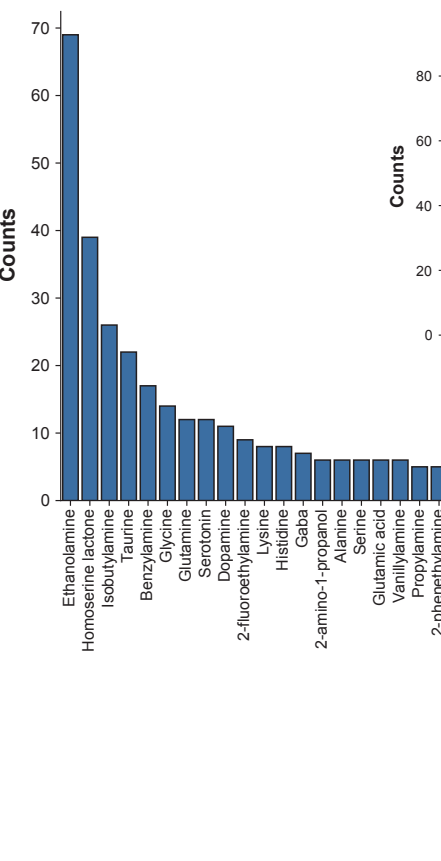
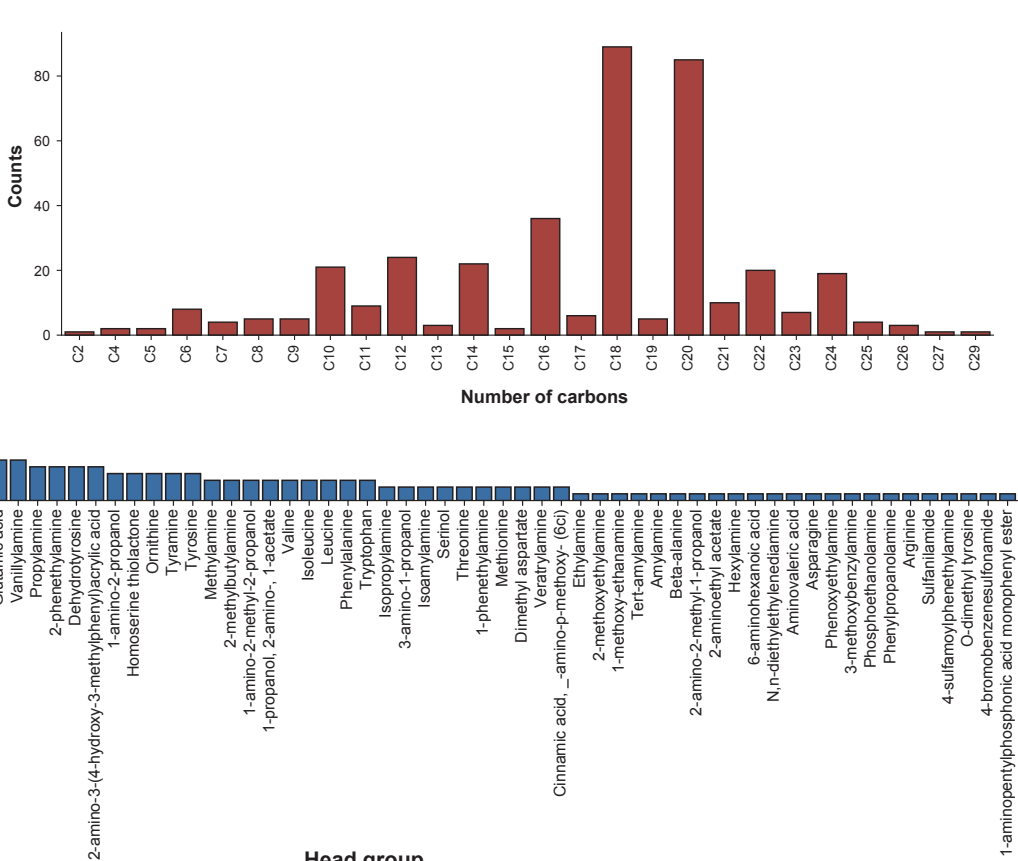


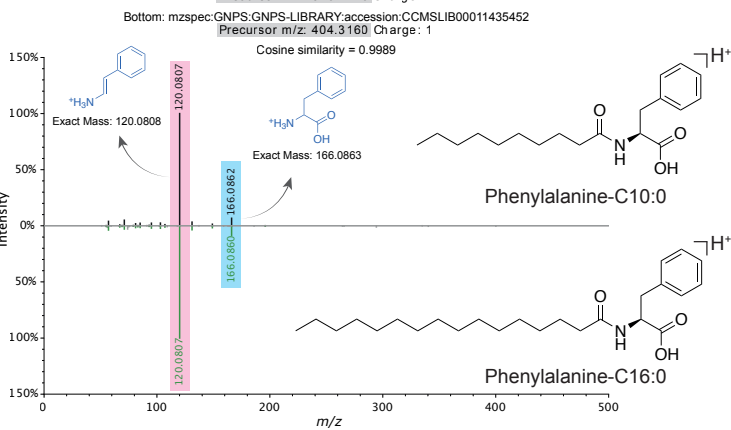
A



B



C



D

QUERY scaninfo(MS2DATA) WHERE
MS2PREC=(208.0968 OR 222.1124 OR 236.128 OR 250.1436 OR 264.1592 OR 278.1748 OR 292.1904 OR 306.206 OR 320.2216 OR 334.2372 OR 348.2528 OR 362.2684 OR 376.284 OR 390.2996 OR 404.3152 OR 418.3308 OR 432.3464 OR 446.362 OR 460.3776 OR 474.3932 OR 488.4088 OR 502.4244 OR 516.44 OR 530.4556 OR 544.4712 OR 558.4868 OR 572.5024 OR 586.518 OR 600.5336 OR 620.0978 OR 634.1134 OR 648.129 OR 662.1446 OR 676.1602 OR 690.1758 OR 704.1914 OR 718.207 OR 732.2226 OR 746.2382 OR 760.2538 OR 774.2694 OR 788.285 OR 802.3006 OR 816.3162 OR 830.3318 OR 844.3474 OR 858.363 OR 872.3786 OR 886.3942 OR 900.4098 OR 914.4254 OR 928.441 OR 942.4566 OR 956.4722 OR 970.4878 OR 984.5034 OR 998.519 OR 1012.5348 OR 1026.5504 OR 1040.566 OR 1054.5816 OR 1068.5972 OR 1082.6128 OR 1096.6284 OR 1110.644 OR 1124.6596 OR 1138.6752 OR 1152.6908 OR 1166.7064 OR 1180.722 OR 1194.7376 OR 1208.7532 OR 1222.7688 OR 1236.7844 OR 1250.7996 OR 1264.8152 OR 1278.8308 OR 1292.8464 OR 1306.862 OR 1320.8776 OR 1334.8932 OR 1348.9088 OR 1362.9244 OR 1376.94 OR 1390.9556 OR 1404.9712 OR 1418.9868 OR 1432.1024 OR 1446.118 OR 1460.1336 OR 1474.1492 OR 1488.1648 OR 1502.1804 OR 1516.196 OR 1530.2116 OR 1544.2272 OR 1558.2428 OR 1572.2584 OR 1586.274 OR 1600.2896 OR 1614.3052 OR 1628.3208 OR 1642.3364 OR 1656.352 OR 1670.3676 OR 1684.3832 OR 1698.3988 OR 1712.4144 OR 1726.4296 OR 1740.4452 OR 1754.4608 OR 1768.4764 OR 1782.492 OR 1796.5076 OR 1810.5232 OR 1824.5388 OR 1838.5544 OR 1852.5696 OR 1866.5852 OR 1880.6008 OR 1894.6164 OR 1908.632 OR 1922.6476 OR 1936.6632 OR 1950.6788 OR 1964.6944 OR 1978.71 OR 1992.7256 OR 2006.7412 OR 2020.7568 OR 2034.7724 OR 2048.788 OR 2062.8036 OR 2076.8192 OR 2090.8348 OR 2104.8504 OR 2118.866 OR 2132.8816 OR 2146.8972 OR 2160.9128 OR 2174.9284 OR 2188.944 OR 2202.9596 OR 2216.9752 OR 2230.9908 OR 2244.1064 OR 2258.122 OR 2272.1376 OR 2286.1532 OR 2300.1688 OR 2314.1844 OR 2328.1996 OR 2342.2152 OR 2356.2308 OR 2370.2464 OR 2384.262 OR 2398.2776 OR 2412.2932 OR 2426.3088 OR 2440.3244 OR 2454.34 OR 2468.3556 OR 2482.3712 OR 2496.3868 OR 2510.4024 OR 2524.418 OR 2538.4336 OR 2552.4492 OR 2566.4648 OR 2580.4804 OR 2594.496 OR 2608.5116 OR 2622.5272 OR 2636.5428 OR 2650.5584 OR 2664.574 OR 2678.5896 OR 2692.6052 OR 2706.6208 OR 2720.6364 OR 2734.652 OR 2748.6676 OR 2762.6832 OR 2776.6988 OR 2790.7144 OR 2804.73 OR 2818.7456 OR 2832.7612 OR 2846.7768 OR 2860.7924 OR 2874.808 OR 2888.8236 OR 2902.8392 OR 2916.8548 OR 2930.8704 OR 2944.886 OR 2958.9016 OR 2972.9172 OR 2986.9328 OR 3000.9484 OR 3014.964 OR 3028.9796 OR 3042.9952 OR 3056.1008 OR 3070.1164 OR 3084.132 OR 3098.1476 OR 3112.1632 OR 3126.1788 OR 3140.1944 OR 3154.21 OR 3168.2256 OR 3182.2412 OR 3196.2568 OR 3210.2724 OR 3224.288 OR 3238.3036 OR 3252.3192 OR 3266.3348 OR 3280.3504 OR 3294.366 OR 3308.3816 OR 3322.3972 OR 3336.4128 OR 3350.4284 OR 3364.444 OR 3378.4596 OR 3392.4752):TOLERANCEPPM=20 AND
MS2PROD=166.0863:TOLERANCEPPM=20:INTENSITYPERCENT=5 AND
MS2PROD=120.0808:TOLERANCEPPM=20:INTENSITYPERCENT=50

E

I Repository-scale MassQL

451 public Orbitrap datasets
~133 million MS/MS spectra

64 queries

176,732 MS/MS spectra retrieved

II Clustering with MSCluster

- Reduce redundancy: combine nearly identical spectra with same precursor mass
- Keep spectra that appear at least 2x
- 1,474 unique potential *N*-acyl lipids

III Filter based on cosine similarity

- Compounds with reference spectra
- Compounds without reference spectra

Standard cosine similarity:
precursor mass top = precursor mass bottom

Modified cosine similarity:
precursor mass top ≠ precursor mass bottom

• **Cosine > 0.7:** 851 unique *N*-acyl lipids

Spectral library created

IV FASST Search

- Orbitrap and QToF datasets
- 356,542 MS/MS spectra retrieved
- 61,833 files
- 950 datasets
- 39,525 MS/MS in human-related datasets
- 28,497 MS/MS in rodent-related datasets
- 29,105 MS/MS in microbeMASST
- 3,754 MS/MS in plantMASST
- 6,537 MS/MS in foodMASST

V Filter FASST Search:cosine similarity with original spectra

Cosine = 0.8
Filtered spectrum

Cosine = 0.65
Raw spectrum

Keep only matches in which the cosine similarity with **raw spectrum** are >0.7

VI Final numbers

31,299 MS/MS in human-related datasets
21,866 MS/MS in rodent-related datasets
22,589 MS/MS in microbeMASST
2,931 MS/MS in plantMASST
5,576 MS/MS in foodMASST