# REPOSITORY-SCALE PROPAGATED SPECTRAL LIBRARY OF SUSPECTS

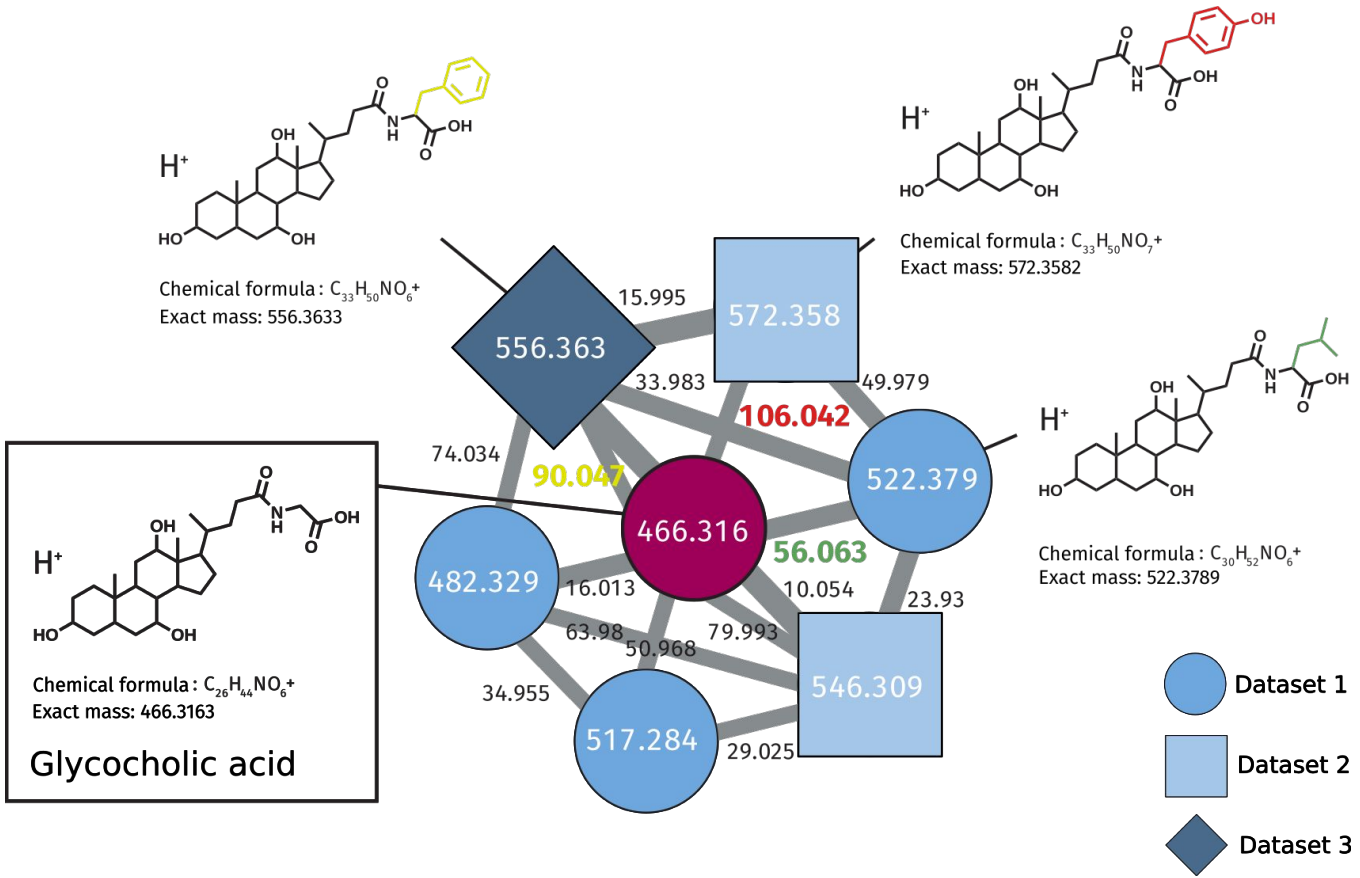
Wout Bittremieux — SoCal Science @Home Lollapalooza

 @bittremieux

 bittremieux.be

 wbittremieux@health.ucsd.edu

# MOLECULAR NETWORKING




Aron, A. *et al*. Reproducible molecular networking of untargeted mass spectrometry data using GNPS. *Nature Protocols* **15**, 1954–1991 (2020).

# REPOSITORY-SCALE ANALYSIS

**Living data** = periodic reanalysis of GNPS datasets

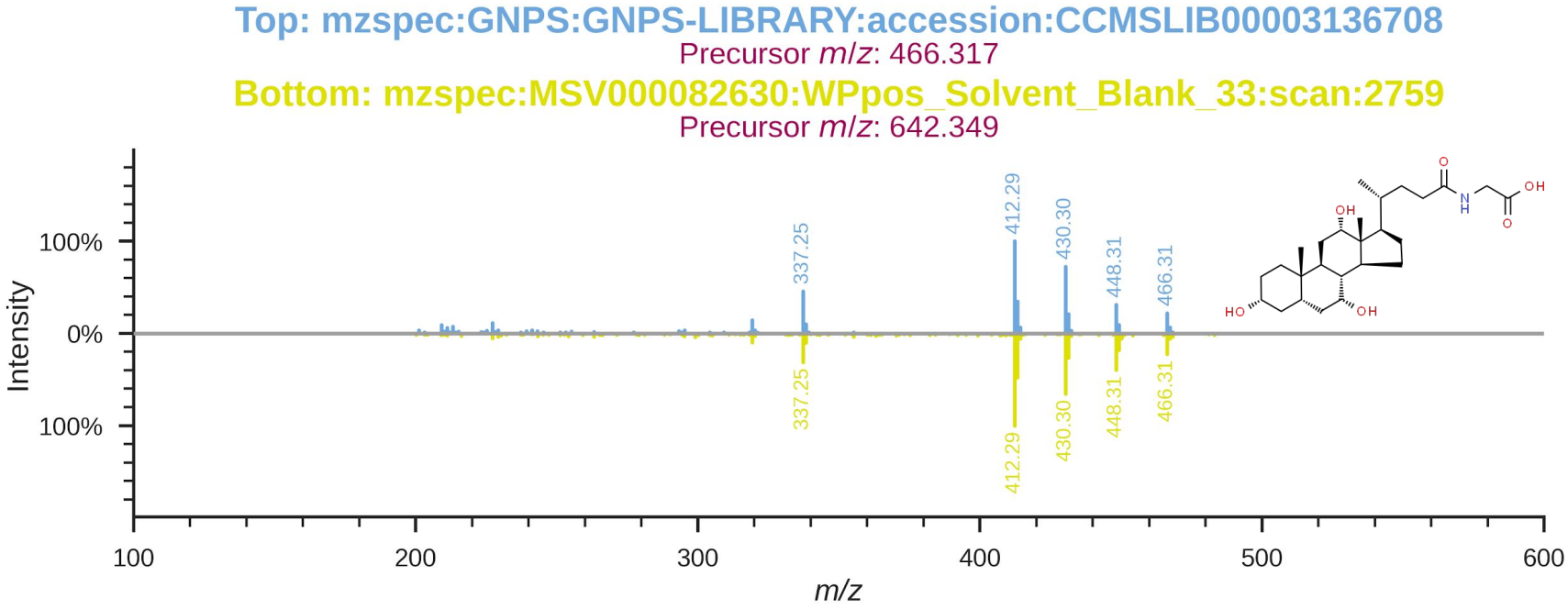
- 1,335 public datasets (Nov. 2020)
- 963,616 spectrum pairs from molecular networking
- 90,937 unique “suspects” from 34,263 MS runs

**Suspects**: provide clues about molecular identities,
but careful validation (synthesis) required for confirmation

Wang, M. *et al*. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature Biotechnology* **34**, 828–837 (2016).

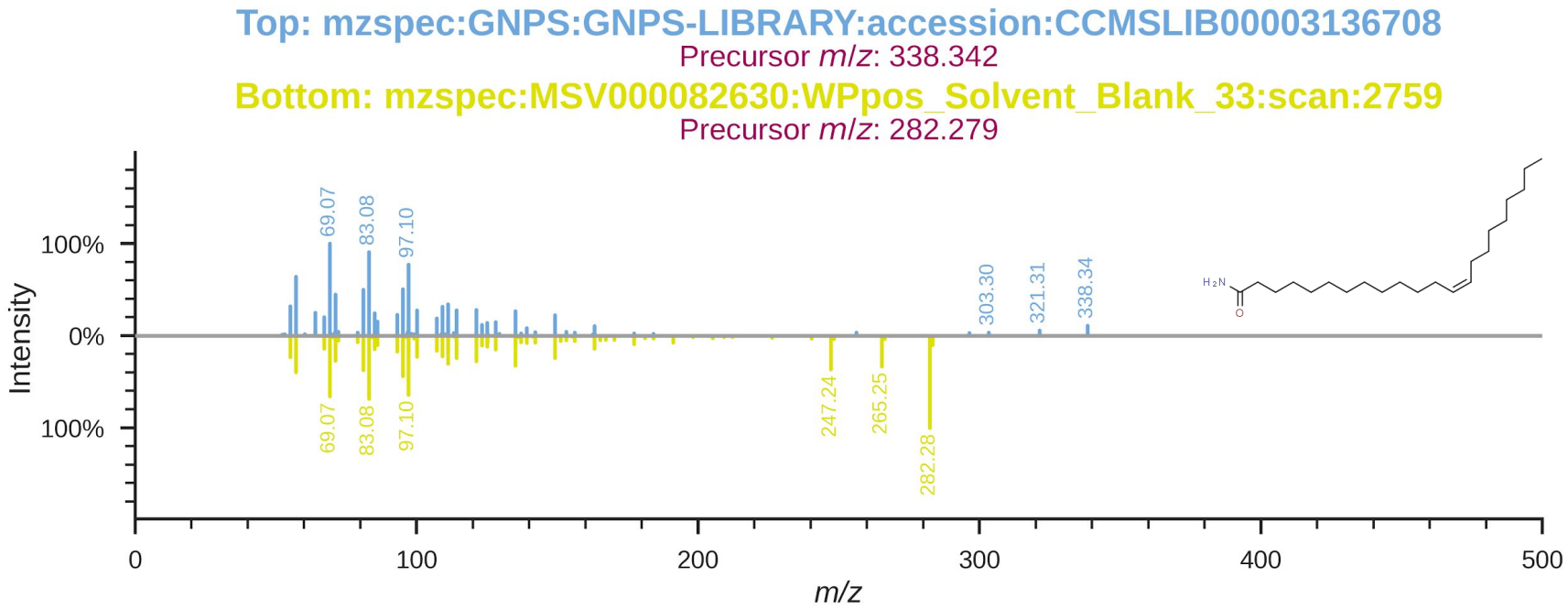
# SUSPECTS

Suspect related to glycocholic acid with delta *m/z* 176.032
(potential explanation: glucuronidation)



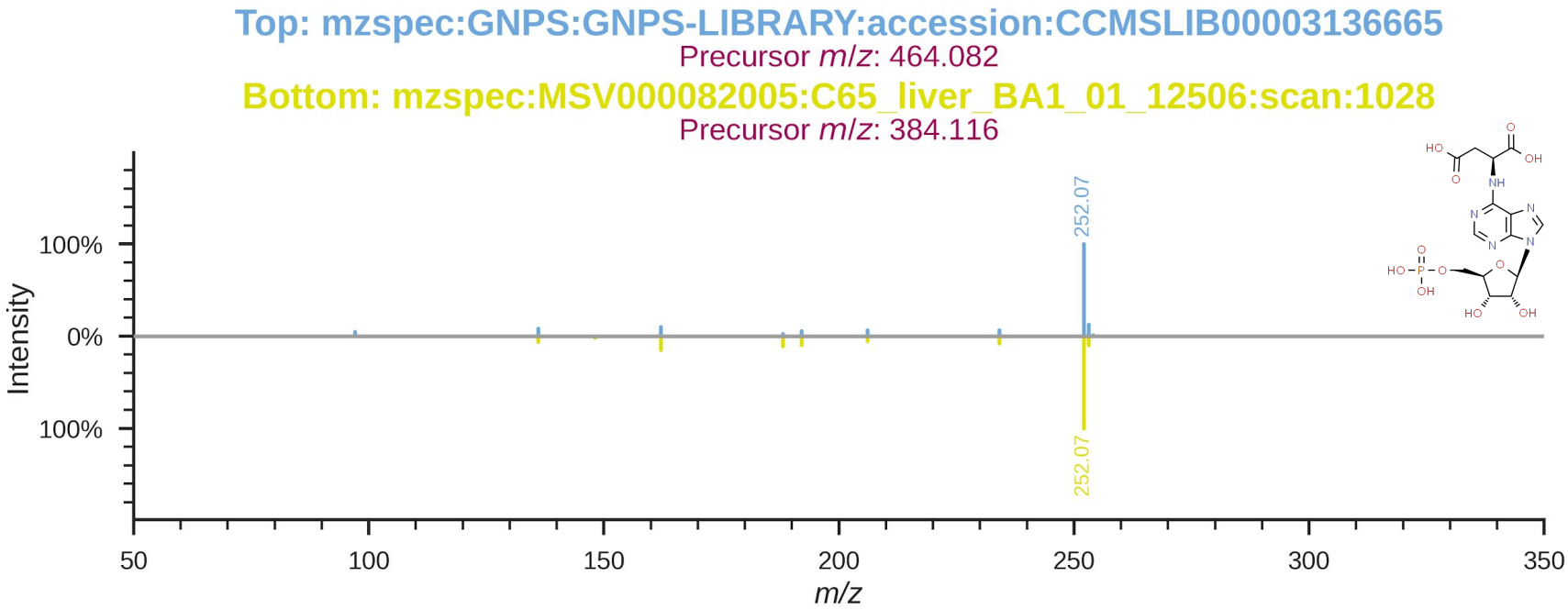
# SUSPECTS

Suspect related to 13-Docosenamide, (Z)- with delta *m/z* -56.063 (potential explanation: $C_4H_8$)

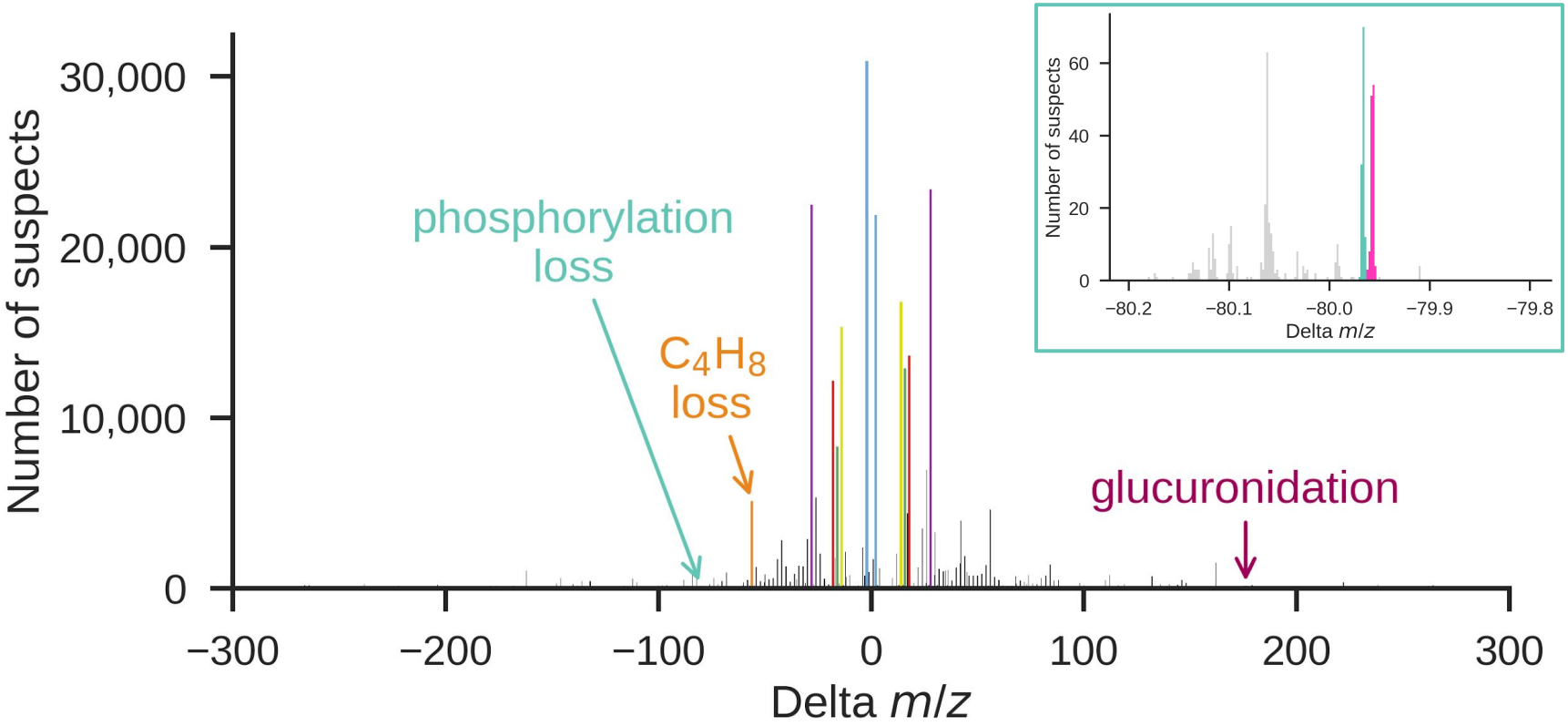


# SUSPECTS

Suspect related to adenylosuccinic acid with delta *m/z* -79.966
(potential explanation: phosphorylation)



# DELTA MASS


Number of suspects
30,000
20,000
10,000
0
−300
−200
−100
0
100
200
300
Delta m/z
phosphorylation loss
C4H8 loss
glucuronidation
Number of suspects
60
40
20
0
−80.2
−80.1
−80.0
−79.9
−79.8
Delta m/z

# SUSPECT LIBRARY

- Suspect library
  - 90,937 suspects (current number of unique matches to reference library spectra: 14,777)
  - 1,347 unique delta masses
- Insights into organic chemistry via data-driven library creation
- Provided on GNPS for spectral library searching
- Looking for collaborators to test the suspect library!

# ACKNOWLEDGMENTS


## Dorrestein lab


UC San Diego

fwo Research Foundation Flanders
Opening new horizons