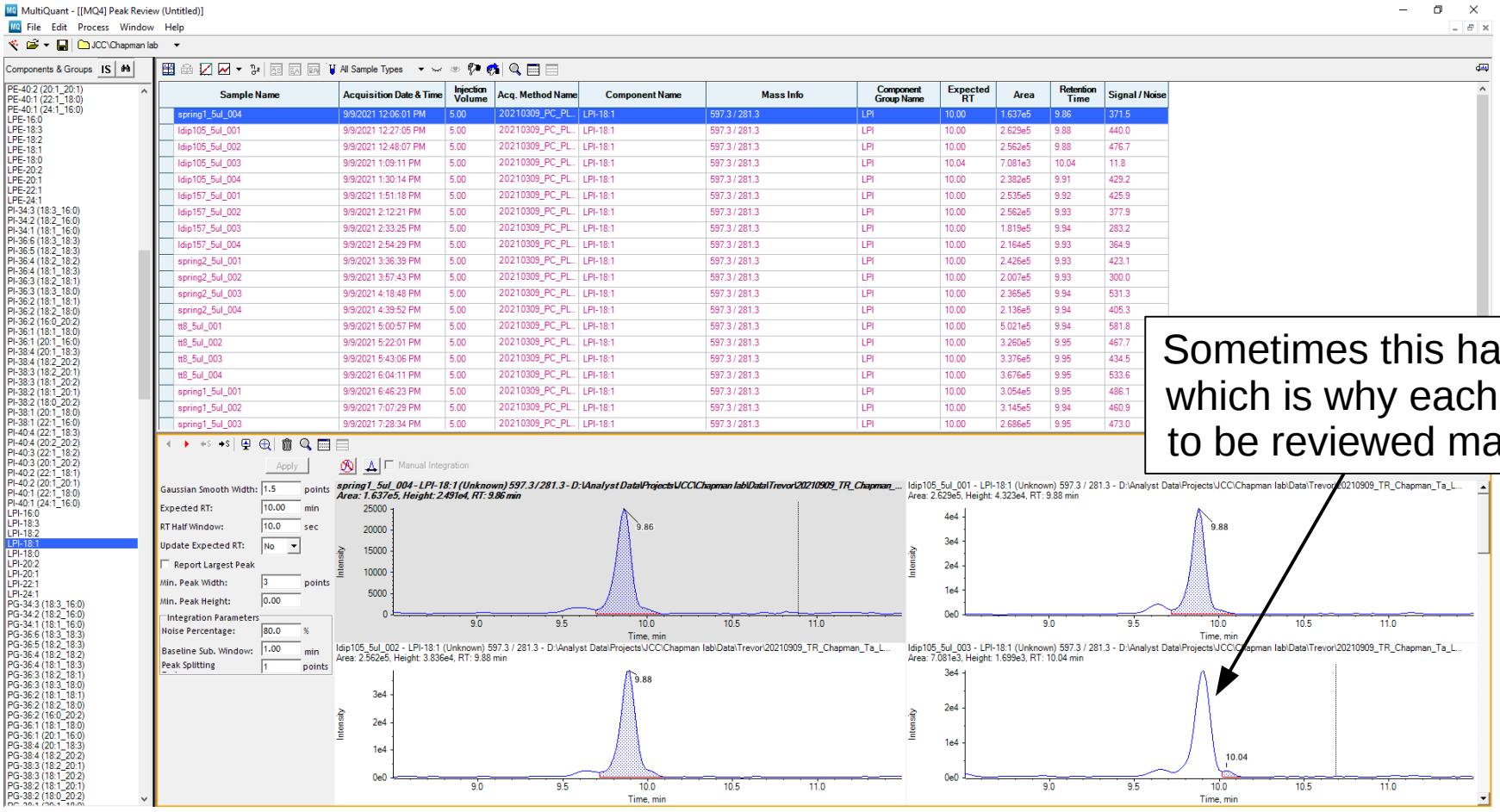
(Also 1st step for all the other lipids)



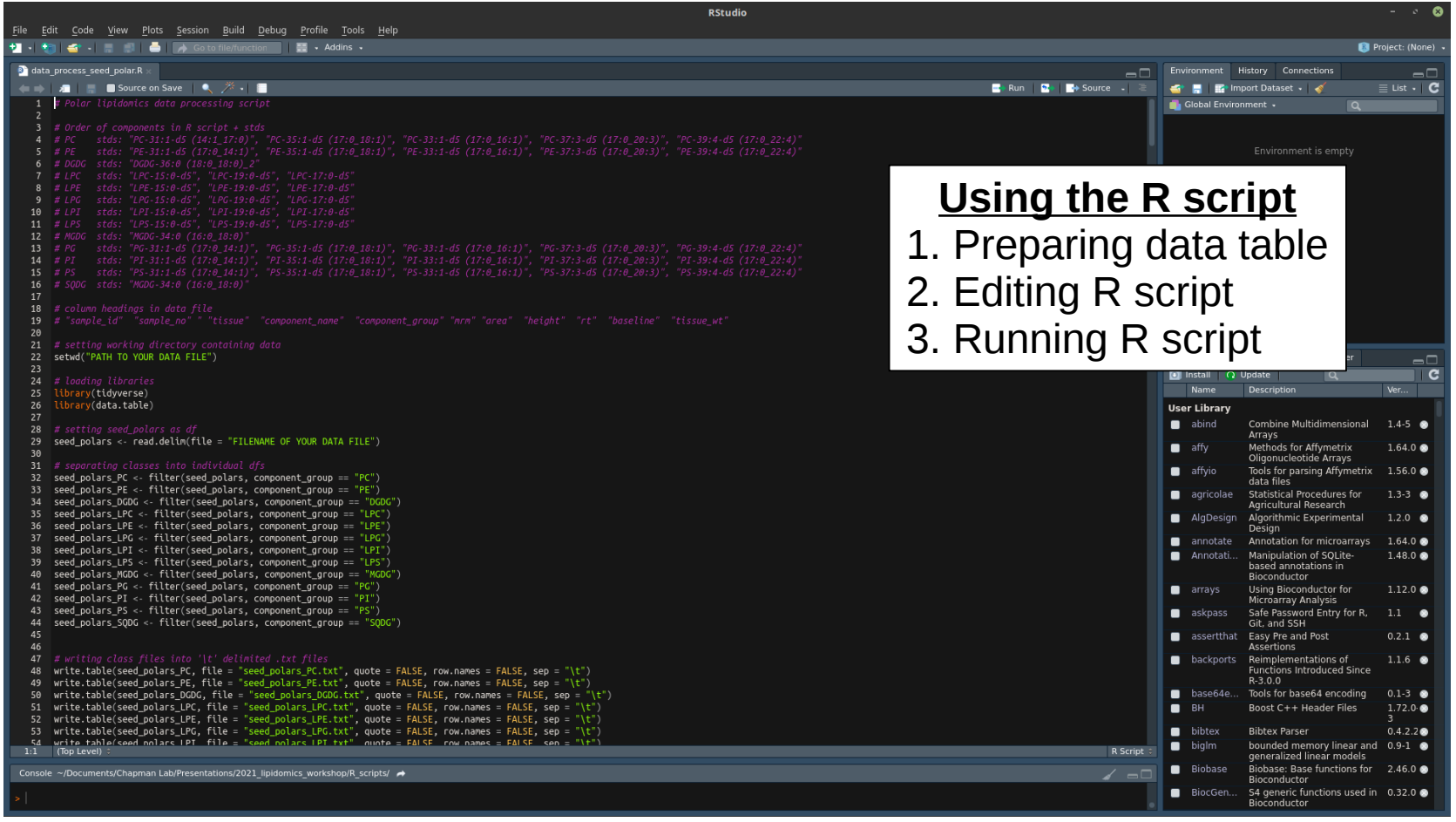
Go through each peak to be sure integrated correctly (This will take the most time)

2nd step in TAG analysis: Creating Multiquant method to integrate HPLC peaks

(Also 1st step for all the other lipids)



3rd step in data analysis: Using R/RStudio to quantify integated HPLC peaks



3rd step in data analysis: Using R/RStudio to quantify integated HPLC peaks

(in-person R/RStudio demonstration here)

Lipidomics Workflow — from extraction to analysis

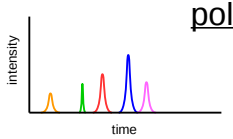
MTBE Extraction

OR

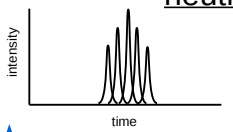
+ internal stds

HPLC-MS/MS

polar lipids



neutral lipids



Data Analysis

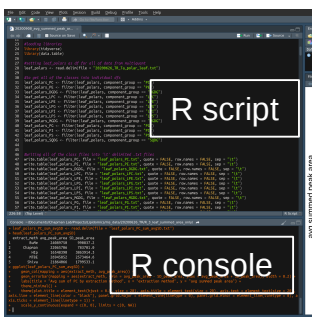
R + RStudio

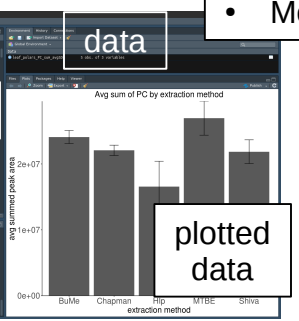
Data table → Class-specific data tables → detected molecular species; internal standards

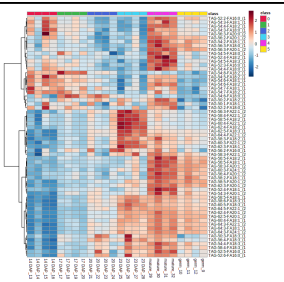
normalization by IS & tissue wt → average & SD of **molecular species**

sum of molecular species → average & SD of **summed species**

- processed data table
- plotted data
- Metaboanalyst-ready table







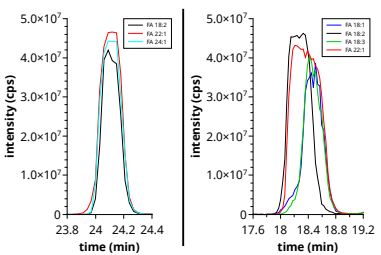
TAG MRMs

python

FA1 / FA1 / FA1
FA1 / FA1 / FA2
FA1 / FA2 / FA2
… … …

TAG specific steps

TAG-FA review for quantification



selected TAG MRMs

Other Resources

R resources

R: https://www.r-project.org/

RStudio: https://www.rstudio.com/products/rstudio/

R for Data Science (online ebook): https://r4ds.had.co.nz/index.html

Tidyverse library: https://www.tidyverse.org/

Data.Table library: https://cran.r-project.org/web/packages/data.table/vignettes/datatable-intro.html

W3: https://www.w3schools.com/r/default.asp

Python resources

Anaconda: https://www.anaconda.com/

Visual Studio Code: https://code.visualstudio.com/

Python Data Science Handbook (online ebook): https://jakevdp.github.io/PythonDataScienceHandbook/

Numpy (included with Anaconda): https://numpy.org/

Pandas (included with Anaconda): https://pandas.pydata.org/

Matplotlib (included with Anaconda): https://matplotlib.org/

W3: https://www.w3schools.com/python/default.asp

Lipidomics tools

mMass (viewing spectra/chromatogram from .txt file): http://www.mmass.org/

MarvinSketch (chemical drawing): https://chemaxon.com/products/marvin

Metaboanalyst (statistical analysis): https://www.metaboanalyst.ca/

Lipid metabolism

AOCS Lipid Library: https://lipidlibrary.aocs.org/

ARALIP: The Arabidopsis Acyl-Lipid Metabolism Website: http://aralip.plantbiology.msu.edu/about

Acyl-Lipid Metabolism: https://doi.org/10.1199/tab.0133

LIPID MAPS: https://www.lipidmaps.org/