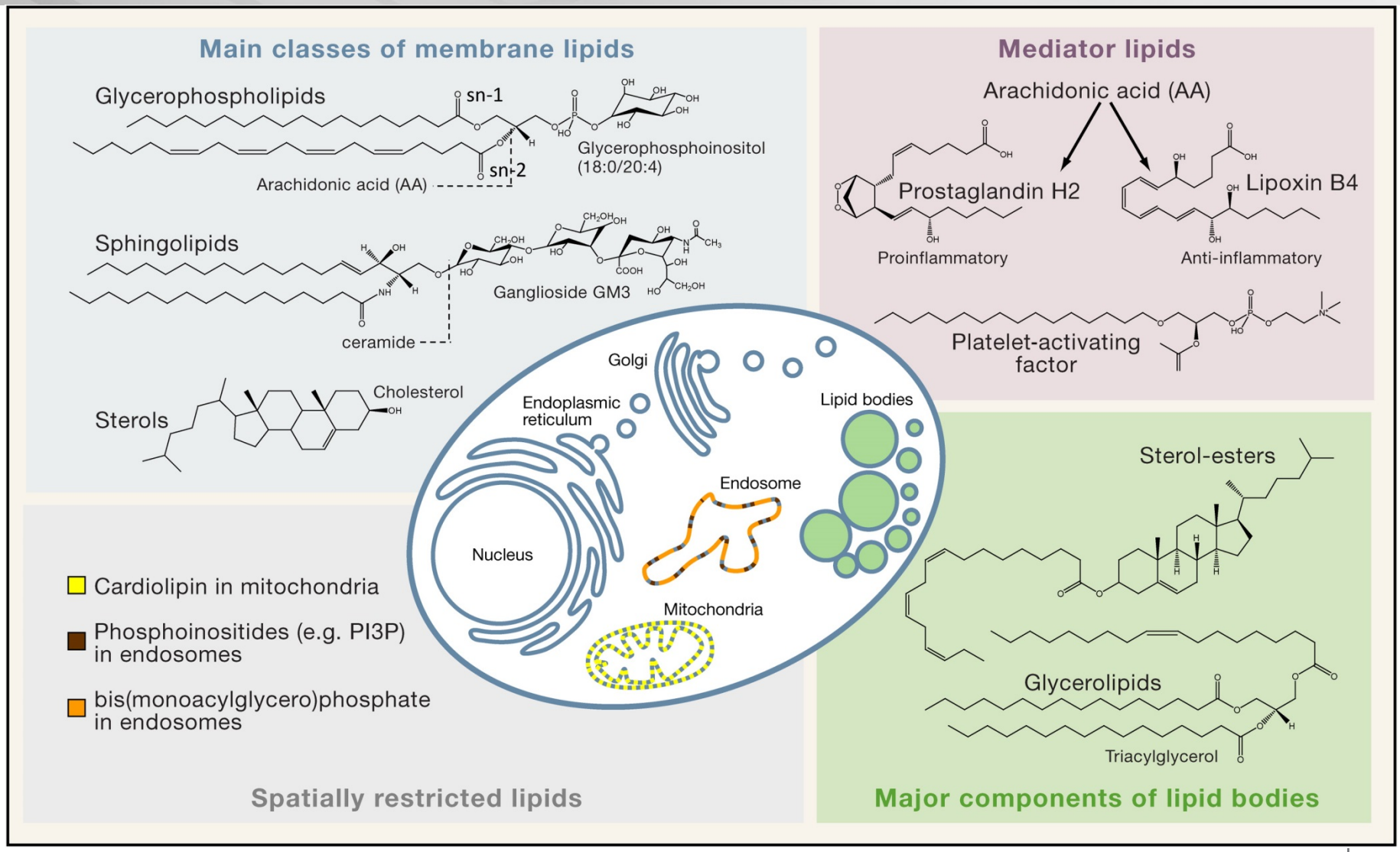
Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

Lipid Identification

SAMUEL H PAYNE
PNNL

U.S. DEPARTMENT OF
ENERGY

Lipids in Life

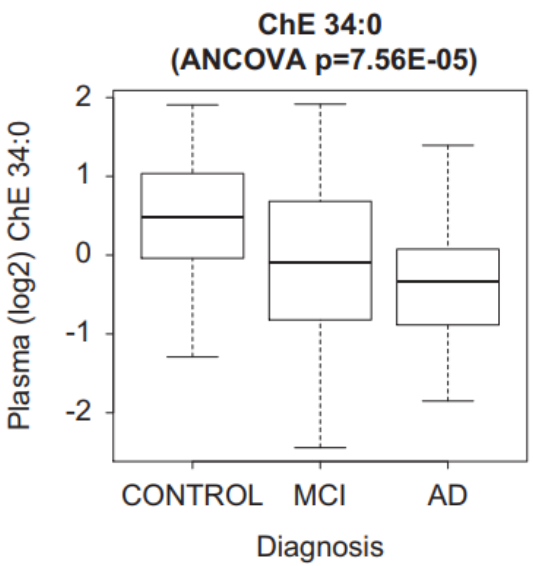


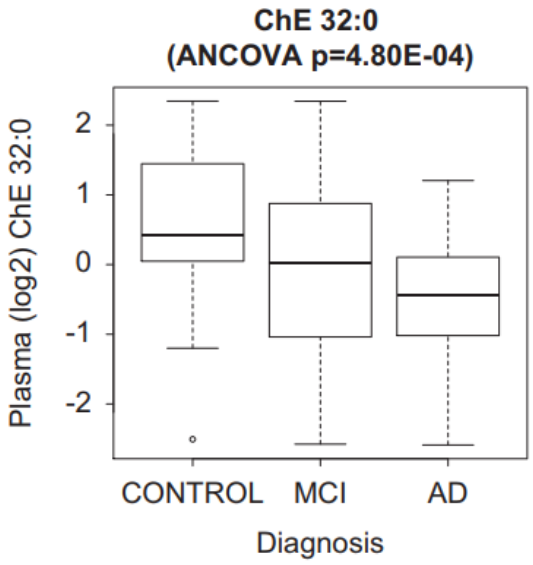
Wenk, Cell 2010

Lipids as Biomarkers

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

- Alzheimer's association
  - Cholesteryl esters





OPEN

Citation: Transl Psychiatry (2015) 5, e494; doi:10.1038/tp.2014.127
www.nature.com/tp

ORIGINAL ARTICLE

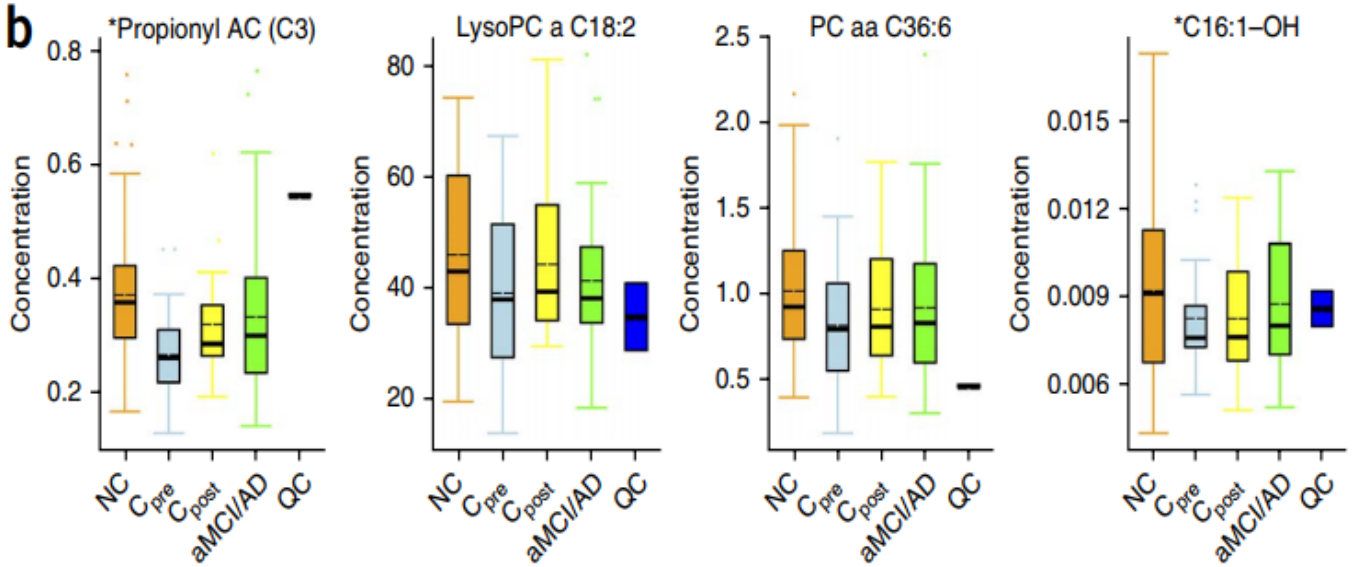
Plasma lipidomics analysis finds long chain cholesteryl esters to be associated with Alzheimer's disease

P Proitsi[1], M Kim[2], L Whiley[2], M Pritchard[1], R Leung[1], H Soininen[3], I Kloszewska[4], P Mecocci[5], M Tsolaki[6], B Vellas[7], P Sham[8], S Lovestone[9], JF Powell[1,11], RJB Dobson[1,10,11] and C Legido-Quigley[2,11]

Lipids as Biomarkers

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

- Alzheimer's association
  - Phospholipid panel



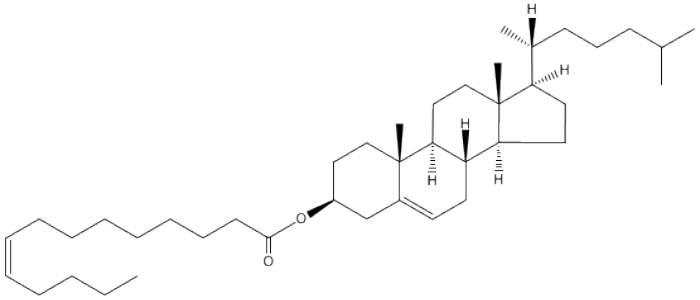
Plasma phospholipids identify antecedent memory impairment in older adults

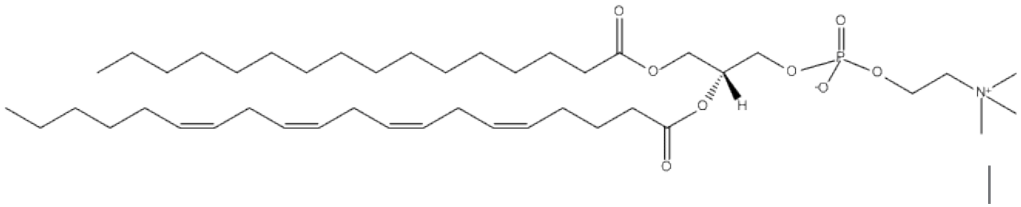
Mark Mapstone[1], Amrita K Cheema[2,3], Massimo S Fiandaca[4,5], Xiaogang Zhong[6], Timothy R Mhyre[5], Linda H MacArthur[5], William J Hall[7], Susan G Fisher[8,14], Derick R Peterson[9], James M Haley[10], Michael D Nazar[11], Steven A Rich[12], Dan J Berlau[13,14], Carrie B Peltz[13], Ming T Tan[6], Claudia H Kawas[13] & Howard J Federoff[4,5]

NATURE MEDICINE VOLUME 20 | NUMBER 4 | APRIL 2014

Challenges of Lipids

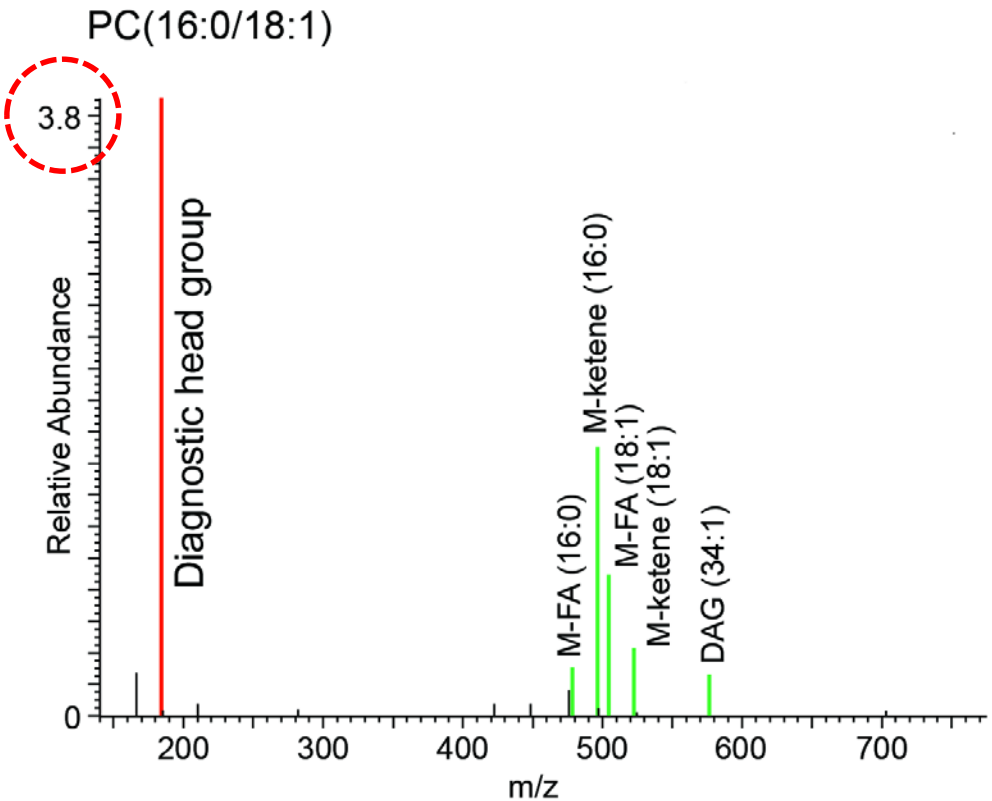
- Biochemically diverse
  - Poor separation in chromatography
  - Require +/- ionization for full fragmentations
- No genome template
  - No ‘database search’ equivalent
  - Genome models poorly predict lipid species





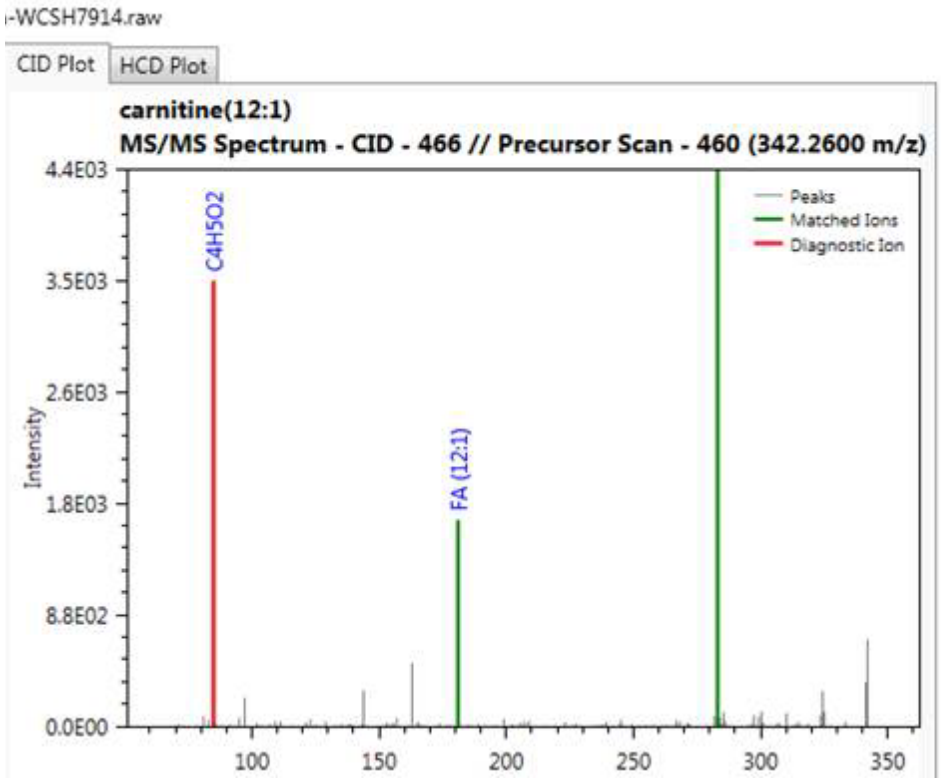
Challenges of Lipid Spectra

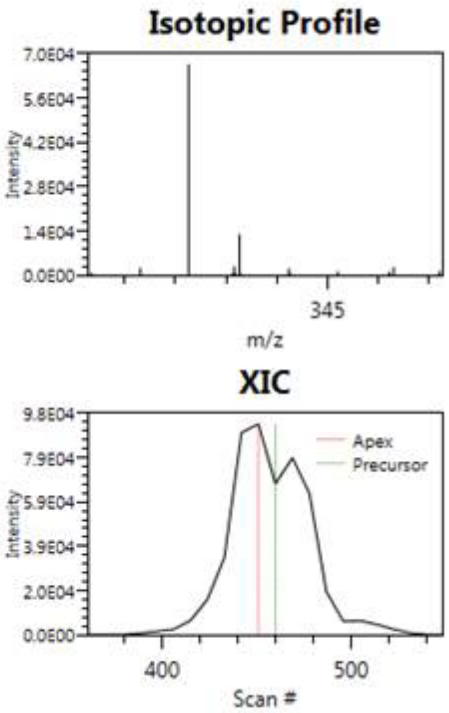
- Few peaks
- Low intensity peaks



Challenges of Lipid Spectra

- Co-elution of isomers

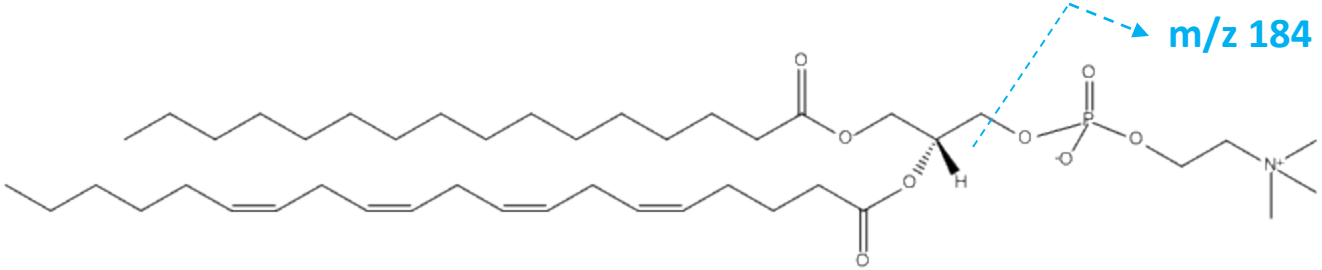




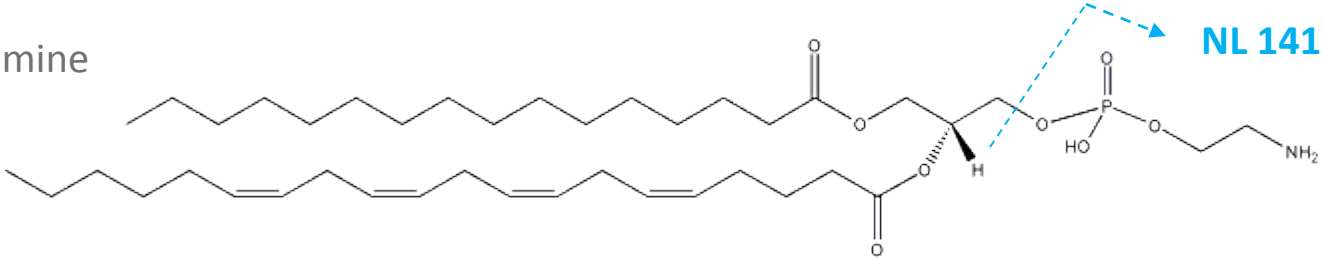
Challenges of Lipid Spectra

- Diagnostic peaks differ by class

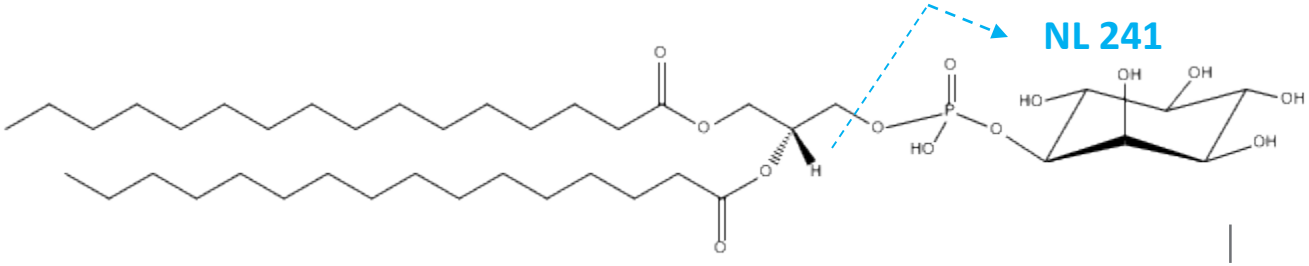
phosphatidylcholine
(16:0/20:4)



phosphatidylethanolamine
(16:0/20:4)



phosphatidylinositol
(16:0/16:0)



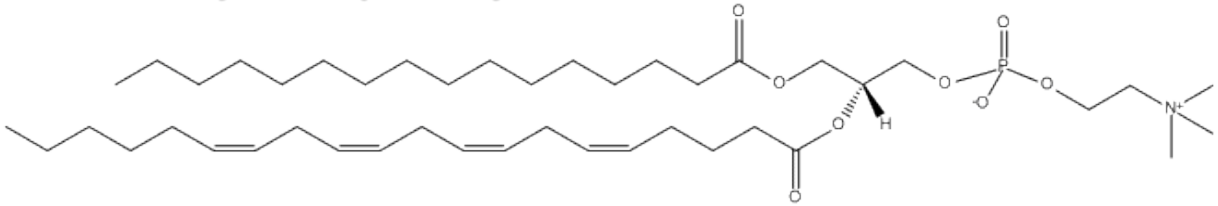
Informatics for Lipidomics

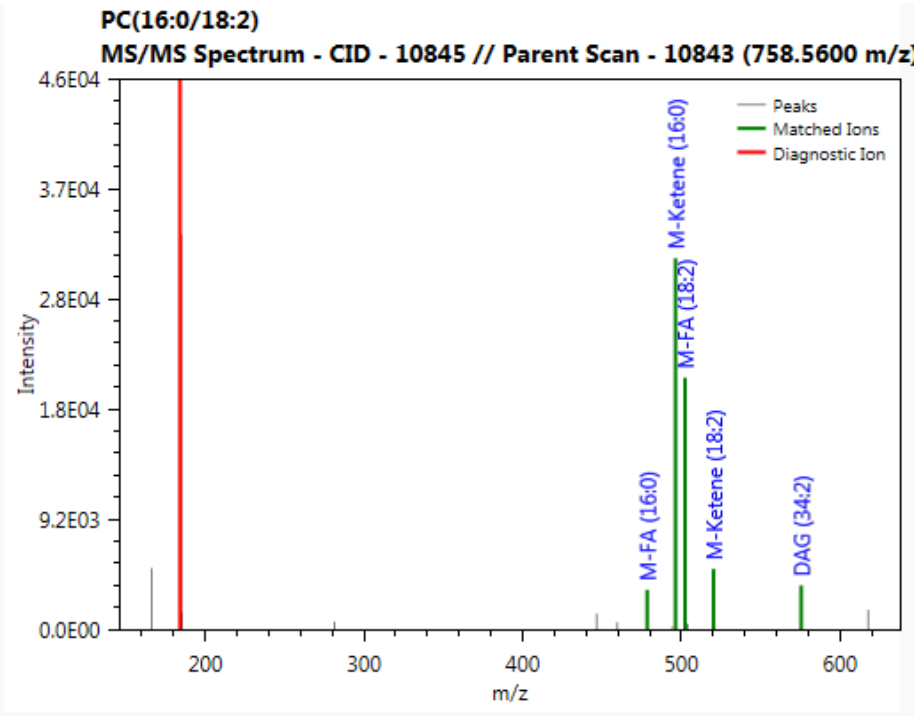
- How to assist in manual annotation
  - Graphical display, support software
- How to replace manual annotation
  - False discovery estimation

Capturing the Typical

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

diacyl-Phosphatidylcholine





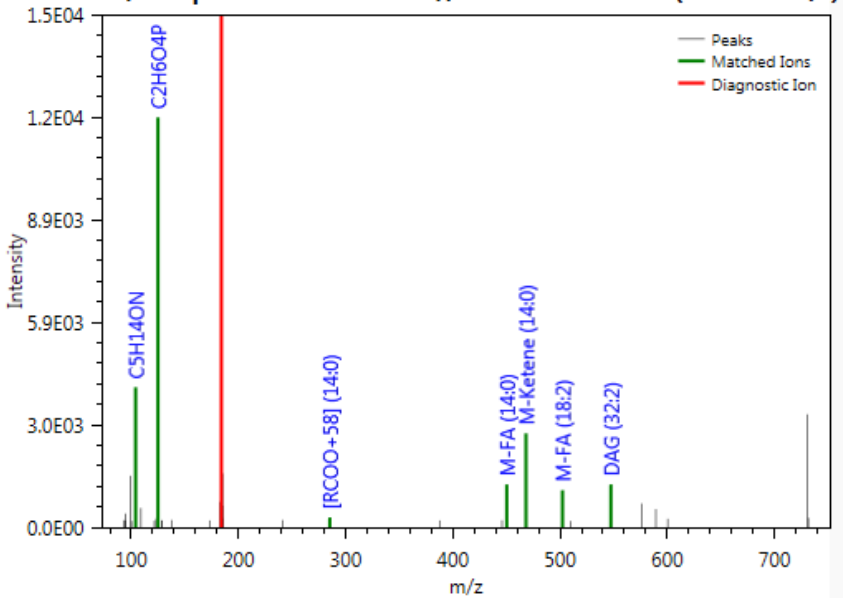
Fragment ion	C	H	N	O	P
C5H15O4NP	5	15	1	4	1
[DAG]	X+3	2(X+3)-3-2Y	0	4	0
loss of ketene (1)	X+8	2(X+8)+3-2Y	1	7	1
loss of ketene (2)	X+8	2(X+8)+3-2Y	1	7	1
loss of fatty acid (1)	X+8	2(X+8)+1-2Y	1	6	1
loss of fatty acid (2)	X+8	2(X+8)+1-2Y	1	6	1

Multiple Metrics

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle *Since 1965*

HCD Plot | CID Plot

PC(14:0/18:2)
MS/MS Spectrum - HCD - 5737 // Parent Scan - 5734 (730.5300 m/z)

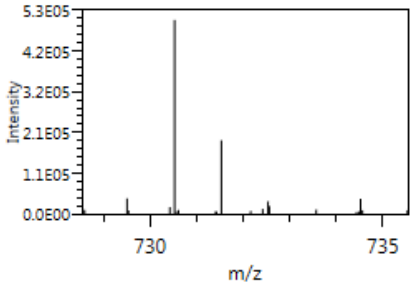


Observed MS/MS | Theoretical MS/MS

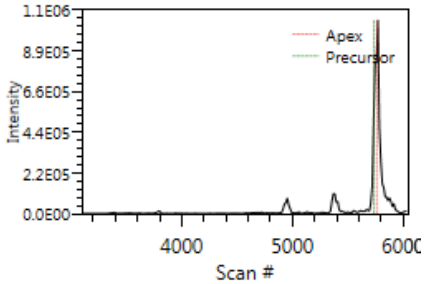
Theoretical MS/MS Peaks

m/z	Description
104.1075	C5H14ON
125.0004	C2H6O4P
184.0739	C5H15O4NP
211.2062	FA (14:0)
240.1001	C8H19O5NP
258.1106	C8H21O6NP
263.2375	FA (18:2)
267.2324	[RCOO+58]-H2O (14:0)
285.243	[RCOO+58] (14:0)

Isotopic Profile



XIC



Start Scan: 5760 Stop Scan: 5760

Target Lipid Info

Adduct	Hydrogen
Target m/z	730.5387
Empirical Formula	C40H77N1O8P1

Observed Info

Observed m/z	730.5297
PPM Error	-12.37
Scan LC Apex	5760
Normalized Elution Time	0.633
Intensity	8.31E05
Area Under Curve	831,285

Liquid

- Scores MS/MS spectra
  - Expected fragmentation pattern and intensity
- Precursor XIC
- Precursor isotopic envelope
- Customizable database

Bioinformatics, 33(11), 2017, 1744–1746
doi: 10.1093/bioinformatics/btx046
Advance Access Publication Date: 31 January 2017
Applications Note

Data and text mining

LIQUID: an-open source software for identifying lipids in LC-MS/MS-based lipidomics data

Jennifer E. Kyle*,†, Kevin L. Crowell†, Cameron P. Casey, Grant M. Fujimoto, Sangtae Kim, Sydney E. Dautel, Richard D. Smith, Samuel H. Payne and Thomas O. Metz*

File | Single Target | Global Analysis

File loaded: Dey2_lipids_Bottom_1_1__POS_16Jun14_Polaroid_14-02-09.raw

HCD Mass Error: 30

CID Mass Error: 500

Ionization Mode: Positive

Load Lipid Targets

# of Targets: 15505

Results Per Scan: 1

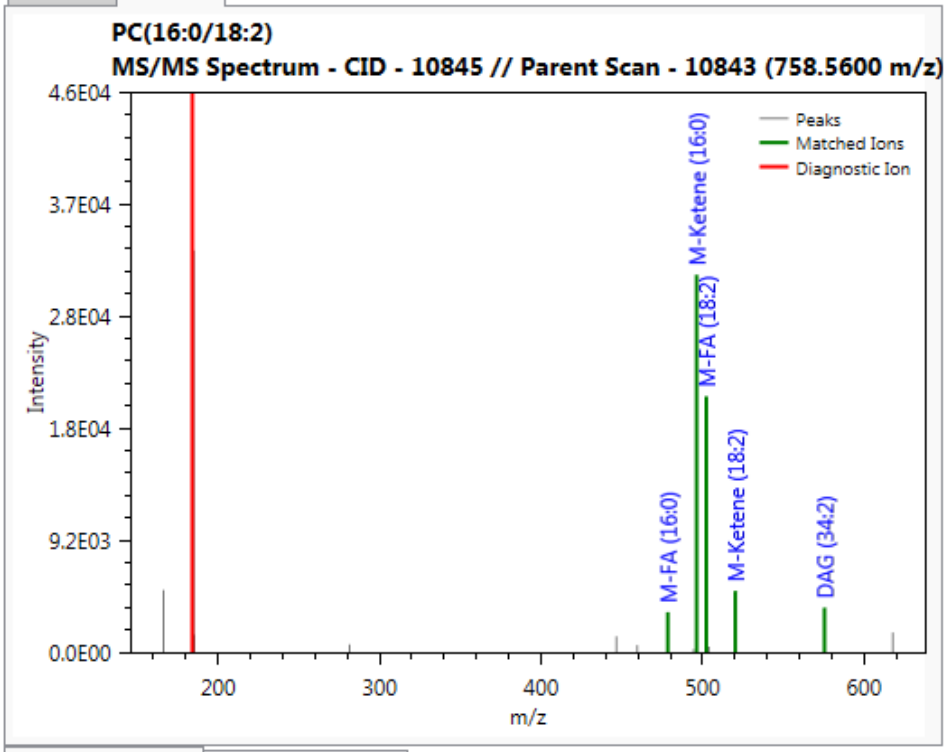
Process Targets!

	Scan	Lipid	Score
☐	10787	PE(16:0/18:2)	30.04
☑	10805	PC(20:2/22:6)	46.77
☐	10807	PC(O-16:0/22:6)	35.01
☐	10813	GalCer(d18:0/16:0)	17.13
☐	10815	PC(15:1/18:2)	29.11
☐	10822	PC(O-16:0/14:0)	20.39
☐	10824	PC(17:0/22:6)	45.84
☐	10835	PE(P-16:0/20:4)	30.16
☐	10841	PC(P-18:0/22:6)	31.40
☑	10844	PC(16:0/18:2)	35.13
☐	10850	PC(16:0/22:6)	49.81
☐	10861	PC(17:0/20:4)	52.43
☑	10870	PC(18:1/22:6)	51.00
☐	10876	PC(16:0/P-16:0)	26.03
☐	10883	PC(16:0/18:0)	32.70
☐	10887	PC(17:0/18:2)	47.99
☐	10889	PC(17:0/22:5)	38.73
☑	10896	PC(16:0/20:3)	31.98
☐	10898	PC(18:0/20:5)	33.03
☐	10900	PC(18:2/20:2)	45.36
☑	10902	PC(18:0/22:6)	40.69
☐	10904	PC(18:1/18:1)	20.60

Export Selected Results

Export All Results

HCD Plot | CID Plot

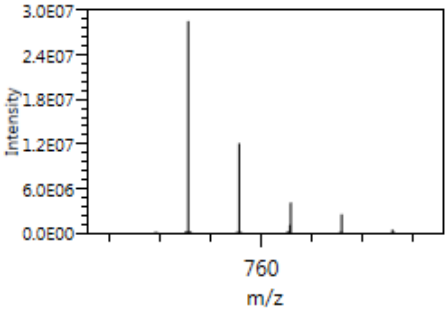


Observed MS/MS | Theoretical MS/MS

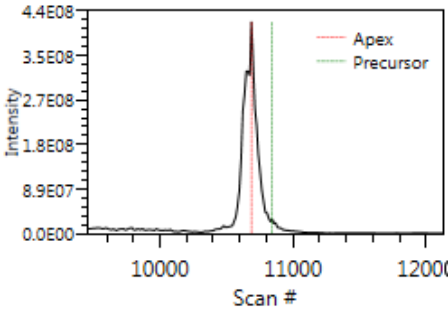
Annotated MS/MS Peaks

Type	m/z	Intensity	Annotation
HCD	104.1061	123251	C5H14ON
HCD	124.9988	363835	C2H6O4P
HCD	184.0717	18239154	C5H15O4NP
HCD	478.3228	21064	M-FA (16:0)
HCD	496.3375	48478	M-Ketene (16:0)
HCD	575.4953	35802	DAG (34:2)
CID	184.0462	1511341	C5H15O4NP
CID	478.4988	3298	M-FA (16:0)
CID	496.3448	31038	M-Ketene (16:0)

Isotopic Profile



XIC



Start Scan: 10531 Stop Scan: 11051

Target Lipid Info

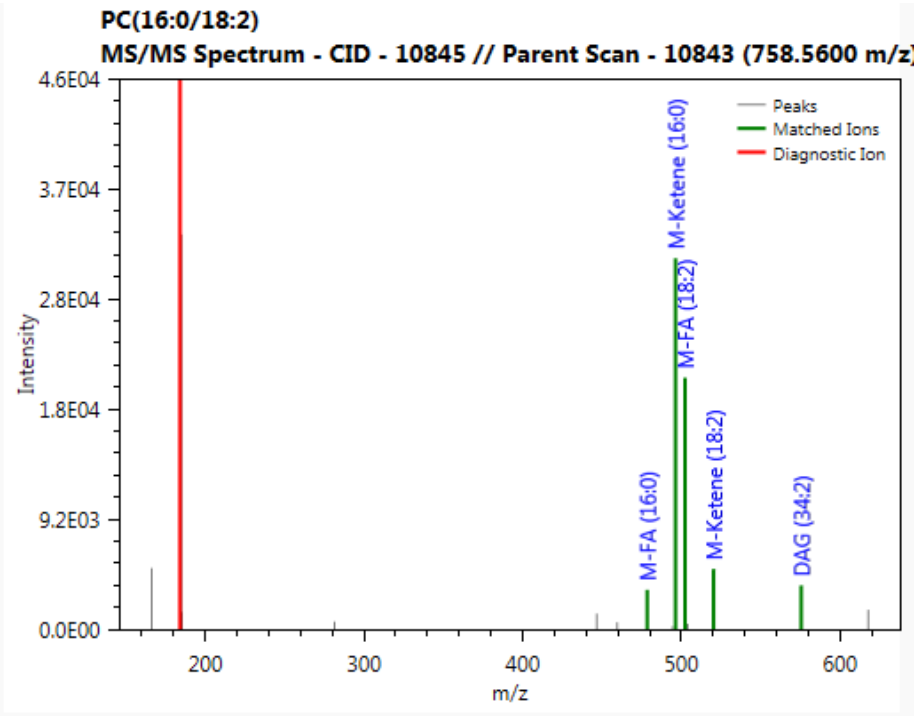
Adduct	Hydrogen
Target m/z	758.5700
Empirical Formula	C42H81N1O8P1

Observed Info

Observed m/z	758.5632
PPM Error	-8.91
Scan LC Apex	10687
Normalized Elution Time	0.751
Intensity	4.22E08
Area Under Curve	3,816,604,019

Scoring MS/MS spectra

- Learning and using expected fragments
  - Probability of **peak** at **intensity**

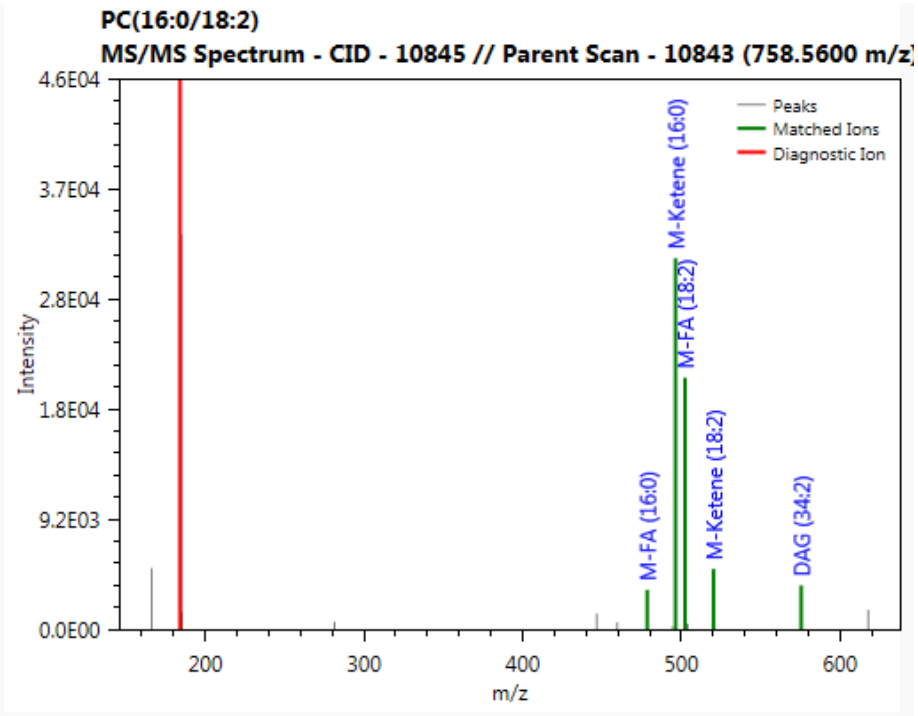


Probability table

Fragment ion	High	Mid	Low	NOT
Diagnostic	1.0	0	0	0
[DAG]	0.1	0.1	0.7	0.1
loss of ketene	0.1	0.3	0.5	0.1
loss of fatty acid	0.0	0.1	0.8	0.1

Future Scoring

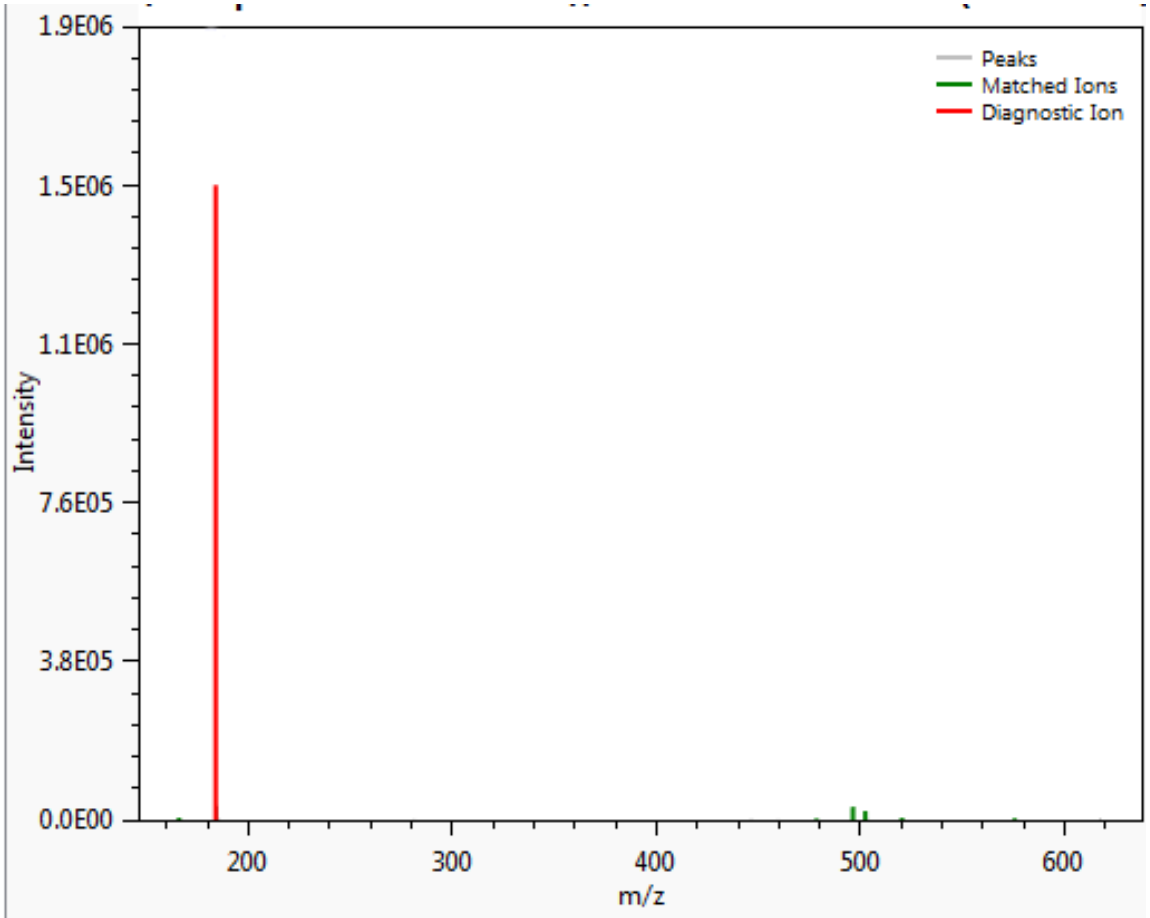
- Bayesian probability
  - Probability of peak 1 given observation of peak2



Peak1	Peak2	P(p1|p2)
Hi	Hi	0.1
Mid	Hi	0.2
Low	Hi	0.6
NOT	Hi	0.1

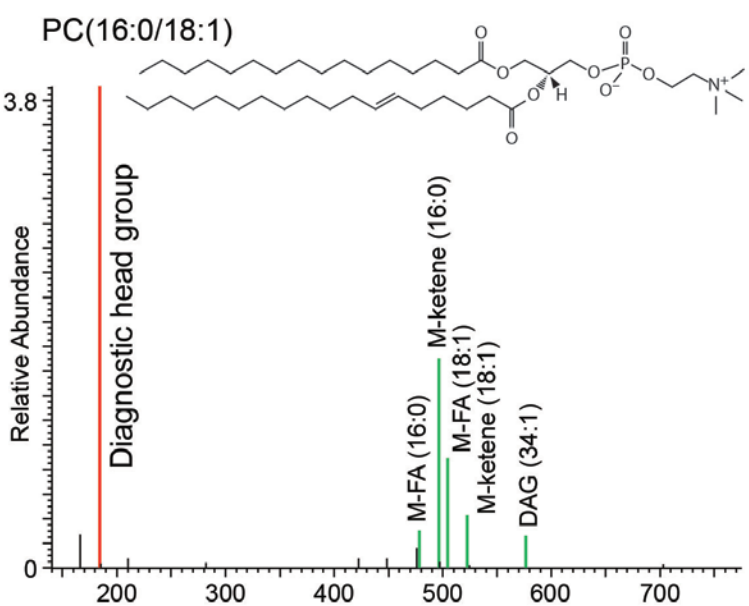
Future Scoring

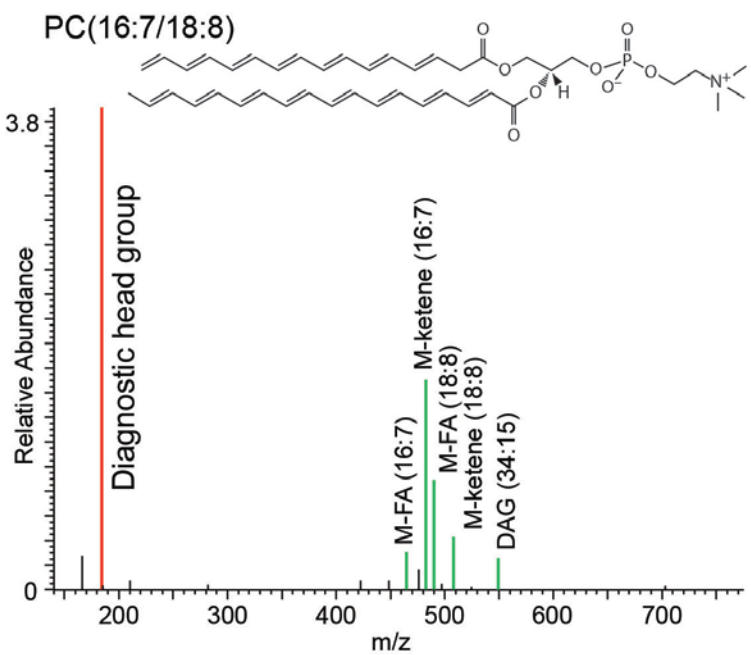
- Scaling Intensity (single dominant peak)



False Discovery Estimation

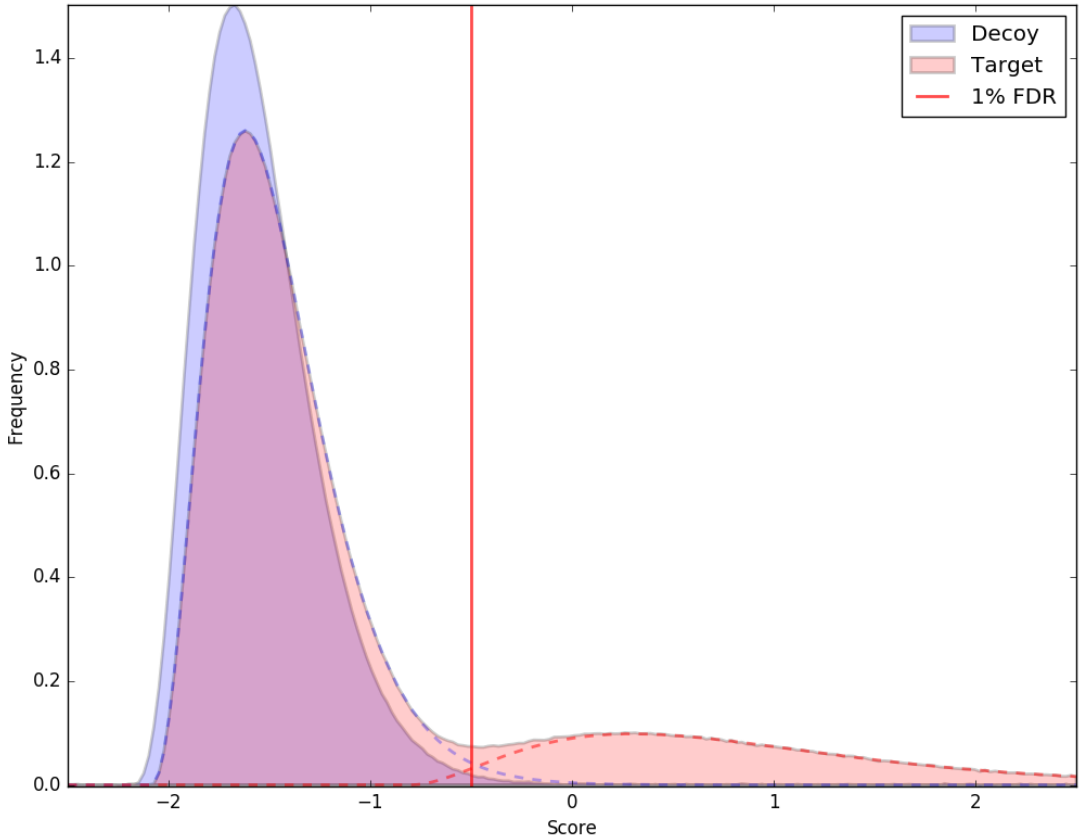
- Decoy based
  - Create unrealistic molecules in database





Good Decoy

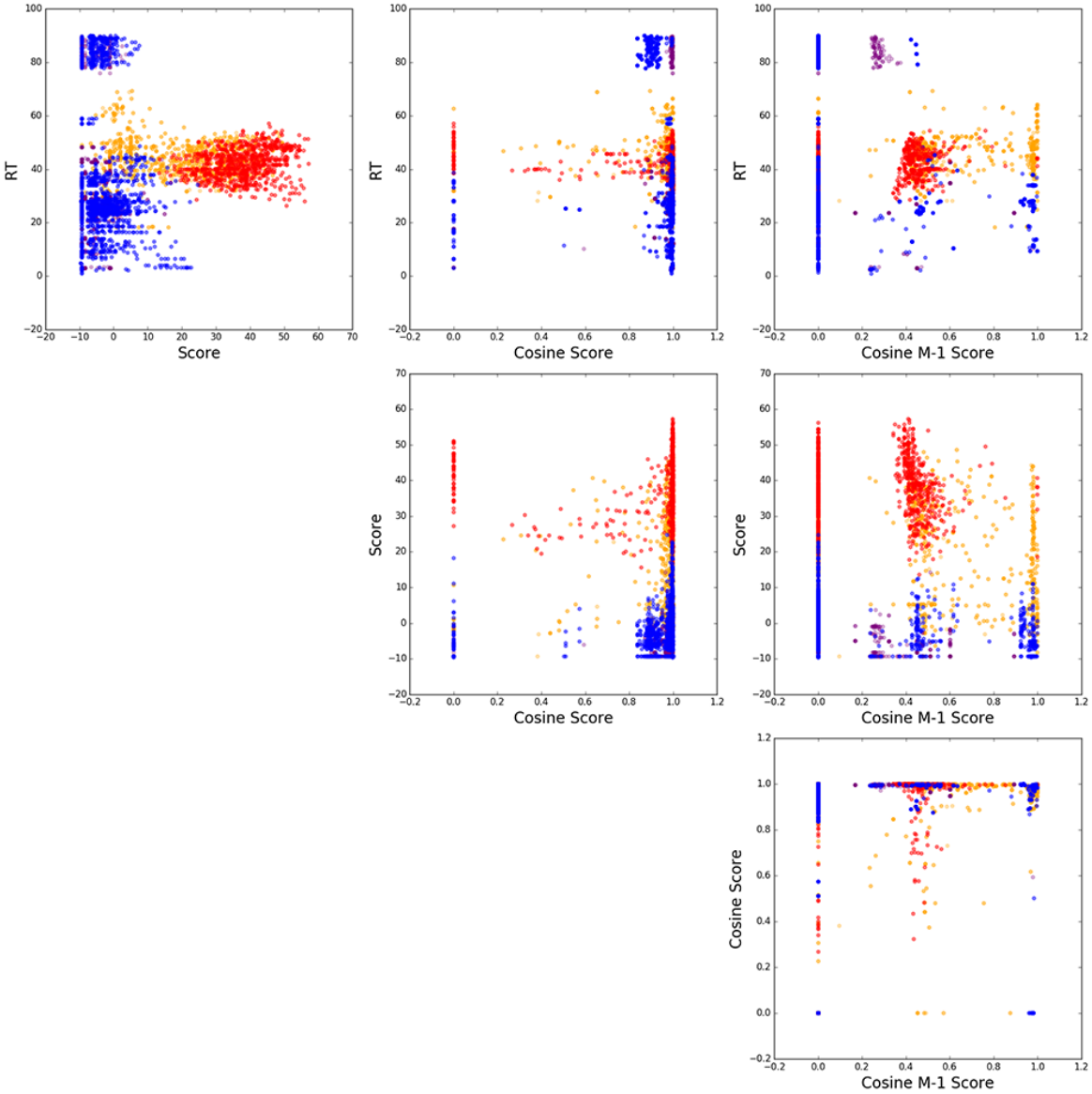
- True Negative == False Positive
  - (adapted from Peptide Prophet)



Separating Target and Decoy

- Red (TP)
- Blue (TN)

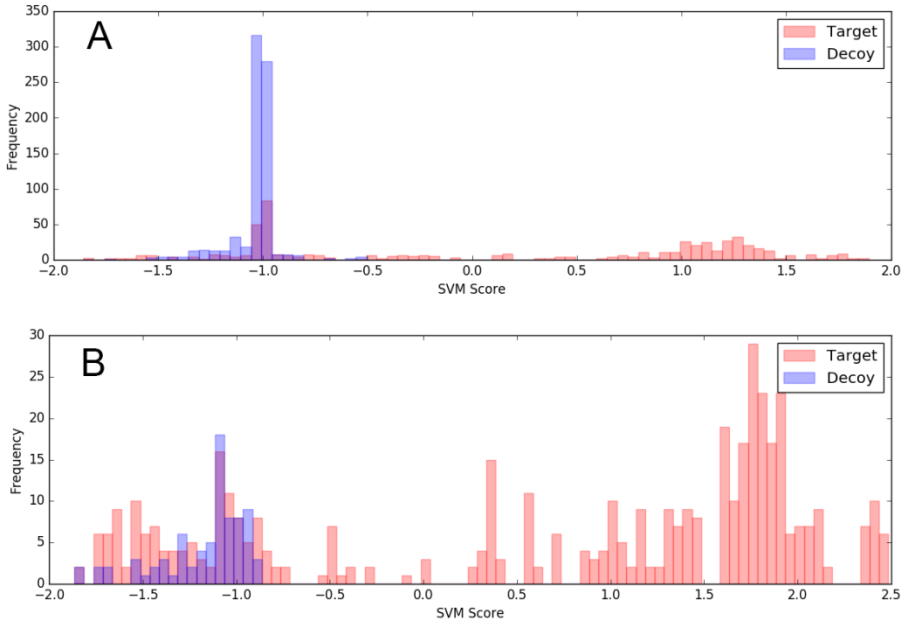
- Multiple dimension = SVM



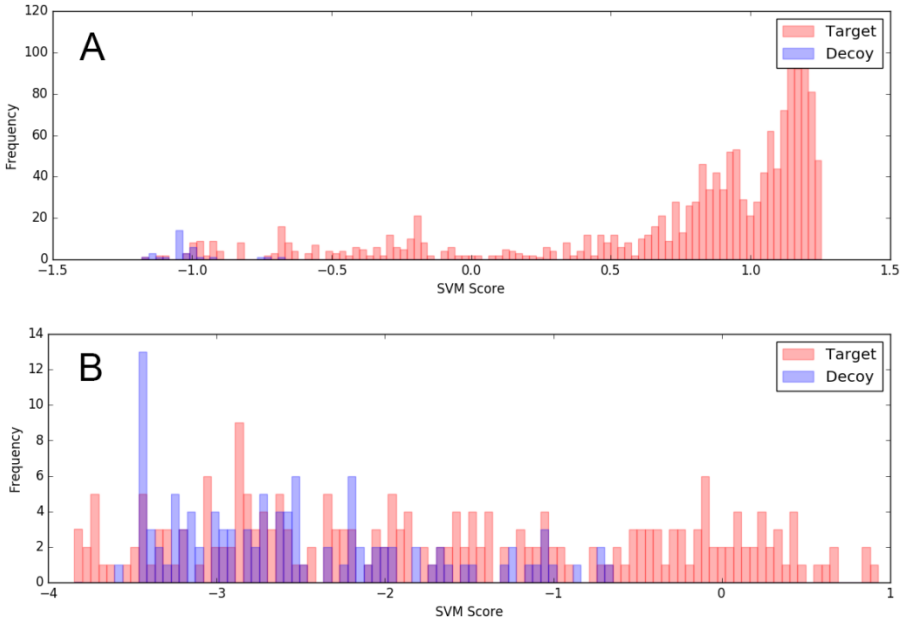
Lipid Decoys

- Non uniform score for lipid classes
- Bin scoring by classes

Good Decoy



Bad Decoy



Future FDR

- Creating a unified score across classes
- Making decoys have equal quality

- IM-MS as a metric
- MS1 only approaches?
- MS2 DIA

Acknowledgements

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

- PNNL Integrative Omics group

- Jennifer Kyle
- Grant Fujimoto
- Joon-Yong Lee