Published in final edited form as:

J Am Soc Mass Spectrom. 2019 August ; 30(8): 1373–1384. doi:10.1007/s13361-019-02207-5.

Mass spectrometry-based integration and expansion of the chemical diversity harbored within a marine sponge

Thomas P. Cantrell¹, Christopher J. Freeman^{2,3}, Valerie J. Paul², Vinayak Agarwal^{1,4,*}, Neha Garg^{1,*}

¹School of Chemistry and Biochemistry, Georgia Institute of Technology, GA 30332, USA

²Smithsonian Marine Station, Smithsonian Institution, FL 34949, USA

³Department of Biology, College of Charleston, SC 29424, USA

⁴School of Biological Sciences, Georgia Institute of Technology, GA 30332, USA

Abstract

Marine sponges and their associated symbionts produce a structurally diverse and complex set of natural products including alkaloids, terpenoids, peptides, lipids, and steroids. A single sponge with its symbionts can produce all of the above-mentioned classes of molecules and their analogs. Most approaches to evaluating sponge chemical diversity have focused on major metabolites that can be isolated and characterized; therefore, a comprehensive evaluation of intra- (within a molecular family; analogs) and inter-chemical diversity within a single sponge remains incomplete. We use a combination of metabolomics tools, including a supervised approach via manual library search and literature search, and an unsupervised approach via molecular networking and MS2LDA analysis to describe the intra and inter-chemical diversity present in *Smenospongia aurea*. Furthermore, we use imaging mass spectrometry to link this chemical diversity to either the sponge or the associated cyanobacteria. Using these approaches, we identify 7 more molecular features that represent analogs of four previously known peptide/polyketide smenamides and assign the biosynthesis of these molecules to the symbiotic cyanobacteria by imaging mass spectrometry. We extend this analysis to a wide diversity of molecular classes including indole alkaloids and meroterpenes.

INTRODUCTION

Characterizing the chemical diversity within a holobiont, defined as a biological unit including both a eukaryotic host and the prokaryotic and eukaryotic communities living within it, comes with complex analytical challenges. Intra-community interactions can be symbiotic, in which the members satisfy nutritional deficiencies or fulfill other important biological or ecological roles and are hence tightly associated, or can be commensal in nature, leading to variable and transitional associations. Intra-community interactions rely on a molecular language. The genome of each individual community member encodes small

*Correspondence to: vagarwal@gatech.edu; neha.garg@chemistry.gatech.edu, Address reprint requests to, Name: Neha Garg, Georgia Institute of Technology, Engineered Biosystems Building, 950 Atlantic Drive, Atlanta, GA 30332-2000, Phone: 404.385.5677, Fax: 404.385.4620, neha.garg@chemistry.gatech.edu.

organic molecules and larger biomolecules that are the alphabets of this molecular language. Challenges in the molecular characterization of a holobiont includes cataloging the large diversity of molecules present in them, gaining insights into their chemical structures, and deciphering the individual producers of these molecules. Advances in mass spectrometry instrumentation and data analysis workflows are making these questions increasingly accessible [1–6]. On one hand, high-resolution mass spectrometers enable the acquisition of large metabolomic datasets and mass spectrometry imaging (MSI) workflows allow for querying the spatial distribution of analytes in a biological sample [7, 8]. On the other hand, metabolomic data curation tools are enriching mass spectral libraries with the goal of improving metabolite identifications and providing unifying platforms for bringing metabolomics data in the public space [5, 6, 9–11]. A combination of contemporary analytical tools is providing a transformative view of the chemistry and biology of complex biomes.

Sponges (phylum Porifera) are sessile benthic marine invertebrates that are ubiquitous across the globe. Sponges are the richest source of small organic molecules, colloquially referred to as *natural products*, in marine ecosystems [12]. Several decades of research have led to discovery of over 10,000 natural products from sponges (<http://pubs.rsc.org/marinlit/>; accessed January 31, 2019). Numerous sponge-derived natural products are endowed with desirable pharmacological bioactivities that have motivated their discovery, structure elucidation, and efforts towards chemical syntheses. Conversely, sponge-derived natural products can be potent toxins and environmental pollutants. Divested from their anthropogenic uses and impacts, sponge natural products play important ecological roles in their native habitat. Sponges are also hosts to a rich and diverse microbiome, as well as other eukaryotes such as fungi [13–15]. Despite intensive interest in sponge natural product chemistry, several key questions remain. Principal among these are: (i) a comprehensive overview of natural product chemical families present in a sponge metabolome, (ii) inventorying, using publicly accessible data, the known and unknown sponge-derived chemistry, and (iii) gaining insight on the identities of the sponge community members that produce these natural products.

We describe here the application of metabolomic curation tools and MSI to investigate the metabolome of the Caribbean marine sponge *Smenospongia aurea* (class- Demospongia, order- Dictyoceratida). *S. aurea* is highly abundant tropical shallow water marine sponge and has been the subject of extensive chemical investigation over the last four decades which have led to description of several natural products [16–20]. Our findings describe the *scale of the problem*, that is, just how diverse the *S. aurea* metabolome is and how little of it has been mapped, and provide insights into the production of natural products by certain members of the *S. aurea* associated microbiome. Data described here have been made publicly accessible which will enable subsequent community curation of the *S. aurea* metabolome, and enable the structural curation of identical or similar molecules in other metabolomics datasets. Molecular annotations are based upon accurate mass, isotopic pattern, matching of fragmentation spectra (*Level 2* annotation standard as proposed by the Metabolomics Society standard initiative) and matching of retention times when standards are available (*Level 1* annotations) [21, 22].

MATERIALS AND METHODS

Mass spectrometry data acquisition method development

Extraction of the sponge tissue is described in detail in Supplementary Information. The extracted metabolites were dissolved in methanol (LC-MS grade) and analyzed with Agilent 1290 Infinity II UHPLC system (Agilent Technologies) using a Kinetex™ 1.7 µm C18 reversed phase UHPLC column (50 × 2.1 mm) coupled to a ImpactII ultra-high resolution Qq-TOF mass spectrometer (Bruker Daltonics, GmbH, Bremen, Germany) equipped with ESI source. The gradient employed for chromatographic separation was 5% solvent B (100% MeCN, 0.1% formic acid) with solvent A as 100% H₂O, 0.1% formic acid for 3 min, a linear gradient of 5% B-95% B in 17 min, held at 95% B for 3 min, 95% B - 5% B in 1 min and held at 5% B for 2 min at a flow rate of 0.5 mL/min throughout the run. MS spectra were acquired in positive ion mode from *m/z* 100–2000 Da. An external calibration with ESI-L Low Concentration Tuning Mix (Agilent Technologies) was performed prior to data collection and internal lock-mass calibrant Hexakis(1H,1H,2H-perfluoroethoxy)phosphazene was used throughout the run. The capillary voltage of 4500 V, nebulizer gas pressure (N₂) of 2 bar, ion source temperature of 200 °C, dry gas flow of 9 L/min, spectral rate of 3 Hz for MS¹ and 6 Hz for MS² was used. For acquiring MS² data, seven most intense ions per MS¹ were selected, and collision induced dissociation energies in Table S1 were used. Basic stepping function was used to fragment ions at 50% and 125% of the CID calculated for each *m/z* from Table S1 with timing of 50% for each step. Similarly, basic stepping of collision RF of 550 and 800 Vpp with a timing of 50% for each step and transfer time stepping of 75 and 90 µs with a timing of 50% for each step was employed. MS/MS active exclusion parameter was set to 2 and released after 30 sec. The mass of internal lock-mass calibrant was excluded from the MS² list. Data analysis by molecular networking and MS2LDA is described in detail in Supplementary Information

Fluorescence microscopy

Sponge tissue sectioning is described in the Supplementary Information. Fluorescence microscopy was performed at 20× on sponge tissue cross-sections prior to either higher resolution fluorescence imaging at 100×, or, MALDI-MSI. Vacuum desiccated samples were imaged with a Plan-Apochromat 20×/0.8 M27, from Carl Zeiss Microscopy (Thornwood, NY, US), under ambient conditions. All fluorescent images were acquired using a Zeiss LSM 780 AxioObserver and processed in Zeiss Zen Black (Version 14.0.12.201). Images were collected in tile scans spanning the entire sample area with 10% frame overlap. The 8 µm sections were imaged at 100× magnification using an alpha Plan-Apochromat 100×/1.46 Oil DIC M27 Elyra objective from Carl Zeiss Microscopy (Thornwood, NY, US). These sections were mounted in Prolong Gold antifade (Invitrogen, Carlsbad, CA, US) following 60 min vacuum desiccation. The sections in Prolong Gold were cured for 24 h under a coverslip and subsequently sealed. Chlorophyll was imaged using 488 nm excitation and emission was collected at 647–721 nm. Using spectral imaging an autofluorescence profile was used to visualize sponge tissue. For the autofluorescence profile of sponge tissue, an excitation of 488 nm was used and emission band was collected from 493 to 556 nm. Fluorescence channels were the same at 20× and 100× resolution. Linear contrast enhancements were applied to images for clarity.

MALDI imaging data acquisition

Optical images were acquired using a gel scanner at 1200 dpi resolution prior to fluorescence imaging and MALDI matrix application. Deposition of MALDI matrix (5 mg/mL DHB solution in MeCN:H₂O 60:40, 0.2% formic acid) was performed utilizing iMatrix sprayer in 20 passes with a spray density of 1 $\mu\text{L}/\text{cm}^2$. The matrix-coated slides were dried under vacuum for 60 min. MALDI-MSI data was acquired using a Bruker rapifleX MALDI TissueTyper™ system operating in reflectron positive mode from m/z 200–2000 Da (Bruker Daltonik GmbH, Bremen, Germany). The smartbeam 3D laser was operated in a single scan mode setting (imaging $-20\text{ }\mu\text{m}$) and the image acquired using a pitch of $20\text{ }\mu\text{m}$. The laser was operated at 10 kHz and 200 shots per pixel were recorded. Calibration was performed prior to MALDI-MSI data acquisition using red phosphorous deposited on the ITO slide. Instrument control was achieved via flexControl 4.0 (Bruker Daltonik GmbH, Bremen, Germany) and imaging acquisition was performed using flexImaging 5.0 (Bruker Daltonik GmbH, Bremen, Germany). Total ion current was utilized as normalization method. To display two-dimensional ion intensity maps, false colors are assigned to ions in the overall average mass spectrum in the flexImaging software or SCiLS Lab Pro 3D software (Bruker Daltonik GmbH, Bremen, Germany).

RESULTS AND DISCUSSION

Chemical extraction and data acquisition

S. aurea was collected by hand at a depth of 4–7 m at the Wonderland Reef in the Florida Keys. *S. aurea* was identified by visual inspection with characteristic large oscula and a fleshy yellow interior (Fig. 1a). Sponge specimens were frozen at $-80\text{ }^{\circ}\text{C}$ pending chemical extraction. Total DNA was isolated from sponge tissue preserved in RNAlater and used as template for 18S rDNA amplification using previously described protocols [23]. Sanger sequencing of multiple clones from the 18S rDNA amplicon clone library and comparison against sequence databases by BLAST confirmed the identity of the sponge specimen. Extraction of lyophilized sponge material with MeOH at room temperature yielded a bright green extract which was analyzed by LC-MS/MS without any derivatization or *a priori* fractionation. Sponge tissue was embedded in a polymer matrix and sectioned at low temperature for microscopy. Fluorescence microscopy was used to visualize the cyanobacterial cells within the sponge tissue (Fig. 1b). The cyanobacterial cells, observed by autofluorescence of cyanobacterial chlorophyll, are represented in green. The cyanobacterial cells are localized to the sponge ectosome, as has been observed for other cyanobacteria harboring sponges. The sponge tissue was visualized by spectral imaging and is shown in red.

Marine sponges are among the most prolific producers of small organic molecules in the marine environment with exceedingly complex metabolomes [12]. To describe metabolome diversity from *S. aurea*, we used a molecular networking-based approach to organize and annotate the metabolome (Fig. 2). We describe our findings with data acquired in the positive ionization mode. As most spectral annotations for natural products and other organic molecules are commonly available in the positive ionization mode only, and, as the motivation of the study is to use and enrich the existing spectral libraries for compound

dereplication (*vide infra*), we have used data collected in the positive mode only. Structurally related metabolites result in MS² fragment ions that are either identical in mass or differ by common mass shifts, representing chemical modifications such as methylation, acetylation, oxidation, and glycosylation, among others. In molecular networking, this structural relatedness is represented in the form of molecular networks comprising of nodes and edges. Each node is a metabolite and connecting nodes are structurally related metabolites (molecular family), or can also be isotopes (for example, resulting from presence of halogen atoms), adducts (for example, sodium adducts), and ions with mixed MS² spectra (spectra arising from fragmentation of multiple precursor ions). Molecular networks were generated using the publicly available web-based infrastructure Global Natural Products Social Molecular Networking (GNPS) (<https://gnps.ucsd.edu>) and the MS² spectra are searched against various spectral libraries incorporated within GNPS, such as, MassBank of North America, European MassBank, ReSpec, CASMI, and other third-party spectral libraries [5]. Using the molecular networking approach for the *S. aurea* metabolomics dataset, 14710 MS² spectra were organized into 1615 nodes that formed 149 individual connected clusters consisting of 861 nodes, also called molecular families (Fig. 2). The rest, ~750 nodes, were singletons, implying that no structural neighbors could be identified for these ions by molecular networking. These singletons have been omitted in Fig. 2 for clarity.

By spectral matching to MS² spectral libraries, only 33 of 1615 nodes (2.0%) resulted in putative annotations. This low number of automated annotations reflects the sparse nature of spectral libraries that are contemporarily available, and emphasizes the need to bring more mass spectrometry data into the public domain with careful annotation of small molecules. Library and literature-based manual annotation of data revealed molecular families corresponding to lipids, fatty acids, amino acid metabolites, polyols, various indole alkaloids, terpenes, peptide/polyketide hybrid molecules, and pyrones, which we describe below. The singular network-based representation elegantly demonstrates the tremendous chemical diversity of the *S. aurea* metabolome (Fig. 2). Next, we intended to decipher how much of the sponge chemical diversity was described by the molecular networking approach, and assess whether we could enrich the molecular annotations in the molecular network and use imaging mass spectrometry to gain insight into the differential spatial organization of molecules within the sponge tissue.

Dereplication of peptide/polyketide hybrid molecules

The MS¹ spectra corresponding to the node with m/z 501.252 (Fig. 3, in yellow), to which no spectral match in MS² spectral libraries was present, demonstrated an isotopic pattern consistent with the presence of one chlorine atom. A MarinLit, Dictionary of Natural Products, and PubChem database compound search of exact mass (500.244 Da) and molecular formula C₂₈H₃₇ClN₂O₄ did not yield an acceptable annotation. Literature search of compounds isolated from marine sponges, specifically *S. aurea*, and manual annotation of the MS² ions revealed that the node with m/z 501.252 corresponds to hybrid peptide/polyketides known as smenamides (theoretical [M+H]¹⁺ = 501.252, found = 501.252, error = 0 ppm) [18]. An extracted ion chromatogram (EIC) for m/z = 501.252 ± 0.01 Da demonstrated two distinct peaks, likely corresponding to isomeric smenamide A (**1**) and smenamide B (**2**) that have been described in the literature (Fig. 3a) [18]. Due to identical

MS/MS fragmentation, spectra corresponding to **1** and **2** were condensed in a single node in the molecular network. The connected node, m/z 503.250, is the Cl^{37} isotope of **1**, **2**.

The chemical structure of **1** and **2** bears a vinylogous chlorine appended to a hydrid polyketide-peptide backbone. Due to the presence of this structural motif in cyanobacterial natural products, such as the jamaicamides, cyanobacterial origin of smenamides within the sponge microbiome has been proposed [18, 24]. Symbiotically associated sponge cyanobacteria are resistant to laboratory cultivation, preventing the direct testing of this hypothesis [25]. We sought to determine the correlation between the spatial distribution of smenamides and cyanobacterial cells within the sponge tissue. Frozen *S. aurea* tissue was embedded in 20% gelatin, sectioned in 10 μm thickness slices using a cryomicrotome and thaw-mounted onto prechilled conductive slides. Maintaining a low temperature during sectioning, $-15\text{ }^{\circ}\text{C}$ in this study, was critical for efficient sectioning without damage to the sponge tissue ultrastructure. Querying the sectioned sponge tissue by fluorescence imaging clearly shows that the cyanobacterial symbionts in *S. aurea* are localized to the sponge ectosome, as visualized by the cyanobacterial chlorophyll autofluorescence (Fig. 3b, in purple), and can be distinguished from the sponge choanosomal matrix, as visualized by the autofluorescence of the sponge tissue (Fig. 3b, in green). On the same sponge section, we next employed MALDI-TOF MSI to query the localization of smenamides. An adduct of **1**, **2**, m/z 561.277, with identical retention time and similar fragmentation spectra in ESI-MS/MS was observed by MALDI-TOF MSI (Fig. S1). The ion image of m/z 561.277 co-localizes with the chlorophyll autofluorescence of cyanobacterial cells in the sponge ectosome (Fig. 3c). These findings support the hypothesis that smenamides are likely produced by cyanobacterial symbionts present within *S. aurea*. Our findings complement other molecular approaches, such as physical fractionation of sponge endosymbionts, metagenome sequencing, and hybridization of sequence specific fluorescent nucleotide probes that have been employed previously to establish cyanobacterial origins of organic molecules in the microbiomes of marine invertebrates [23, 26–29].

As delineated above, **1** and **2** have previously been isolated from *S. aurea*. We next sought to interrogate whether analogs of **1** and **2** are present in the sponge. The *S. aurea* molecular network demonstrated an underappreciated diversity of smenamides within the sponge metabolome (Fig. 2, 3d). Various interconnected nodes suggest that previously unknown analogues of smenamides are also produced in *S. aurea*. The node with m/z 467.267 is the condensed node for previously described isomeric analogs smenamide C (**3**) and D (**4**), wherein the phenylalanine-derived benzylic moiety of **1** and **2** is replaced by the leucine side chain (Fig. 3d) [30]. Crucially, **3** and **4** were isolated from a marine cyanobacterial bloom and not from a marine sponge [30], thus illustrating that the biosynthetic potential to generate structurally similar natural products exists in both free-living cyanobacteria and sponge symbionts [31, 32].

The two nodes, m/z 543.299 (in red, Fig. 3f, $m/z = +42.047$ from 501.252 (**1**, **2**)) and m/z 487.237 (in purple, Fig. 3g, $m/z -14.015$ from 501.252 (**1**, **2**)) correspond, respectively, to a difference of $(+)\text{CH}_2$ and $(-)\text{CH}_2$ structural units from **1** and **2**. Where are these differences localized in the chemical structures of **1** and **2**? Comparison of the MS^2 spectra of molecules with m/z 543.299 (Fig. 3f) and m/z 487.237 (Fig. 3g) against that for **1**, **2** (m/z

501.252, Fig. 3e) allow for putative structural assignments. All MS² ions corresponding to the alkyl chain for **1**, **2** are shifted by +42.047 Da in Fig. 3f. Hence, it is tantalizing to propose that the +42.047 Da (+CH₂)₃ mass difference corresponds to the iterative incorporation of an additional methyl malonate derived polyketide extender unit in the smenamamide biosynthetic scheme to generate the smenamamide analog with *m/z* 543.299. This biosynthetic promiscuity, as revealed by molecular networking and annotation of MS² spectra, is not without precedent in polyketide biosyntheses [33]. On the other hand, for the smenamamide analog at *m/z* 487.237 (Fig. 3g), the (-)14.015 Da (-CH₂) difference corresponds to the loss of methyl group (-CH₃, +H) from **1**, **2**. This loss of methyl group likely occurs from the methylated amide moiety at the left periphery of **1**, **2**, as denoted by the conservation of MS² ions at *m/z* 105.070 and *m/z* 161.0133 between Fig. 3e and Fig. 3g, but the (-)14.015 Da shift in the MS² ions at *m/z* 234.186, *m/z* 262.180, and *m/z* 298.157. All five polyketide MS² ions are conserved between Fig. 3g and Fig. 3h. The node with *m/z* 453.252 represents desmethyl analogs of **3**, **4**. The spectra corresponding to all molecules described above have been added to MS² spectral library hosted by GNPS as smenamides and putative smenamamide analogs.

At this point, the diversity of known smenamides has already been increased from the previously reported four molecules, **1–4**, to include five more structurally similar natural products. The well-defined and rich MS² spectra for smenamides (Fig. 3e–h) made us ask if we could even further expand the smenamamide chemical diversity by searching for the conservation of distinct MS² ions- which correspond to structural elements with the smenamamide core structure- across the *S. aurea* metabolome. To perform this search, we employed the online tool MS2LDA which organizes conservation between MS² spectra in the form of ‘motifs’ [34–36]. These motifs consist of a set of common MS² peaks and/or neutral losses that correspond to shared substructures between analogs. Sharing of motifs between different clusters within a network implies structural conservation even when the clusters are not joined together based on the spectral overlap, expressed as the cosine score, falling below the threshold value for cluster grouping. Motifs for all nodes within the smenamamide cluster were tabulated (Fig. S2a) and their conservation across the *S. aurea* molecular network manually curated. We noted that the MS2LDA motif 528 (Fig. S2a) was conserved between multiple nodes in the smenamamide cluster and nodes from a different second cluster (Fig. 2, 3d). In the absence of putative structure for node with *m/z* 530.26 that connects the two clusters, the chemical substructure corresponding to the motif 528 cannot be predicted. However, manual inspection of the MS² spectra for nodes from the first cluster (*m/z* 530.26) and from the second cluster (*m/z* 532.276, 518.260, 504.245, and 461.202, Fig. S2b) clearly demonstrates that this second cluster of nodes also corresponds to smenamamide analogs. While the spectral overlap between nodes from these two clusters is low enough that they are grouped together in separate clusters (Fig. 3d), conservation of MS2LDA derived motifs allowed us to connect the two clusters together. Thus, two clusters in the *S. aurea* network correspond to smenamides (Fig. 2).

In addition to smenamides, hybrid polyketide-peptide molecules smenathiazole A and smenathiazole B were also detected (Fig. S3). Smenamides and smenathiazoles possess potent cytotoxic and neurotoxic activities [18, 19, 30]. Discovery of natural analogs presents

an opportunity to guide structure-activity relationships aiming at synthesis of diverse pharmacophores for drug discovery.

Dereplication of indole alkaloids and their distribution within the sponge

The *S. aurea* metabolome is exceptionally rich in tryptophan-derived indolic natural products. Of these, two classes of molecules are conspicuous in the molecular network. First are the tryptamine derivatives. A large diversity of tryptamine derivatives, and their brominated analogs were detected in the *S. aurea* metabolome. Prominent molecules, tryptamine derivatives **5–10**, and brominated tryptamine derivatives **11–15** are described in Fig. 4a and Table S2. Curiously, clustering based on spectral overlap alone led to the tryptamine and the brominated tryptamine nodes getting organized in two separate clusters (Fig. 4a). Querying conserved MS2LDA motifs between nodes of these two clusters led to the identification of MS2LDA motif 480 (Fig. S4) which is shared between the nodes corresponding to molecules **6** and **14** in the two clusters (Fig. 4a).

The abundance of tryptamine derivatives was particularly high. Indolic natural products possess myriad biological activities. Whether *S. aurea* uses these simple tryptamines as chemical defenses, or to fulfill other ecological roles is presently unclear. The distribution of **7** was interrogated using MALDI-TOF MSI (Fig. 4b). MSI revealed that **7** is present homogenously throughout the sponge tissue. In addition, brominated tryptamines **12** and **14** displayed isotopic pattern consistent with monobromination in MALDI-MSI and had identical uniform distribution throughout the sponge tissue (Fig. 4c, d).

The other class of indolic natural products that were detected in *S. aurea* include the aplysinopsins, tryptophan metabolites that are widely distributed in the marine environment [37]. A large diversity of marine aplysinopsins has been described previously from various different sources, of which a small fraction resides within *S. aurea* [16, 17]. We used molecular networking to dereplicate aplysinopsins **16–18** and visualized their distribution using MALDI-MSI to interrogate whether the distribution matched that of tryptamines. The tryptamine derivatives (Fig. 4b–d) and aplysinopsins (Fig. 4e and Fig. S5) show identical distribution throughout the sponge. MS² spectra for **16–17** are shown in Fig. S5. Aplysinopsin **16**, demethylaplysinopsin **17**, 3'-deimino-3'-oxo-aplysinopsin **18** were connected to the cluster containing tryptamine derivatives described above (Fig. 4a). In addition, aplysinopsin dimers, namely tubastrindole B (**19**) and C (**20**) were present as a separate cluster, not connected to aplysinopsins and tryptamines (Fig. S6). Tubastrindoles have been previously reported in sponge *S. cerebriformis* [38] and in stony coral *Tubastraea* sp. [39]. We report these molecules in *S. aurea* for the first time.

Mining for meroterpene class of natural products

Meroterpene natural products are comprised of an aromatic moiety upon which a terpene chain is appended followed by cyclization of the terpene chain and further tailoring of the aromatic-terpene scaffold. Marine sponges are exceptionally prolific sources of meroterpenes [40]. Meroterpenes, specifically sesquiterpene-phenol aureols, have been isolated from *S. aurea* and chemical structures established using spectroscopic and X-ray diffraction experiments [16, 17]. We explored whether the meroterpene natural product

chemical diversity was captured by our untargeted metabolomic mining approach, and if we could describe mass spectrometric fragmentation rules to help the discovery of meroterpenes from other biological sources.

An EIC for the MS¹ ion for the molecular formula C₂₁H₃₁O₂⁺, for aureol **21**, revealed the presence of the molecule in high abundance in the sponge extract (theoretical [M+H]¹⁺ = 315.232, found = 315.232, error = 0 ppm). The fragmentation spectra allowed for the annotation of characteristic MS² ions corresponding to the hydroquinone aromatic moiety (at 123.045 Da), as well as the decalin ring system (191.179 Da, Fig. 5a). MS² ions separated by 14.015 Da, corresponding to –CH₂– moieties are characteristic of terpenoid molecules. However, the presence of higher abundance aromatic moiety MS² ions, such as that at 123.045 Da help differentiate meroterpenoids from other terpenoids, such as sterols, that do not have an aromatic appendage. In addition to the aromatic moiety, the cyclized sesquiterpene derived decalin ring could also be annotated as the characteristic MS² ion observed at 191.179 Da. These two MS² ions allow us to differentiate whether tailoring modifications occurring on **21** are predicated upon the aromatic ring, or the cyclized sesquiterpene.

The node corresponding to **21** in the *S. aurea* metabolomic network is connected to several other nodes, perhaps alluding to structurally related meroterpenes to be present in the sponge, in addition to **21** (Fig. 5b). MS² spectra corresponding to each of these nodes was analyzed manually. Spectra corresponding to the most abundant precursor intensity were chosen for manual annotation. MS² spectrum corresponding to the node with *m/z* 357.243 Da, denoting the molecular formula C₂₃H₃₃O₃⁺ (theoretical [M+H]¹⁺ = 357.242, found = 357.243, error = 3 ppm) demonstrated identical MS² product ions as for **21**, with an additional MS² ion at 165.055 Da, which neatly corresponds to **22**, a previously reported O-acetylated derivative of **21** (Fig. 5b, Fig. S7) [17]. Our untargeted metabolomic approach allows for the detection of novel sesquiterpene-hydroquinone derivatives that have not been reported previously in the chemical literature. A curious case presents itself for **23**, a proposed nitrosylated derivative of **21** (theoretical [M+H]¹⁺ = 344.2220, found = 344.2229, error = 2.6 ppm). The observed MS² spectra for **23** leads us to postulate that the nitrosyl group resides on the hydroquinone ring (Fig. S7). A hydroxylated derivative of **22** (**24**, theoretical [M+H]¹⁺ = 373.237, found=373. 238, error=3 ppm), can be annotated based on the characteristic MS² product ion which demonstrates that the extra –OH is predicated upon the hydroquinone ring of **22** (Fig. S7). The oxidative hydroxylation is likely ortho- to the acetylated phenoxy, as would be expected due to the electron directing nature of the phenoxy.

Based on our prior observation that all structural derivatives within a natural product class cannot be clustered together based on MS² spectral overlap alone, we queried whether MS2LDA can reveal additional meroterpene derivatives in the *S. aurea* metabolome. A manual query of the MS2LDA motifs conserved within the cluster of nodes harboring **21–24** led to the identification of a high degree of conservation of the motif 495 (Fig. S8). MS2LDA motif 495 comprises the two fragment ions observed in the MS2 spectra of **21**, as well as the neutral loss of 192 Da which corresponds to the loss of the tetramethyl decalin ring for meroterpenes detected in *S. aurea* (Fig. 5a). Curiously, motif 495 was found

conserved in numerous other nodes in the *S. aurea* network (Fig. S8). Analyses of the spectrum for node with m/z 349.194 Da, which has a characteristic chlorine atom isotopic distribution, demonstrates a chlorinated MS² product ion which characterizes the presence of **25** (theoretical $[M+H]^{1+} = 349.193$, found = 349.194, error = 3 ppm), a previously reported chlorinated derivative of **21** [17] (Fig. 5b, Fig. S7). Similarly, the node with m/z 329.248 likely corresponds to a O-methylated derivative of **21** (**26**, theoretical $[M+H]^{1+} = 329.248$, found = 329.248, error = 0 ppm, Fig. 5b, Fig. S7). The cluster of nodes in which the node with m/z 329.28 is present also harbors nodes that were dereplicated by GNPS to correspond to fatty acids (Fig. 2, Fig. 5b). This co-clustering of meroterpenes with fatty acids is likely affected by the detection of consecutive MS² ions differing by 14.015 Da in both structural families, which corresponds to the successive loss of methylene units during MS/MS fragmentation. Thus, by a combination of clustering of nodes by MS² spectral overlap, and detection of conserved mass spectral motifs which correspond to distinct moieties in chemical structures, we demonstrate the natural product diversity harbored by marine sponges remains vastly underestimated.

Chemical diversity not captured by molecular networking

Even though the molecular networking approach captures significant chemical diversity within a sample and allows putative annotations of related molecules, we were able to identify additional related and new molecules that were not captured in the described molecular network. For example, two previously undescribed aplysinopsin analogs at m/z 243.1240 and m/z 229.1084 were not captured by the aplysinopsin molecular network, likely due to spectral filtering parameters utilized (Fig. S9). Additionally, a brominated pyrroloiminoquinolone, namely makaluvamine O, was identified by manual analysis of ESI-MS/MS data, database search and structural annotation of observed MS² ions (Fig. S10).

Utilizing a multi-pronged mass spectrometry-based strategy, we comprehensively describe for the first time the intra and inter-chemical diversity of a marine sponge. This thorough approach first highlights how complex the metabolomes of marine sponges are. Secondly, descriptions of multiple analogs of the same molecule suggest that the biosynthetic pathways for the natural synthesis of these compounds are likely promiscuous. Thirdly, detection of similar compounds produced by cyanobacterial blooms and the cyanobacteria associated with a marine sponge suggests presence of common biosynthetic pathways between the two and has interesting ecological implications. Such correlations were derived by manual search of the literature. We have deposited the current dataset and the putative annotations in the MassIVE repository. Availability of high-resolution mass spectrometry datasets and compound annotation through these repositories will enable high-throughput cross- correlations of metabolomes of diverse marine organisms including sponges, corals, marine mammals, and fishes, among others. The advances will facilitate molecular characterization of predator-prey interactions, symbiosis, bioaccumulation and other important ecological and environmental interactions, enhancing our knowledge of the marine ecosystem at the metabolic level.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

This work was supported by US National Institutes of Health (R00-ES026620 to V.A.) and Alfred P. Sloan Foundation research fellowship (FG-2018–10920 to V.A.). The authors thank Ipsita Mohanty for sponge DNA barcoding and Andrew C. Mcavoy for collecting data for standards. We thank Erich Bartels (Mote Marine Laboratory) for field assistance. We wish to acknowledge the core facilities at the Parker H. Petit Institute for Bioengineering and Bioscience at the Georgia Institute of Technology for the use of their shared equipment, services and expertise: Andrew Shaw and Aaron Lifland for help with fluorescence microscopy and Aqua T. Asberry for help with cryo-sectioning of sponge samples. We would like to thank Prof. Facundo Fernandez (Georgia Institute of Technology) for use of their MALDI matrix sprayer. We would like to thank Ming Wang and Justin J. van der Hoof, University of California San Diego, for use of MS2LDA workflow.

REFERENCES

1. Baker M: Metabolomics: from small molecules to big ideas. *Nat. Methods* 8, 117 (2011)
2. Patti GJ, Yanes O, Siuzdak G: Innovation: Metabolomics: the apogee of the omics trilogy. *Nat. Rev. Mol. Cell. Biol* 13, 263–269 (2012) [PubMed: 22436749]
3. Scheubert K, Hufsky F, Bocker S: Computational mass spectrometry for small molecules. *J. Cheminform.* 5, 12 (2013) [PubMed: 23453222]
4. Bouslimani A, Sanchez LM, Garg N, Dorrestein PC: Mass spectrometry of natural products: current, emerging and future technologies. *Nat. Prod. Rep* 31, 718–729 (2014) [PubMed: 24801551]
5. Wang M, Carver JJ, Phelan VV, Sanchez LM, Garg N, Peng Y, Nguyen DD, Watrous J, Kapono CA, Luzzatto-Knaan T, Porto C, Bouslimani A, Melnik AV, Meehan MJ, Liu WT, Crusemann M, Boudreau PD, Esquenazi E, Sandoval-Calderon M, Kersten RD, Pace LA, Quinn RA, Duncan KR, Hsu CC, Floros DJ, Gavilan RG, Kleigrew K, Northen T, Dutton RJ, Parrot D, Carlson EE, Aigle B, Michelsen CF, Jelsbak L, Sohlenkamp C, Pevzner P, Edlund A, McLean J, Piel J, Murphy BT, Gerwick L, Liaw CC, Yang YL, Humpf HU, Maansson M, Keyzers RA, Sims AC, Johnson AR, Sidebottom AM, Sedio BE, Klitgaard A, Larson CB, Boya PC, Torres-Mendoza D, Gonzalez DJ, Silva DB, Marques LM, Demarque DP, Pociute E, O'Neill EC, Briand E, Helfrich EJ, Granatosky EA, Glukhov E, Ryffel F, Houson H, Mohimani H, Kharbush JJ, Zeng Y, Vorholt JA, Kurita KL, Charusanti P, McPhail KL, Nielsen KF, Vuong L, Elfeki M, Traxler MF, Engene N, Koyama N, Vining OB, Baric R, Silva RR, Mascuch SJ, Tomasi S, Jenkins S, Macherla V, Hoffman T, Agarwal V, Williams PG, Dai J, Neupane R, Gurr J, Rodriguez AM, Lamsa A, Zhang C, Dorrestein K, Duggan BM, Almaliti J, Allard PM, Phapale P, Nothias LF, Alexandrov T, Litaudon M, Wolfender JL, Kyle JE, Metz TO, Peryea T, Nguyen DT, VanLeer D, Shinn P, Jadhav A, Muller R, Waters KM, Shi W, Liu X, Zhang L, Knight R, Jensen PR, Palsson BO, Pogliano K, Lington RG, Gutierrez M, Lopes NP, Gerwick WH, Moore BS, Dorrestein PC, Bandeira N: Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol* 34, 828–837 (2016) [PubMed: 27504778]
6. Kind T, Tsugawa H, Cajka T, Ma Y, Lai Z, Mehta SS, Wohlgemuth G, Barupal DK, Showalter MR, Arita M, Fiehn O: Identification of small molecules using accurate mass MS/MS search. *Mass Spectrom. Rev* 37, 513–532 (2018) [PubMed: 28436590]
7. Garg N, Zeng Y, Edlund A, Melnik AV, Sanchez LM, Mohimani H, Gurevich A, Miao V, Schiffler S, Lim YW, Luzzatto-Knaan T, Cai S, Rohwer F, Pevzner PA, Cichewicz RH, Alexandrov T, Dorrestein PC: Spatial Molecular Architecture of the Microbial Community of a Peltigera Lichen. *mSystems*. 1, (2016)
8. Buchberger AR, DeLaney K, Johnson J, Li L: Mass Spectrometry Imaging: A Review of Emerging Advancements and Future Insights. *Anal. Chem* 90, 240–265 (2018) [PubMed: 29155564]
9. Gerlich M, Neumann S: MetFusion: integration of compound identification strategies. *J. Mass Spectrom* 48, 291–298 (2013) [PubMed: 23494783]
10. Haug K, Salek RM, Conesa P, Hastings J, de Matos P, Rijnbeek M, Mahendraker T, Williams M, Neumann S, Rocca-Serra P, Maguire E, Gonzalez-Beltran A, Sansone SA, Griffin JL, Steinbeck C:

MetaboLights--an open-access general-purpose repository for metabolomics studies and associated meta-data. *Nucleic Acids Res.* 41, D781–786 (2013) [PubMed: 23109552]

11. Duhrkop K, Shen H, Meusel M, Rousu J, Bocker S: Searching molecular structure databases with tandem mass spectra using CSI:FingerID. *Proc. Natl. Acad. Sci. U S A* 112, 12580–12585 (2015) [PubMed: 26392543]
12. Blunt JW, Carroll AR, Copp BR, Davis RA, Keyzers RA, Prinsep MR: Marine natural products. *Nat. Prod. Rep* 35, 8–53 (2018) [PubMed: 29335692]
13. Webster NS, Thomas T: The Sponge Hologenome. *MBio.* 7, e00135–00116 (2016) [PubMed: 27103626]
14. Thomas T, Moitinho-Silva L, Lurgi M, Bjork JR, Easson C, Astudillo-Garcia C, Olson JB, Erwin PM, Lopez-Legentil S, Luter H, Chaves-Fonnegra A, Costa R, Schupp PJ, Steindler L, Erpenbeck D, Gilbert J, Knight R, Ackermann G, Victor Lopez J, Taylor MW, Thacker RW, Montoya JM, Hentschel U, Webster NS: Diversity, structure and convergent evolution of the global sponge microbiome. *Nat. Commun* 7, 11870 (2016) [PubMed: 27306690]
15. Pita L, Rix L, Slaby BM, Franke A, Hentschel U: The sponge holobiont in a changing ocean: from microbes to ecosystems. *Microbiome.* 6, 46 (2018) [PubMed: 29523192]
16. Djura P, Stierle DB, Sullivan B, Faulkner DJ: Some Metabolites of the Marine Sponges *Smenospongia-Aurea* and *Smenospongia* (=Polyfibrospongia) *Echina. J. Org. Chem* 45, 1435–1441 (1980)
17. Hu JF, Schetz JA, Kelly M, Peng JN, Ang KKH, Flotow H, Leong CY, Ng SB, Buss AD, Wilkins SP, Hamann MT: New anti-infective and human 5-HT₂ receptor binding natural and semisynthetic compounds from the Jamaican sponge *Smenospongia aurea*. *J. Nat. Prod* 65, 476–480 (2002) [PubMed: 11975483]
18. Teta R, Irollo E, Della Sala G, Pirozzi G, Mangoni A, Costantino V: Smenamides A and B, chlorinated peptide/polyketide hybrids containing a dolapyrrolidinone unit from the Caribbean sponge *Smenospongia aurea*. Evaluation of their role as leads in antitumor drug research. *Mar. Drugs* 11, 4451–4463 (2013) [PubMed: 24217287]
19. Esposito G, Teta R, Miceli R, Ceccarelli LS, Della Sala G, Camerlingo R, Irollo E, Mangoni A, Pirozzi G, Costantino V: Isolation and Assessment of the in Vitro Anti-Tumor Activity of Smenothiazole A and B, Chlorinated Thiazole-Containing Peptide/Polyketides from the Caribbean Sponge, *Smenospongia aurea*. *Mar. Drugs* 13, 444 (2015) [PubMed: 25603342]
20. Esposito G, Teta R, Della Sala G, Pawlik JR, Mangoni A, Costantino V: Isolation of Smenopyrone, a Bis- γ -Pyrone Polypropionate from the Caribbean Sponge *Smenospongia aurea*. *Mar. Drugs* 16, 285 (2018)
21. Members MSIB, Sansone S-A, Fan T, Goodacre R, Griffin JL, Hardy NW, Kaddurah-Daouk R, Kristal BS, Lindon J, Mendes P, Morrison N, Nikolau B, Robertson D, Sumner LW, Taylor C, van der Werf M, van Ommen B, Fiehn O: The Metabolomics Standards Initiative. *Nat. Biotechnol* 25, 846 (2007) [PubMed: 17687353]
22. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, Fan TW-M, Fiehn O, Goodacre R, Griffin JL, Hankemeier T, Hardy N, Harnly J, Higashi R, Kopka J, Lane AN, Lindon JC, Marriott P, Nicholls AW, Reilly MD, Thaden JJ, Viant MR: Proposed minimum reporting standards for chemical analysis. *Metabolomics.* 3, 211–221 (2007) [PubMed: 24039616]
23. Agarwal V, Blanton JM, Podell S, Taton A, Schorn MA, Busch J, Lin Z, Schmidt EW, Jensen PR, Paul VJ, Biggs JS, Golden JW, Allen EE, Moore BS: Metagenomic discovery of polybrominated diphenyl ether biosynthesis by marine sponges. *Nat. Chem. Biol* 13, 537–543 (2017) [PubMed: 28319100]
24. Edwards DJ, Marquez BL, Nogle LM, McPhail K, Goeger DE, Roberts MA, Gerwick WH: Structure and biosynthesis of the jamaicamides, new mixed polyketide-peptide neurotoxins from the marine cyanobacterium *Lyngbya majuscula*. *Chem. Biol* 11, 817–833 (2004) [PubMed: 15217615]
25. Simmons TL, Coates RC, Clark BR, Engene N, Gonzalez D, Esquenazi E, Dorrestein PC, Gerwick WH: Biosynthetic origin of natural products isolated from marine microorganism-invertebrate assemblages. *Proc. Natl. Acad. Sci. U. S. A* 105, 4587–4594 (2008) [PubMed: 18250337]

26. Unson MD, Holland ND, Faulkner DJ: A brominated secondary metabolite synthesized by the cyanobacterial symbiont of a marine sponge and accumulation of the crystalline metabolite in the sponge tissue. *Mar. Biol* 119, 1–11 (1994)
27. Flowers AE, Garson MJ, Webb RI, Dumdei EJ, Charan RD: Cellular origin of chlorinated diketopiperazines in the dictyoceratid sponge *Dysidea herbacea* (Keller). *Cell Tissue Res.* 292, 597–607 (1998) [PubMed: 9582417]
28. Schmidt EW, Nelson JT, Rasko DA, Sudek S, Eisen JA, Haygood MG, Ravel J: Patellamide A and C biosynthesis by a microcin-like pathway in *Prochloron didemni*, the cyanobacterial symbiont of *Lissoclinum patella*. *Proc. Natl. Acad. Sci. U. S. A* 102, 7315–7320 (2005) [PubMed: 15883371]
29. Platt P, Gautschi J, Thacker R, Musafija-Girt M, Crews P, Gerwick W: Identification of the cellular site of polychlorinated peptide biosynthesis in the marine sponge *Dysidea* (Lamellodysidea) *herbacea* and symbiotic cyanobacterium *Oscillatoria spongelliae* by CARD-FISH analysis. *Mar. Biol* 147, 761–774 (2005)
30. Via CW, Glukhov E, Costa S, Zimba PV, Moeller PDR, Gerwick WH, Bertin MJ: The metabolome of a cyanobacterial bloom visualized by MS/MS-based molecular networking reveals new neurotoxic smenamide analogs (C, D, and E). *Front. Chem* 6, (2018)
31. Andrianasolo EH, Gross H, Goeger D, Musafija-Girt M, McPhail K, Leal RM, Mooberry SL, Gerwick WH: Isolation of Swinholide A and Related Glycosylated Derivatives from Two Field Collections of Marine Cyanobacteria. *Org. Lett* 7, 1375–1378 (2005) [PubMed: 15787510]
32. Ueoka R, Uria AR, Reiter S, Mori T, Karbaum P, Peters EE, Helfrich EJN, Morinaka BI, Gugger M, Takeyama H, Matsunaga S, Piel J: Metabolic and evolutionary origin of actin-binding polyketides from diverse organisms. *Nat. Chem. Biol* 11, 705 (2015) [PubMed: 26236936]
33. Hertweck C: Decoding and reprogramming complex polyketide assembly lines: prospects for synthetic biology. *Trends Biochem. Sci* 40, 189–199 (2015) [PubMed: 25757401]
34. van der Hooft JJJ, Wandy J, Barrett MP, Burgess KEV, Rogers S: Topic modeling for untargeted substructure exploration in metabolomics. *Proc. Natl. Acad. Sci. U. S. A* 113, 13738 (2016) [PubMed: 27856765]
35. van der Hooft JJJ, Wandy J, Young F, Padmanabhan S, Gerasimidis K, Burgess KEV, Barrett MP, Rogers S: Unsupervised Discovery and Comparison of Structural Families Across Multiple Samples in Untargeted Metabolomics. *Anal. Chem* 89, 7569–7577 (2017) [PubMed: 28621528]
36. Wandy J, Zhu Y, van der Hooft JJJ, Daly R, Barrett MP, Rogers S: Ms2lda.org: web-based topic modelling for substructure discovery in mass spectrometry. *Bioinformatics.* 34, 317–318 (2017)
37. Bialonska D, Zjawiony JK: Aplysinsins - Marine Indole Alkaloids: Chemistry, Bioactivity and Ecological Significance. *Mar. Drugs* 7, 166 (2009) [PubMed: 19597579]
38. Dai J, Jimenez JJ, Kelly M, Barnes S, Lorenzo P, Williams P: Dictazolines A and B, bis-spiroimidazolidinones from the marine sponge *Smenospongia cerebriformis*. *J Nat Prod.* 71, 1287–1290 (2008) [PubMed: 18547116]
39. Iwagawa T, Miyazaki M, Okamura H, Nakatani M, Doe M, Takemura K: Three novel bis(indole) alkaloids from a stony coral, *Tubastraea* sp. *Tetrahedron Lett.* 44, 2533–2535 (2003)
40. Smith TE: Biogenetic Relationships of Bioactive Sponge Merotriterpenoids. *Mar. Drugs* 15, 285 (2017)

Figure 1. Anatomy of *S. aurea*. (a) Marine sponge *S. aurea*. (b) Photomicrographs of an 8 μm section of *S. aurea* embedded in gelatin, acquired by fluorescence microscopy at 100×. The cyanobacterial cells are colored in green and sponge skeleton is colored in red. The scale bar is at 20 μm.

Figure 2.

The molecular network for the metabolome of *S. aurea*. Node clusters are labeled with the structural class of natural products that they correspond to, with the chemical structures of a few representative molecules discussed in the manuscript shown. For clarity, 750 singleton nodes were removed from this illustration.

Figure 3.

Dereplication and distribution of peptide/polyketide hybrid smenamides. (a) EIC of molecule with m/z 501.252 corresponding to isomeric **1**, **2**. (b) Fluorescence microscopy-based localization of cyanobacteria by autofluorescence of chlorophyll (in purple) and of sponge tissue (in green). (c) MALDI-MSI analysis of localization of **1** and **2**, which overlaps with the autofluorescence of cyanobacteria in panel b. The ion image was generated using SCLIS Lab Pro 3D software as a discriminatory m/z with a p value of 0.05. (d) Molecular network for smenamides, connected by spectral overlap (grey edges) and by MS2LDA motif conservation (green dashes) reveal new analogs of previously known smenamides **1–4**. (e) MS² spectra of **1** and **2** and structural annotation of MS² ions. (f) MS² spectra of previously unknown molecule with m/z 543.299 (node in red color). Based on the annotation of MS²

ions, the polyketide unit of **1**, **2** contains additional $(+)(\text{CH}_2)_3$ corresponding to m/z of +42.047 in the precursor mass and MS^2 ions corresponding to polyketide unit. **(g)** MS^2 spectra of unknown molecule with m/z 487.237 (node in purple color). Based on structural annotation of the observed MS^2 ions, the loss of the methyl group from the amide nitrogen at the left periphery of **1**, **2** satisfies the observed MS^2 spectra. **(h)** MS^2 spectra of molecule with m/z 453.252 (node in violet color) reveal previously unknown desmethyl analog of **3**, **4**. Note the conservation of MS^2 ion corresponding to the polyketide chain between panels **g** and **h**, implying that the loss of the methyl group occurs from the amine nitrogen at the left periphery of **3**, **4**. The nodes colored in green are described in supplementary Fig. S2.

Figure 4.

(a) Cluster of nodes and annotations for various tryptamine and bromotryptamine derivatives, **5–10** and **11–15**, respectively, and aplysinopsins **16–18**. Note that the MS2LDA motif 480 connects the two cluster of nodes that annotate to these molecules. (b) MALDI-MSI-based distribution of non-brominated tryptamine **7**, (c) brominated tryptamine **12** and (d) brominated tryptamine **14** and (e) aplysinopsin **16** is shown.

Figure 5.

Meroterpenoid natural product chemistry from *S. aurea*. (a) MS² spectra for **21** demonstrating the characteristic fragmentation ions at 123.045 Da and 191.179 Da which correspond to the aromatic hydroquinone and the terpene decalin ring, respectively. The two MS² ions that are circled constitute the MS2LDA motif 495, as described in Fig. S8. (b) Clustering of nodes corresponding to meroterpenoid natural products in *S. aurea*. The MS² spectra for **22–26** are shown in Fig. S7. The MS2LDA motif 495 connects the cluster of nodes that correspond to **21–24**, as well as numerous other nodes in *S. aurea* metabolic network. The MS2LDA motif 495 is described in Fig. S8.