Supplementary Note - Expansion of Box1

Mass Spectrometry Searches using MASST.

Authors: Mingxun Wang[1,2], Alan K. Jarmusch[1], Fernando Vargas[1,14], Alexander A. Aksenov[1], Julia M. Gauglitz[1], Kelly Weldon[1,3], Daniel Petras[1], Ricardo da Silva[1], Robert Quinn[1, 5], Alexey V. Melnik[1], Justin J.J. van der Hooft[1,6], Andrés Mauricio Caraballo Rodríguez[1], Louis Felix Nothias[1], Christine M. Aceves[1], Morgan Panitchpakdi[1], Elizabeth Brown[1], Francesca Di Ottavio[12], Nicole Sikora[1], Emmanuel O. Elijah[1], Lara Labarta-Bajo[14], Emily C. Gentry[1], Shabnam Shalapour[15], Kathleen E. Kyle[10], Sara P. Puckett[11], Jeramie D. Watrous[13], Carolina S. Carpenter[3], Amina Bouslimani[1], Madeleine Ernst[1], Austin D. Swafford[3], Elina I. Zúñiga[14], Marcy J. Balunas[11], Jonathan L. Klassen[10], Rohit Loomba[3,15], Rob Knight[3,4,8], Nuno Bandeira[3,8,9], Pieter C. Dorrestein[1,3,4,7,9]

Addresses:

1. Collaborative Mass Spectrometry Innovation Center, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego.
2. Ometa Labs LLC.
3. Center for Microbiome Innovation, University of California San Diego.
4. Department of Pediatrics, University of California San Diego.
5. Michigan State University
6. Bioinformatics Group, Wageningen University, Wageningen, The Netherlands.
7. Department of Pharmacology, School of Medicine, University of California San Diego, La Jolla, CA 92093, USA
8. Department of Computer Science and Engineering, University of California San Diego.
9. Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego.
10. Department of Molecular and Cell Biology, University of Connecticut, Storrs, CT, USA
11. Division of Medicinal Chemistry, Department of Pharmaceutical Sciences, University of Connecticut, Storrs, CT, USA
12. Faculty of Bioscience and Technology for Food, Agriculture, and Environment, University of Teramo, TE, Italy
13. Department of Medicine, University of California, San Diego, California, USA
14. Division of Biological Sciences, University of California San Diego, La Jolla, San Diego, CA, 92093, USA
15. Division of Gastroenterology, University of California San Diego.

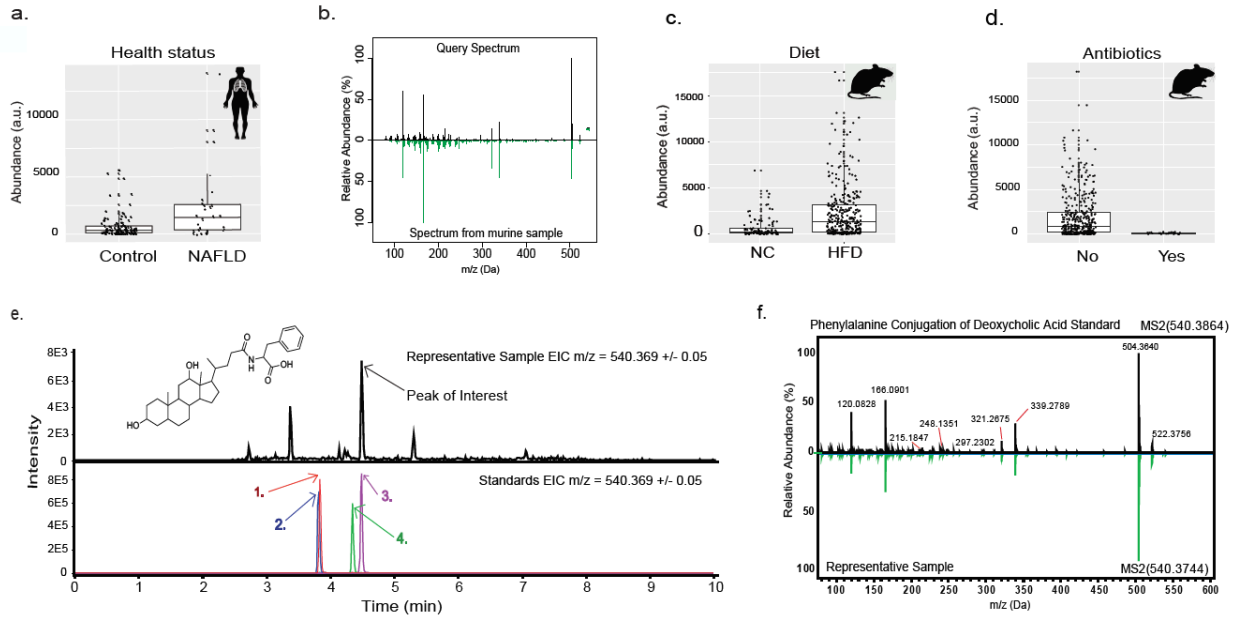
Expansion and references for the ten applications of MASST from Box 1.

1) Are specific molecular features detected via mass spectrometry in one clinical cohort also observed elsewhere? Nonalcoholic fatty liver disease (NAFLD) is the leading cause of

chronic liver disease in the United States, afflicting 80-100 million Americans. It can be broadly sub-divided into two categories: Nonalcoholic fatty liver (NAFL), which is the non-progressive form of NAFLD with minimal risk of progression to cirrhosis, and nonalcoholic steatohepatitis (NASH), with substantial risk of progression to cirrhosis.[1] The fecal samples of individuals with NAFLD with and without advanced fibrosis (n=28) with corresponding healthy controls of close relatives (twins, siblings-siblings and parents-offspring, n=110) were analyzed using untargeted data dependent mass spectrometry. Partial least squares- discriminant analysis, Random Forest and other statistical methods consistently suggested that a mass spectrometry feature with *m/z* 540.3677 was most strongly associated with diagnosed NAFLD (Box Figure 1a). We had no other knowledge about this molecular feature other than its precursor mass and mass fragments. MASST revealed that the MS/MS spectrum associated with this feature was found in 43 datasets (12 mouse fecal, 4 rat fecal, 1 *Escherichia coli* and 20 human fecal studies and several others). Deeper inspection of the datasets revealed 2 studies of mouse models of NAFLD. MASST allows one to find these datasets that can be investigated further. Upon re-exploring the data for one of the latter studies[1], this molecule was found in higher abundance in mice fed a high-fat diet, a condition that induced NAFLD in these animals thus supporting the discovery that this molecule may indeed be NAFLD-related (Box Figure 1 b and c). Furthermore, this molecule disappeared when animals were treated with antibiotics[1], suggesting a microbial origin of this molecule. To gain additional insight, the data were subjected to molecular networking which revealed that the MS/MS spectrum was related chenodeoxycholate but with a mass shift of phenylalanine. This putative molecule would be similar to the related to microbially derived cholyphenylalanine by a loss of an oxygen.[2] As the mass spectrometry is largely blind to regiochemistry, four different phenylalanine amidate conjugates were synthesized. The phenylalanine conjugate of the deoxycholic acid co-migrated with this feature, classifying the annotation as level one according to the 2007 Metabolomics Standards Initiative (Box Figure 1e).[3]

Taken together, these findings constitute a discovery that the most significant differentiating molecule associated with NAFLD in fecal samples is a microbially-derived bile acid. This molecule is observed in the human cohort, but is also recapitulated in the murine model of the disease. This serves as an important confirmation that did not require any additional experimental analysis due to availability and reuse of existing public data. A video with the face behind the science for this MASST analysis job can be found here https://youtu.be/sHHlVTCoQJY

MASST_GNPS job link:
https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=c781fd2cdd534578b30a37b4fc692a9f



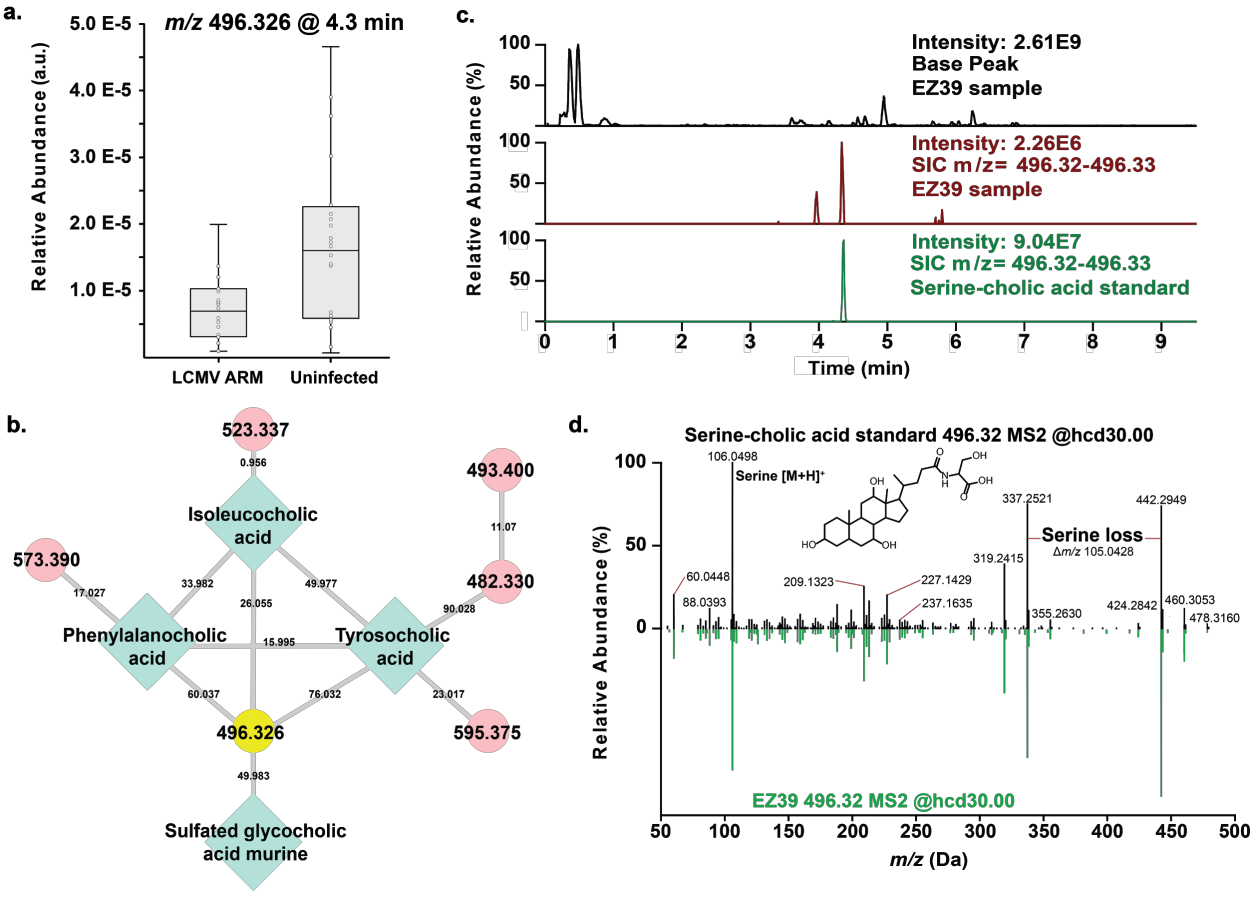
Box Figure 1. Feature with *m/z* 540.3677 associated with a nonalcoholic fatty liver disease (NALFD). a. Box and whisker plot for the human data, p=9.91E-07 (T-test); b. Mirror plot of the MS/MS spectrum from the NALFD cohort used as MASST query compared to a spectrum found in murine cirrhosis study; c. The levels of this molecular feature in fecal samples from mice fed a high fat diet (HFD) vs normal chow (NC), p=2.7E-08; d. The levels of this molecular feature when mice on a HFD are treated with antibiotics p=7.2E-37; e. Extracted ion chromatograms (EICs) of the feature of interest for a representative sample (top panel) and four synthetic standards (bottom panel, phenylalanine amidate conjugate of 1) ursodeoxycholic acid; 2) hyodeoxycholic acid; 3) deoxycholic acid; 4) chenodeoxycholic acid). f. Mirror plot of MS/MS of the feature of interest observed in a representative sample compared to that of the synthetic standard of phenylalanine conjugate of deoxycholic acid.

2) Can findings about a molecule identified in model organism studies be translated to humans? One major application expected will be the translation of molecular information from animal models to humans. The first example of such a translational application was shown during the development of MASST.[2] In a murine model of acute infection with lymphocytic choriomeningitis virus Armstrong (LCMV ARM)[4], it was observed that the ileum of infected and uninfected mice contained an MS/MS spectrum with a precursor mass *m/z* 496.326, which was significantly reduced in abundance at day 8 post-infection when compared to uninfected controls (Box Figure 2a). There were neither matches to any reference spectrum nor could any other match to this molecular feature be found in the metabolomics literature. To test whether this molecule is also found in humans, thus supporting the translational potential, a MASST search was performed. MASST revealed that the same MS/MS spectrum was found in another murine dataset and two human datasets; the American Gut Project (2 males and 7 females all of whom were >45 years of age)[5] and in four samples of children less than 2 years in an infant eating behavior study. Because this molecular feature was also observed in fecal samples in humans, this molecule was prioritized for structural determination. A molecular network suggested this molecule (Yellow circle, Box Figure 2) was related to a recently discovered set of microbially

synthesized bile acids (Green diamonds, Box Figure 2).[2] Molecular networking suggested that serine was conjugated to cholic acid similar to the phenylalanine, leucine/isoleucine and tyrosine conjugates. Precursor mass shifts between the amino acid conjugates of cholic acid support that serine was conjugated (Box Figure 2). Indeed, comparison of our annotation to a synthetic standard of cholylserine showed identical precursor masses (*m/z*), retention time (RT), and MS/MS spectrum which is level one identification according to the 2007 Metabolomics Standards Initiative (Box Figure 2).[3] A video with the face behind the science for this MASST analysis job can be found here https://youtu.be/SExVUrD56-s

MASST_GNPS job link:
https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=bac3d3788e704af59e4a15a5146e4d6b

Molecular networking job
https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=1809afb755794081b128585409100343



Box Figure 2 Modulation of a cholylserine bile acid in a mouse model of acute infection with LCMV. a. 10 week-old C57BL/6J mice were intravenously infected with $2x10^6$ pfu LCMV ARM. Untargeted LC-MS/MS and subsequent identification of a molecular feature with an *m/z* at 496.326 was performed in the ileums of infected mice and uninfected controls at day 8 post-infection. The relative intensities were based on the total ion current of that feature. Pooled data

from 2 independent experiments are shown**.** **p<0.01 by Mann-Whitney test. b. The molecular family within the molecular network for this MS/MS spectrum is shown in the aforementioned mice. Green diamonds are nodes with spectral annotations to the reference library, red circles do not have annotations. Yellow circle is the node of interest. c. The chromatogram with retention time of the synthetic standard for cholylserine vs a representative uninfected ileum sample are shown. d. The MS/MS spectrum of the synthetic standard compared to the to the MS/MS of the sample (green).

3) Can MASST be used to reveal the presence and distribution of environmental toxins? Domoic acid became famous through the novel "The Birds" by Daphne du Maurier and a film from Alfred Hitchcock. This neurotoxin caused seagulls to attack humans. In real life, it has been responsible for numerous poisonings of humans and sea animals, has caused several fatalities[6], has a negative economic impact such as shutting down crabbing in California and is monitored in many coastal areas on the West and Northeast coast of the United States. From a MASST search, spectral matches to domoic acid were found in seven different public datasets, and although more than 1,000 metabolomics datasets were searched, all seven matches were marine related, including culturing experiments of the diatom *Pseudonitzschia,* one of the known domoic acid producers. Spectral matches to domoic acid were found in datasets from the California coast, including data from the Scripps pier and surrounding beaches in San Diego, an area where domoic acid has been observed frequently. In addition to places where it has been observed, matches to domoic acid were also present in five of the LC-MS/MS runs from data from surface seawater collected in Narragansett Bay, Rhode Island, originally deposited in MetaboLights (MTBLS293)[7] and a dataset from Hawaiian coral reef water. This is surprising because these areas have few reported *Pseudonitzschia* blooms. Thus, the re-purposing of non-targeted scientific surveys of marine systems, that are costly to generate, might offer a strategy to prioritize regions for expanded toxin monitoring. A video with the face behind the science for this MASST analysis job can be found here https://youtu.be/vm6UkYwDGn4

MASST_GNPS job link
https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=4d3efb880fda4b8bae36dbfe5b5d6e42

4) In what datasets can we find a published MS/MS spectrum? A class of potentially very important compounds has been described in the literature: N-acyl amide lipids were predicted to exist via bioinformatics analysis, and subsequent heterologous expression of a gene cluster from human-associated bacteria demonstrated their existence.[8] Because of their structural similarity to endogenous signaling molecules of eukaryotes, these compounds are capable of interacting with G-protein-coupled receptors (GPCRs), thus providing a potential way for microbiota to manipulate physiology of the host.[8] This finding brings about tremendous opportunities for development of new drugs and therapies. Previously, we have revealed the presence of these molecules in human stool samples.[5] Further MASST search with the MS/MS of 3-hydroxyhexadecanoyl glycine and 3-hydroxypentadecanoyl lysine suggests that these molecules have a very wide ecological distribution. These MS/MS spectra appear in roughly 10% of all public datasets. In addition to human fecal samples, they were also found in data from fecal samples of mice, rats, human and bovine teeth, isolates of various bacteria including multiple *Streptomyces, Amycolatopsis, Pseudomonas, Neisseria, Achromobacter and Bacillus* species, but also, in multiple marine samples including coral reefs, open ocean waters, marine sediments, as

well as soils and even human habitat. Other molecules that have such wide distributions include phospholipids. Perhaps the GPCR response to these molecules is an evolutionary result of microbial co-existence, a hypothesis that was formulated on the basis of searching these published MS/MS spectra. A video with the face behind the science for this MASST analysis job can be found here https://youtu.be/8W2BCxtszIA

MASST_GNPS job link:
3-hydroxyhexadecanoyl glycine
https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=c853cad1dee04d25a82ed7d0ad1faf61
3-hydroxypentadecanoyl lysine
https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=f0dbb65d92624d0090b2219dd8789ea8

5) Are specific natural products observed in cultured microbes also observed in non-laboratory settings? Many of our therapeutics in use today are derived from microbial natural products.[9] Many natural products, including natural products widely used in the clinic, have not been detected in environmental samples (as opposed to cultures in the lab) and therefore their ecological roles have not been widely established.[10] Thus, questions that the natural products research and related fields often consider to answer is whether a molecule found in the laboratory might also be present in natural environments, and if so, where? An example of such a natural product belongs to the viscosin family of pseudomonads derived molecules: orfamides.[11,12] Several biological activities have been reported for this family of microbial cyclic lipopeptides, supporting their potential as biocontrol agents.[11,12] While they have been isolated from microbial cultures, they have not been observed in a non-laboratory setting. A MASST search with the MS/MS spectrum of *m/z* 1295.84, corresponding to orfamide A, revealed four datasets that contained this molecular ion. These datasets include not only *Pseudomonas* isolate collections,[36] but also field-collected *Trachymyrmex septentrionalis* fungus gardens collected from the eastern USA. Ant fungus gardens are intricate multi-microbial systems that are both a home and food source for ants and their larvae.[13-16] These results suggest the presence and, perhaps, an as-yet-undetermined role of pseudomonads in natural ant fungus garden ecosystems. To investigate this hypothesis, we analyzed the publicly available *T. septentrionalis* fungus gardens in NCBI and confirmed the widespread existence of pseudomonads in these environments, consistent with previous work in related systems.[13-16] Furthermore, we isolated several pseudomonads from *T. septentrionalis* fungus gardens. In combination, these findings suggest that orfamides and related molecules might also play a previously unrecognized role in ant fungus garden ecosystems. A video with the face behind the science for this MASST analysis job can be found here https://youtu.be/4Zb5gZIabBU

MASST_GNPS job link:
https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=bf28bda6ccdd4b699f9a5f2a67b93ee6

6) Where do we find agricultural fungicides in the environment? Is there evidence that people may be in contact with these fungicides? Fungicides are widely used in agricultural systems to control for devastating losses worth millions of dollars in crops worldwide. Exposure to fungicides by humans may occur through contact with the skin, ingestion or inhalation.[17-20] A MASST search with the MS/MS spectrum of azoxystrobin, a fungicide of the strobilurin family

commonly used in agriculture world-wide revealed matches in expected datasets, including two datasets containing a standard to azoxystrobin, a dataset that sampled the surface of fruits, and several food datasets, especially fruits and vegetables. Deeper inspection of the food data-sets revealed that azoxystrobin was found in a mandarin, pressed juices, herbed goat cheese, restaurant grain dish, cucumber (with peel), roma tomato (with peel), duck, orange juice, peanut butter, red sauce, tomato pesto, dried parsley, muffin, mandarin orange flesh, mandarin orange peel, potatoes with herbs, raw green onion, green grapes, and raisin. There was no match to any environmental datasets, but a large number of matches were observed from human skin samples in six different datasets. In total 78 LC-MS/MS files from 15 out of 135 volunteers in all six studies revealed the presence of a spectral match to azoxystrobin from sampling sites included hands, faces and feet suggesting that a significant population might be exposed to azoxystrobin. A video with the face behind the science for this MASST analysis job can be found here https://youtu.be/hGemmjdeOY0

MASST_GNPS job link:
https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=f3fbe2baf0be4641bf14e8f60a79494d

7) Are known toxins from food found in/on people? The mycotoxin roquefortine C is a potent neurotoxin at high concentrations, but is present at low concentrations in foods such as blue cheeses.[21,22] The LD50 is 169-189 mg/kg by intraperitoneal administration. Roquefortine C levels ranging from 0.05 to 12 mg/kg have been reported in cheeses. While it has low toxicity in humans due in part to low bioavailability, the routes of exposure to toxicants and their potential sources are of interest for food safety. As a toxicant exemplar we identified roquefortine C in a number of blue cheese samples. Through a MASST search we also obtained MS/MS matches to two NIH natural products standards reference collections that include roquefortine C as a standard, a *Penicillium* culture, as well as additional blue cheese bacterial and fungal isolates, a fungal collection, and ocean microbial cultures and human stool (infants and adults). The sources and routes of exposure can be thus proposed, beginning with detection in microbial culture, in fermented food products, and ultimately in stool from people. A video with the face behind the science for this MASST analysis job can be found here https://youtu.be/04RSsOY0oGM

MASST_GNPS job link:
https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=c1da96a2d4a74568a3394d2dadf2ffff

8) Can we use approximate matches to a natural product to find datasets that may contain analogs? Staurosporine is a natural indolocarbazole discovered in 1977 from a culture of *Streptomyces staurosporeus* (now *Lentzea albida*). obtained from a soil sample collected in Japan.[23] A decade after that discovery, its strong cytotoxic effect against cancer cells through its potent modulation of tyrosine kinases was demonstrated.[24] Here, we leverage the spectra of the staurosporine reference spectrum with MASST search on GNPS to assess its occurrence in the public MS/MS datasets, and to discover potential new derivatives. A MASST search was performed using the staurosporine available in the GNPS library (CCMSLIB00000001655). 14 datasets showed the presence of related spectra, mostly associated with microbes. Spectral matches to staurosporine, were observed in samples originating from four Actinobacteria

datasets, which was consistent with previous investigations,[25,26] but also directly in soil and marine sediments. MASST in analog mode revealed several candidate analogs. Interestingly, amongst these potential analogs, only one spectral match corresponded to previously described derivatives (17-OH staurosporine, +15.98 Da), while all the other annotations were undescribed putative staurosporine derivatives (including +CH_2, +14.00 Da; +NO, +29.98 Da; +CHN_2O; +32.99 Da). The search for staurosporine derivatives among the public datasets with MASST took less than 15 min, and suggests that there are still yet to be discovered reservoirs of unique staurosporine derivatives. A video with the face behind the science for this MASST analysis job can be found here https://youtu.be/04RSsOY0oGM

GNPS_MASST job link:
https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=e5b04be34a2f41cb9b681e472d9a1765

9) Can MASST be used to track sunscreens in human and environmental samples? Understanding the impact of humans on earth's ecosystems is of increasing importance. Sunscreens became part of the formulation of several skin care products, and they are used worldwide on a regular basis to protect the skin from the damaging effects of ultraviolet radiations including sunburn and skin cancer. [27] However, when it leaves our skin where does it end up? A MASST search of the MS/MS spectra of two active ingredients of sunscreen - avobenzone and octocrylene – reveals, as expected, their presence in many human skin datasets. Avobenzone and octocrylene were detected in skin samples from 19 public datasets, including 15 datasets from the United States, 2 datasets from surfers from Morocco and England, 1 dataset from Japan and 1 dataset from individuals living in Venezuela. Additionally, avobenzone and octocrylene were detected in saliva, teeth and stool. While skin photoprotection is crucial for human health, many questions have been raised regarding the accumulation of sunscreens in the environment. [27-30] To look for sunscreen in environmental samples, MASST found matches to the indoor environment or personal objects such as offices, houses, cars, bikes, phones, wallets, keys, mattresses, plants, meat for human consumption, corals, and even in coral reef in remote areas such as Moorea.[31-33] While there have been no identifiable toxic effects on humans [34], these results show that human-made chemicals may be widely distributed without truly understanding the potential long-term impact. A video with the face behind the science for this MASST analysis job can be found here https://youtu.be/Sjv00dpMSQ8

Avobenzone *m/z* 311.165 MASST_GNPS job link:
https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=56a764c64f524a30b5796f5d9124b832

Octocrylene *m/z* 362.211 MASST_GNPS job link:
https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=84346268fdca4144bf71d9cdb42fef02

10) Can we find evidence of opioids exposure in public data? Opioids are used not only as psychotropic drugs but are also prescribed as analgesics for the relief of severe pain. In the US, the use of these substances often begins with a prescription for pain relief.[35,36] Methadone is a synthetic opioid used to treat drug addiction and is also used in palliative care for patients with cancer, HIV, and postoperative related pain.[36] By searching the MS/MS of methadone using

MASST, we found a matching MS/MS spectrum in five datasets. Methadone was detected in stool samples of two subjects from the American gut project dataset, both are males, one is middle adulthood and the other is early adulthood. It was found in 5 patients affected by inflammatory bowel disease (IBD). Those patients suffer from chronic abdominal or musculoskeletal pain and opioids (including methadone) are commonly prescribed as pain relief.[35] 12 of the samples of the ~1000 teeth contained a spectral match to methadone. This observation is supported by the fact that dentists in the US, commonly prescribe opioids as analgesics, particularly for surgical tooth extraction.[36,37] However, it is worth highlighting that two of the teeth samples also matched reference spectra of cocaine; the co-occurrence suggests the rationale for methadone could stem from recreational use or relapse of treatment. Finally, in addition to human samples, we found methadone in one sediment and four water samples belonging to a dataset called Earth Microbiome Project (EMP)_500_Metabolomics, a study on environmental samples collected throughout the US.[38, 39] This suggests that methadone (and likely other drugs) are present in our environment. A video with the face behind the science for this MASST analysis job can be found here https://youtu.be/9hTsXJ6l1Is

Methadone MASST GNPS job link:
https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=afa2d37fddb44af2ab016c4c7493b3fc

Cocaine MASST GNPS job link:
https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=c40a73f588b24975b541788d7f086510

References for Box 1.

1. Shalapour S, Lin XJ, Bastian IN, Brain J, Burt AD, Aksenov AA, Vrbanac AF, Li W, Perkins A, Matsutani T, Zhong Z, Dhar D, Navas-Molina JA, Xu J, Loomba R, Downes M, Yu RT, Evans RM, Dorrestein PC, Knight R, Benner C, Anstee QM, Karin M. Inflammation-induced IgA+ cells dismantle anti-liver cancer immunity. Nature.2017 551(7680):340-345.
2. Robert A. Quinn, Alexey V. Melnik, Alison Vrbanac, Kathryn A. Patras, Mitchell Christy, Zsolt Bodai, Alexander Aksenov, Anupriya Tripathi, Lawton Chung, Greg Humphrey, Morgan Panitchpakdi, Ricardo da Silva, Julian Avila-Pacheco, Clary Clish, Sena Bae, Himel Mallick, Eric A. Franzosa, Jason Lloyd-Price, Robert Bussell, Taren Thron, Andrew T. Nelson, Mingxun Wang, Fernando Vargas, Julia M. Gauglitz, Michael J. Meehan, Emily Gentry, Timothy Arthur, Michael Downes, Ting Fu, Alexis Komor, Orit Poulsen, Brigid S. Boland, John T. Chang, William J. Sandborn, Meerana Lim, Neha Garg, Julie Lumeng, Barbara I. Kazmierczak, Ruchi Jain, Marie Egan, Kyung E. Rhee, Ron Evans, Manuela Raffatellu, Hera Vlamakis, Curtis Huttenhower, Gabriel G. Haddad, Dionicio Siegel, Sarkis Mazmanian, Victor Nizet, Rob Knight and Pieter C. Dorrestein Global Chemical Impacts of the Microbiome Include Microbial Conjugated Bile Acids. in review.
3. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, Fan TW, Fiehn O, Goodacre R, Griffin JL, Hankemeier T, Hardy N, Harnly J, Higashi R, Kopka J, Lane AN, Lindon JC, Marriott P, Nicholls AW, Reily MD, Thaden JJ, Viant MR. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics. 2007 3(3):211-221.

4. Buchmeier MJ, Welsh RM, Dutko FJ, Oldstone MB. The virology and immunobiology of lymphocytic choriomeningitis virus infection. Adv Immunol. 1980 30:275-331
5. McDonald D, Hyde E, Debelius JW, Morton JT, Gonzalez A, Ackermann G, Aksenov AA, Behsaz B, Brennan C, Chen Y, DeRight Goldasich L, Dorrestein PC, Dunn RR, Fahimipour AK, Gaffney J, Gilbert JA, Gogul G, Green JL, Hugenholtz P, Humphrey G, Huttenhower C, Jackson MA, Janssen S, Jeste DV, Jiang L, Kelley ST, Knights D, Kosciolek T, Ladau J, Leach J, Marotz C, Meleshko D, Melnik AV, Metcalf JL, Mohimani H, Montassier E, Navas-Molina J, Nguyen TT, Peddada S, Pevzner P, Pollard KS, Rahnavard G, Robbins-Pianka A, Sangwan N, Shorenstein J, Smarr L, Song SJ, Spector T, Swafford AD, Thackray VG, Thompson LR, Tripathi A, Vázquez-Baeza Y, Vrbanac A, Wischmeyer P, Wolfe E, Zhu Q; American Gut Consortium, Knight R. American Gut: an Open Platform for Citizen Science Microbiome Research. mSystems. 2018 3(3). pii: e00031-18.
6. Lefebvre KA, Robertson A. Domoic acid and human exposure risks: a review. Toxicon. 2010 56(2):218-30.
7. Metabolights dataset https://www.ebi.ac.uk/metabolights/MTBLS293
8. Cohen LJ, Esterhazy D, Kim SH, Lemetre C, Aguilar RR, Gordon EA, Pickard AJ, Cross JR, Emiliano AB, Han SM, Chu J, Vila-Farres X, Kaplitt J, Rogoz A, Calle PY, Hunter C, Bitok JK, Brady SF. Commensal bacteria make GPCR ligands that mimic human signalling molecules. Nature. 2017 549(7670):48-53.
9. Newman, D. J.; Cragg, G. M., Natural Products as Sources of New Drugs from 1981 to 2014. Journal of Natural Products 2016, 79 (3), 629-661.
10. Pye, C. R.; Bertin, M. J.; Lokey, R. S.; Gerwick, W. H.; Linington, R. G., Retrospective analysis of natural products provides insights for future discovery trends. Proceedings of the National Academy of Sciences of the United States of America 2017, 114 (22), 5601-5606.
11. Jang, J. Y.; Yang, S. Y.; Kim, Y. C.; Lee, C. W.; Park, M. S.; Kim, J. C.; Kim, I. S., Identification of Orfamide A as an Insecticidal Metabolite Produced by Pseudomonas protegens F6. Journal of Agricultural and Food Chemistry 2013, 61 (28), 6786-6791.
12. Nguyen, D. D.; Melnik, A. V.; Koyama, N.; Lu, X. W.; Schorn, M.; Fang, J. S.; Aguinaldo, K.; Lincecum, T. L.; Ghequire, M. G. K.; Carrion, V. J.; Cheng, T. L.; Duggan, B. M.; Malone, J. G.; Mauchline, T. H.; Sanchez, L. M.; Kilpatrick, A. M.; Raaijmakers, J. M.; De Mot, R.; Moore, B. S.; Medema, M. H.; Dorrestein, P. C., Indexing the Pseudomonas specialized metabolome enabled the discovery of poaeamide B and the bananamides. Nature Microbiology 2017, 2 (1), 9.
13. Mehdiabadi, N. J.; Schultz, T. R., Natural history and phylogeny of the fungus-farming ants (Hymenoptera: Formicidae: Myrmicinae: Attini). Myrmecological News 2010, 13, 37-55.
14. Aylward FO, Burnum KE, Scott JJ, Suen G, Tringe SG, Adams SM, Barry KW, Nicora CD, Piehowski PD, Purvine SO, Starrett GJ, Goodwin LA, Smith RD, Lipton MS, Currie CR. Metagenomic and metaproteomic insights into bacterial communities in leaf-cutter ant fungus gardens. ISME J. 2012 6(9):1688-701.
15. Aylward FO, Suen G, Biedermann PH, Adams AS, Scott JJ, Malfatti SA, Glavina del Rio T, Tringe SG, Poulsen M, Raffa KF, Klepzig KD, Currie CR. Convergent bacterial microbiotas in the fungal agricultural systems of insects. MBio. 2014 5(6):e02077.
16. Aylward FO, Burnum KE, Scott JJ, Suen G, Tringe SG, Adams SM, Barry KW, Nicora CD, Piehowski PD, Purvine SO, Starrett GJ, Goodwin LA, Smith RD, Lipton MS, Currie CR. Metagenomic and metaproteomic insights into bacterial communities in leaf-cutter ant fungus gardens. ISME J. 2012 6(9):1688-701.

17. Nicolopoulou-Stamati P, Maipas S, Kotampasi C, Stamatis P, Hens L. Chemical Pesticides and Human Health: The Urgent Need for a New Concept in Agriculture. Front Public Health. 2016 4:148.
18. Yoon MY, Cha B, Kim JC. Recent trends in studies on botanical fungicides in agriculture. Plant Pathol J. 2013 29(1):1-9.
19. Bartlett DW, Clough JM, Godwin JR, Hall AA, Hamer M, Parr-Dobrzanski B. The strobilurin fungicides. Pest Manag Sci. 2002 58(7):649-62. Review. Erratum in: Pest Manag Sci. 2004 60(3):309.
20. Casida JE, Durkin KA. Pesticide Chemical Research in Toxicology: Lessons from Nature. Chem Res Toxicol. 2017 30(1):94-104.
21. Tiwary AK, Puschner B, Poppenga RH. Using roquefortine C as a biomarker for penitrem A intoxication. J Vet Diagn Invest. 2009 21(2):237-9.
22. Arnold DL, Scott PM, McGuire PF, Harwig J, Nera EA (1978) Acute toxicity studies on roquefortine and PR toxin, metabolites of Penicillium roqueforti, in the mouse. Food Cosmet Toxicol 16: 369–371.
23. Omura, S. *et al.* A new alkaloid AM-2282 OF Streptomyces origin. Taxonomy, fermentation, isolation and preliminary characterization. *J. Antibiot.* **30**, 275–282 (1977).
24. Tamaoki T, Nomoto H, Takahashi I, Kato Y, Morimoto M, Tomita F. Staurosporine, a potent inhibitor of phospholipid/Ca++dependent protein kinase. Biochem Biophys Res Commun. 1986 135(2):397-402.
25. Duncan KR, Crüsemann M, Lechner A, Sarkar A, Li J, Ziemert N, Wang M, Bandeira N, Moore BS, Dorrestein PC, Jensen PR. Molecular networking and pattern-based genome mining improves discovery of biosynthetic gene clusters and their products from Salinispora species. Chem Biol. 2015 22(4):460-471.
26. Jensen PR, Moore BS, Fenical W. The marine actinomycete genus Salinispora: a model organism for secondary metabolite discovery. Nat Prod Rep. 2015 32(5):738-51.
27. Raffa RB, Pergolizzi JV Jr, Taylor R Jr, Kitzen JM; NEMA Research Group. Sunscreen bans: Coral reefs and skin cancer. J Clin Pharm Ther. 2019 44(1):134-139.
28. Schneider SL, Lim HW. Review of environmental effects of oxybenzone and other sunscreen active ingredients. J Am Acad Dermatol. 2019 80(1):266-271.
29. Siller A, Blaszak SC, Lazar M, Olasz Harken E. Update About the Effects of the Sunscreen Ingredients Oxybenzone and Octinoxate on Humans and the Environment. Plast Surg Nurs. 2018 38(4):158-161.
30. Ruszkiewicz JA, Pinkas A, Ferrer B, Peres TV, Tsatsakis A, Aschner M. Neurotoxic effect of active ingredients in sunscreen products, a contemporary review. Toxicol Rep. 2017 4:245-259.
31. Kapono CA, Morton JT, Bouslimani A, Melnik AV, Orlinsky K, Knaan TL, Garg N, Vázquez-Baeza Y, Protsyuk I, Janssen S, Zhu Q, Alexandrov T, Smarr L, Knight R,Dorrestein PC. Creating a 3D microbial and chemical snapshot of a human habitat. Sci Rep. 2018 8(1):3669.
32. Petras D, Nothias LF, Quinn RA, Alexandrov T, Bandeira N, Bouslimani A, Castro-Falcón G, Chen L, Dang T, Floros DJ, Hook V, Garg N, Hoffner N, Jiang Y, Kapono CA, Koester I, Knight R, Leber CA, Ling TJ, Luzzatto-Knaan T, McCall LI, McGrath AP, Meehan MJ, Merritt JK, Mills RH, Morton J, Podvin S, Protsyuk I, Purdy T, Satterfield K, Searles S, Shah S, Shires S, Steffen D, White M, Todoric J, Tuttle R, Wojnicz A, Sapp V, Vargas F, Yang J, Zhang C,

Dorrestein PC. Mass Spectrometry-Based Visualization of Molecules Associated with Human Habitats. Anal Chem. 2016 88(22):10775-10784.
33. Bouslimani A, Melnik AV, Xu Z, Amir A, da Silva RR, Wang M, Bandeira N, Alexandrov T, Knight R, Dorrestein PC. Lifestyle chemistries from phones for individual profiling. Proc Natl Acad Sci U S A. 2016 113(48):E7645-E7654.
34. Ruszkiewicz JA, Pinkas A, Ferrer B, Peres TV, Tsatsakis A, Aschner M. Neurotoxic effect of active ingredients in sunscreen products, a contemporary review. Toxicol Rep. 2017 4:245-259.
35. Long MD, Barnes EL, Herfarth HH, Drossman DA. Narcotic use for inflammatory bowel disease and risk factors during hospitalization. Inflamm Bowel Dis. 2012 18(5):869-76.
36. Baker JA, Avorn J, Levin R, Bateman BT. Opioid Prescribing After Surgical Extraction of Teeth in Medicaid Patients, 2000-2010. JAMA. 2016 315(15):1653-4.
37. Baker JA, Avorn J, Levin R, Bateman BT. Opioid Prescribing After Surgical Extraction of Teeth in Medicaid Patients, 2000-2010. JAMA. 2016 315(15):1653-4.
39. David Hanigan, E. Michael Thurman, Imma Ferrer, Yang Zhao, Susan Andrews, Jinwei Zhang, Pierre Herckes, and Paul Westerhoff Methadone Contributes to N-Nitrosodimethylamine Formation in Surface Waters and Wastewaters during Chloramination Environmental Science & Technology Letters 2015 2 (6), 151-157.
40. Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, Prill RJ, Tripathi A, Gibbons SM, Ackermann G, Navas-Molina JA, Janssen S, Kopylova E, Vázquez-Baeza Y, González A, Morton JT, Mirarab S, Zech Xu Z, Jiang L, Haroon MF, Kanbar J, Zhu Q, Jin Song S, Kosciolek T, Bokulich NA, Lefler J, Brislawn CJ, Humphrey G, Owens SM, Hampton-Marcell J, Berg-Lyons D, McKenzie V, Fierer N, Fuhrman JA, Clauset A, Stevens RL, Shade A, Pollard KS, Goodwin KD, Jansson JK, Gilbert JA, Knight R; Earth Microbiome Project Consortium. A communal catalogue reveals Earth's multiscale microbial diversity. Nature. 2017 551(7681):457-463.