analytical chemistry

Article

pubs.acs.org/ac

Microbial Strain Prioritization Using Metabolomics Tools for the Discovery of Natural Products

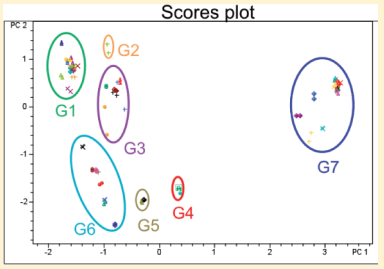
Yanpeng Hou,[†] Doug R. Braun,[†] Cole R. Michel,[†] Jonathan L. Klassen,[‡] Navid Adnani,[†] Thomas P. Wyche,[†] and Tim S. Bugni*[,†]

[†]Pharmaceutical Sciences Division, School of Pharmacy, University of Wisconsin, Madison, Wisconsin 53705, United States

[‡]Department of Bacteriology, University of Wisconsin, Madison, Wisconsin 53705, United States

Supporting Information

ABSTRACT: Natural products profoundly impact many research areas, including medicine, organic chemistry, and cell biology. However, discovery of new natural products suffers from a lack of high throughput analytical techniques capable of identifying structural novelty in the face of a high degree of chemical redundancy. Methods to select bacterial strains for drug discovery have historically been based on phenotypic qualities or genetic differences and have not been based on laboratory production of secondary metabolites. Therefore, untargeted LC/MS-based secondary metabolomics was evaluated to rapidly and efficiently analyze marine-derived bacterial natural products using LC/MS-principal component analysis (PCA). A major goal of this work was to demonstrate that LC/MS-PCA was effective for strain prioritization in a drug discovery program. As proof of concept, we evaluated LC/MS-PCA for strain selection to support drug discovery, for the discovery of unique natural products, and for rapid assessment of regulation of natural product production.



Natural products are genetically encoded small molecules that have had a profound impact on many research areas including medicine, organic chemistry, and cell biology.[1] Between 1981 and 2010, natural products or their derivatives accounted for 74% and 59% of antibacterial and anticancer new chemical entities (NCEs), respectively;[2] natural products remain an important source for drug discovery[3] and continue to inspire synthetic organic chemistry with their superlative architectural complexity.[4] Natural products have also made important contributions to cell biology due to their outstanding potency and specificity. For example, rapamycin helped elucidate the many complexities of mTOR (mammalian target of rapamycin) signaling.[5]

While bacterially produced natural products continue to be an important source for therapeutic discovery, finding novel natural products has become more difficult, and new methods are greatly needed. Genomics-based methods, such as genome mining[6] and metagenomics,[7,8] hold great promise for discovery of novel natural products and new biosynthetic pathways but, at present, are difficult to integrate with contemporary targeted high-throughput screening (HTS).[1,9,10] With respect to culture-dependent methods, immense bacterial and chemical diversity remains undiscovered.[11] Whole genome sequencing of bacteria and fungi has demonstrated that only a small fraction of the "parvome" has been discovered.[12,13] In particular, the marine environment contains a wealth of undiscovered bacteria and bacterial natural products.[14] We have focused on actinomycetes from underexplored niches, marine invertebrates such as sponges and ascidians, as a source of bacterial diversity and chemical diversity for drug discovery.

With respect to drug discovery, analytical technology development has greatly assisted with building fractionated natural product libraries that are compatible with HTS.[15,16] The success of HTS is inherently dependent on chemical diversity and a lack of chemical redundancy.[3,17] Historically, natural product extract sources were chosen either randomly or on the basis of their ecology and/or taxonomy. For bacteria chemical diversity was not determined prior to extraction, leading to redundant strains and compounds in many natural product extract libraries.[18] In order to overcome this historical weakness that led to high rates of rediscovery, methods based on the genetic potential of a microorganism to produce natural products were used.[19,20] Importantly, cultivated strains may appear identical, but produce different secondary metabolites.[21] Alternatively, strains that appear different by morphology and 16S sequencing could produce the same secondary metabolites. Therefore, we hypothesized that a chemoinformatics method based on secondary metabolite production in the lab would be more valuable and would greatly increase the value of a screening library for HTS. We embraced methods from metabolomics since they were designed, in part, to analyze large numbers of compounds without complete knowledge of the structure.

Metabolomics is the global measurement of the small molecule metabolites in a biological system and reflects the phenotype of (and is therefore complementary to) its underlying genomic, transcriptomic, and proteomic networks.

Received: October 7, 2011
Accepted: April 22, 2012
Published: April 23, 2012

dx.doi.org/10.1021/ac202623g | *Anal. Chem.* 2012, 84, 4277−4283

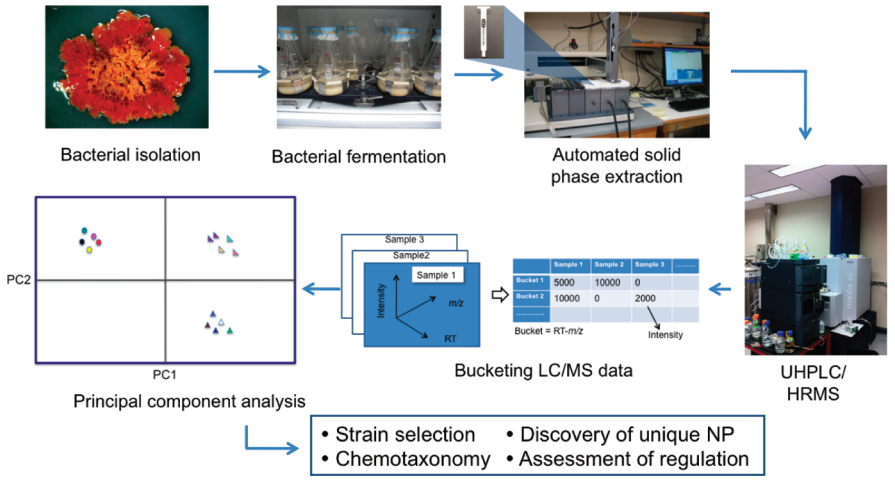


Figure 1. Flowchart for UHPLC/HRMS-based secondary metabolomics.

Metabolomics research typically implements analytical tools such as LC/MS to globally measure small molecule metabolites.[22−25] Combining principal component analysis (PCA) with LC/MS is an attractive method to provide a visual representation of variance between LC/MS profiles. We hypothesized that bacterial strains producing the same secondary metabolites would group together, whereas those producing different metabolites would be separated, thereby providing a method to select bacteria having distinct chemistries without having to identify each component of their corresponding extracts. A major distinction between the work presented here and other metabolomics studies is that we focused on secondary metabolites rather than primary metabolites.

The goal of this study was to evaluate LC/MS-PCA based secondary metabolomics to more broadly investigate secondary metabolites from marine invertebrate-associated bacteria to assist with strain selection/dereplication to support drug discovery efforts, to distinguish taxonomically identical species, to discover new natural products, and to study regulation of secondary metabolite production. Compared with conventional manual comparison among LC/MS traces, PCA significantly increased the efficiency of these studies (Figure 1).

EXPERIMENTAL SECTION

Ascidian Collection and Bacterial Cultivation. See Supporting Information.

Sample Preparation for UHPLC/HRESI-TOF-MS. *Procedure for Fermentations.* An aliquot (1.5 mL) of each fermentation was transferred to a clean microcentrifuge tube (1.7 mL) and centrifuged at 10 000 rpm for 1 min. The supernatant (1 mL) was transferred into a clean vial and placed on a Gilson GX-271 liquid handling system. The supernatant was subjected to automated SPE (Biotage: EVOLUTE ABN, 25 mg absorbent mass, 1 mL reservoir volume), washed using H_2O (1 mL) to remove media components/primary metabolites, and eluted with MeOH (1 mL) directly into an LC/MS-certified vial. While removal of primary metabolites by SPE cannot be guaranteed, we have not observed primary metabolites in our analyses to date. All compounds that have been observed as affecting the PCA have been secondary metabolites. EVOLUTE ABN was selected because it was optimized as a phase to concentrate drug-like compounds. Two biological replicates were prepared for each bacterial strain.

Procedure for Agar-Based Solid Media. For *Streptomyces* spp., we sampled directly off solid agar media. Two cores (8 mm diameter) of bacteria and agar were obtained from each plate, placed directly into MeOH (2 mL), and extracted for 30 min. The extract was transferred into a clean vial and evaporated using a SpeedVac concentrator; the residue was dissolved in MeOH (100 μL), followed by addition of H_2O (1 mL). The solution was then placed on a Gilson GX-271 liquid handling system and subjected to automated SPE as described above.

UHPLC/HRMS Analysis. LC/MS data were acquired using a Bruker MaXis ESI-Q-TOF mass spectrometer coupled with a Waters Acquity UPLC system operated by Bruker Hystar software. A Phenomenex 2.6 μm core−shell column was selected because it provides comparable results to fully porous sub-2-μm particles while generating much lower back pressure at the same velocity. Therefore, the core−shell particles can be used not only on a UHPLC system but also on a standard HPLC system.[26−28] A gradient of MeOH and H_2O (containing 0.1% formic acid) was employed with a flow rate of 0.3 mL/min on an RP C-18 column (Phenomenex Kinetex 2.6 μm, 2.1 mm × 100 mm). The gradient started from MeOH/H_2O (10%/90%), followed by a linear gradient to reach MeOH/H_2O (97%/3%) in 12 min, and held for 2 min at MeOH/H_2O (97%/3%). Full scan mass spectra (m/z 150−1550) were measured in positive ESI mode. MS data were acquired under (+)-ESI based on experience and a recent paper,[29] where 93% of the tested 719 microbial natural product and mycotoxin reference standards were detected using (+)-ESI positive mode. The mass spectrometer was operated using the following parameters: capillary, 4.5 kV; nebulizer pressure, 4.0 bar; dry gas flow, 6.0 L/min; dry gas temperature, 200 °C; scan rate, 2 Hz. Tune mix (Agilent, ESI-L low concentration) was introduced through a divert valve at the end of each chromatographic run for automatic internal calibration. Data-dependent MS/MS data were acquired using the following additional parameters: number of precursors, 3; absolute threshold, 2000; summation factor, 1.0×. Isolation width and collision energy were applied on the basis of isolation mass value and charge state against a table of isolation and fragmentation lists (see Table S1 in Supporting Information).

Data Processing and PCA. Bruker Data Analysis 4.0 was used for analysis of chromatograms. Molecular formulas were predicted using Bruker SmartFormula algorithm, which uses

dx.doi.org/10.1021/ac202623g | *Anal. Chem.* 2012, 84, 4277−4283

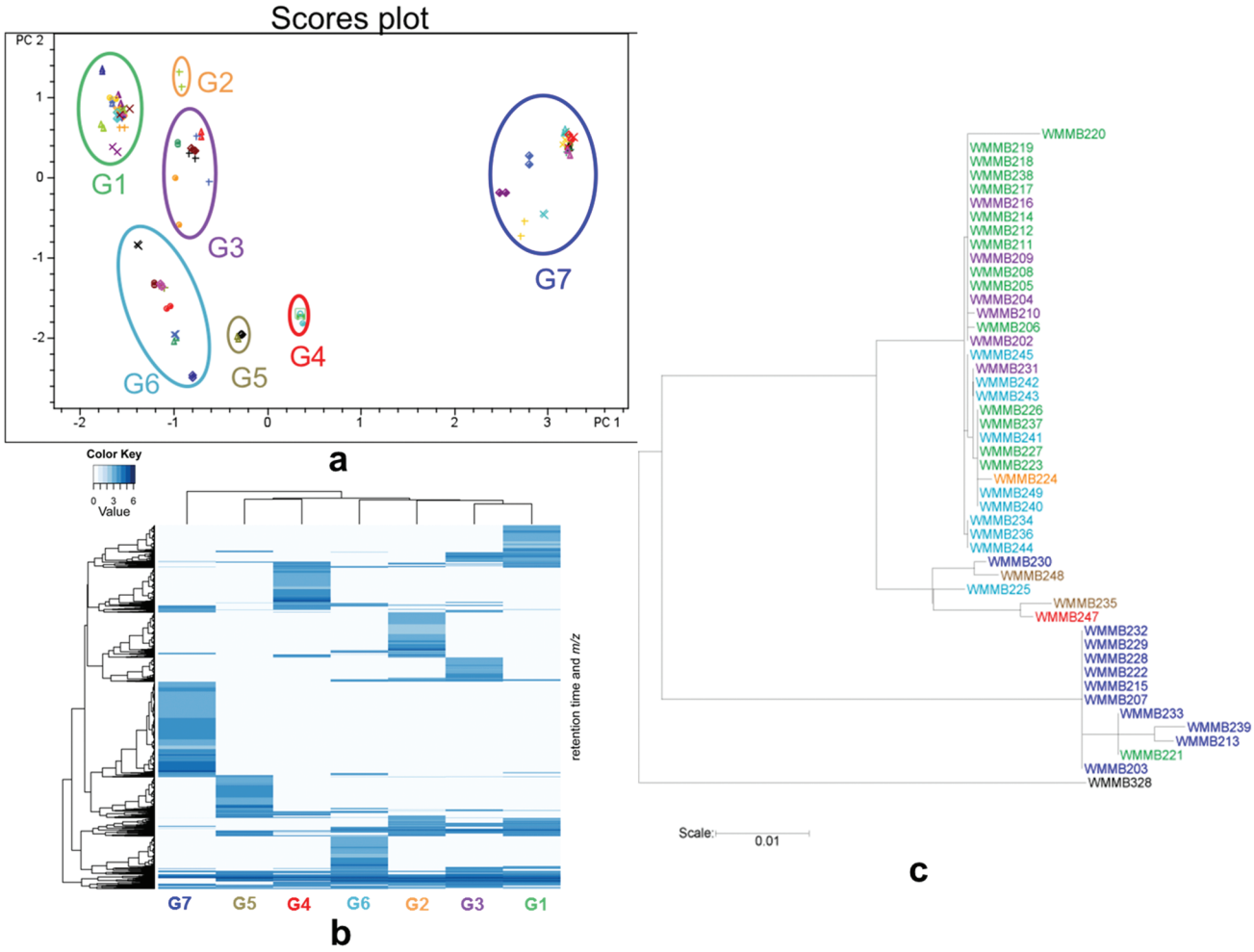


Figure 2. Strain selection by secondary metabolomics. (a) PCA scores plot (PC1 vs PC2) of 47 strains (Table S2 in Supporting Information). Strains, designated by color and shape, were clustered on the basis of natural products detected by MS. G1−G7 represent group 1−group 7, respectively, and are displayed in colored circles. (b) A heat map to display distinct metabolic profiles among the seven groups. (c) Phylogenetic tree of the 47 strains. A comparison with the groups formed in the PCA scores plot is shown. Each group of strains was designated by a colored circle in the PCA scores plot; strains within each circle were matched to the same color in the phylogenetic tree. WMMB-328 is a *Streptomyces* sp. and was used as an outgroup.

both exact mass and isotopic patterns.[30] Bucketing LC/MS data and PCA was performed using Bruker ProfileAnalysis 2.0. The detailed parameters for advanced bucketing and PCA can be found in the Additional Experimental Section in the Supporting Information. Freely available software packages, such as XCMS[31] and MZMine,[32] can also be used for raw LCMS data preprocessing to generate feature tables, which can be used for PCA in software, such as SIMCA, MATLAB, and freely available packages for R. We compared using XCMS in combination with PCA analysis with R to confirm that results were similar. Compared with open-source software, Profile Analysis 2.0 provides dynamically linked scores plot, loadings plot, and bucket statistics plot, which is convenient and efficient for data mining.

RESULTS AND DISCUSSION

Data Preprocessing and PCA. In Profile Analysis 2.0, a molecule feature finding algorithm was employed to extract compound signals from background noise.[31] Advanced bucketing applied an internal hierarchical clustering algorithm to construct buckets with given sizes of retention time (RT) and m/z around picked compounds followed by generation of a bucket table including RT−m/z pairs (buckets) and intensities. Importantly, the size of the retention time window must be greater than the possible retention time variation between runs. Compounds with different charge states were deconvoluted into a single RT−m/z pair (bucket).[33] Both line and profile spectra were subjected to PCA and showed identical results. Thus, line spectra were selected, which significantly reduced calculation time. The bucket table was then subjected to PCA, and the scores plots and the corresponding loadings plots were analyzed using Profile Analysis 2.0. Profile Analysis greatly assisted with analyzing PCA results. For example, the scores plot presents a view of the variance in the data, and the further groups separate, the more different the secondary metabolites. The loadings plot is geometrically related to the scores plot and described the variance observed in the scores plot. Analysis of any point in the loadings plot through selection with a mouse provides a corresponding bucket statistics plot that shows the intensity of the selected compound in all analyses that were used for PCA. Additionally, a dynamically linked table highlights the compound selected in the loadings plot. The table contains the m/z−RT information, intensity, and the accurate mass, which allows a quick analysis of putative

dx.doi.org/10.1021/ac202623g | *Anal. Chem.* 2012, 84, 4277−4283

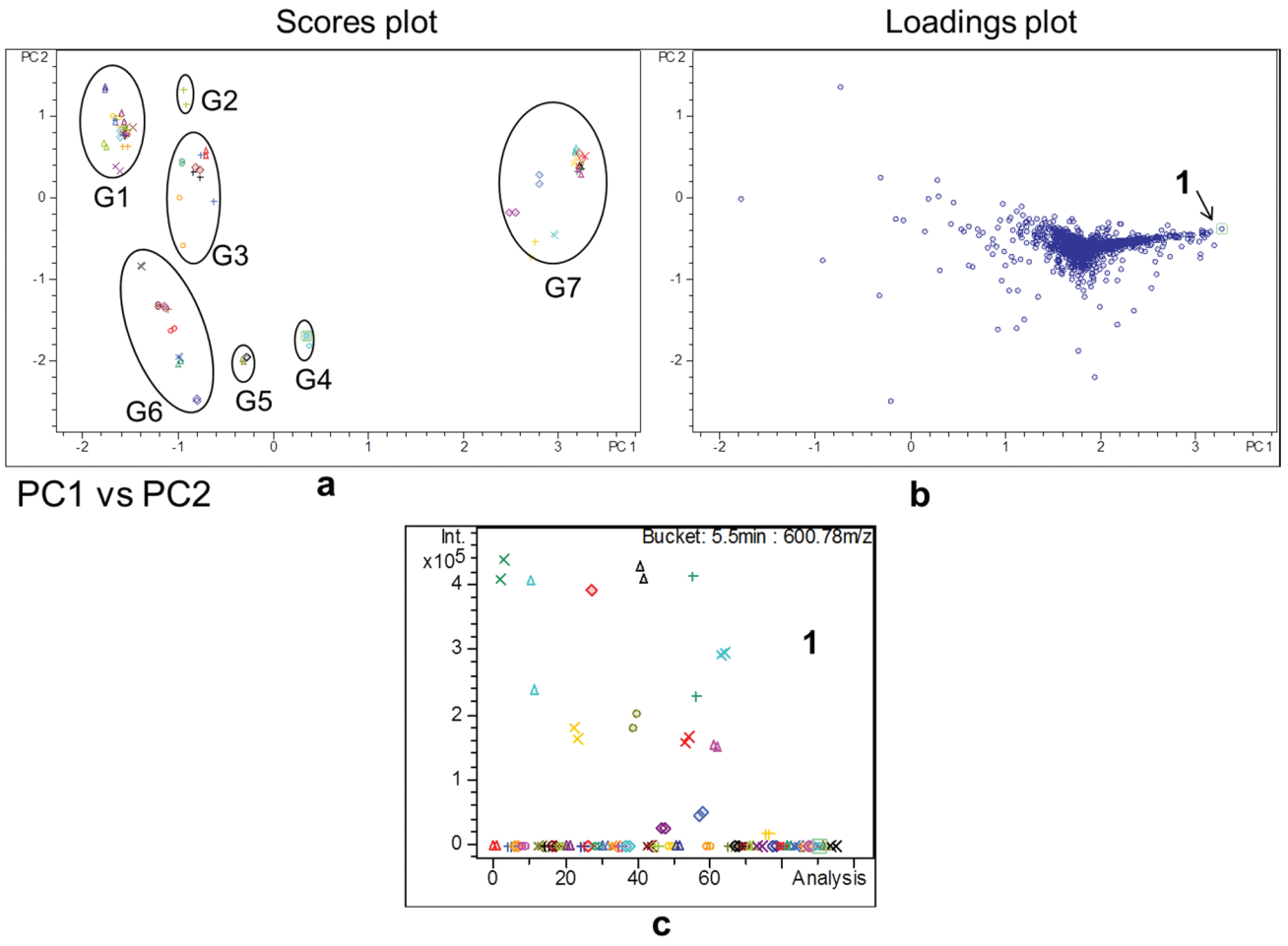


Figure 3. Discovery of natural products unique to the group 7. (a) PCA scores plot (PC1 vs PC2) of the 47 stains (Table S1 in Supporting Information). (b) PCA loadings plot. The PCA loadings plot is geometrically related to the scores plot and can be used to identify the compounds responsible for group patterns in the PCA scores plot. Compound **1** (RT 5.5 min m/z 600.7804, $[M + 2H]^{2+}$) is one of the major contributors to group 7. (c) The bucket statistic plot for **1** showing normalized intensities of **1** in the 47 strains. Compound **1** was only detected in group 7.

molecular formulas using SmartFormula. The ability to interactively analyze secondary metabolites greatly assisted these studies and will be highlighted below.

Strain Selection by Secondary Metabolomics. Without careful construction of a screening library, large screening campaigns lead to identifying large numbers of known compounds. Therefore, an effective strain selection approach is required to maximize chemical diversity and minimize redundancy for discovery of therapeutic leads.[34,35] As a proof of concept for strain selection, 47 strains (32 *Verrucosispora* spp., 5 *Micromonospora* spp., and 10 *Nocardia* spp. cultivated from tropical ascidians, *Trididemnum orbiculatum*, *Didemnum psammathode*, and *Ecteinascidia turbinata*) were analyzed by LC/MS and PCA. (See Table S1 in Supporting Information for more details on the strains.) The LC/MS data used for this research has been made publicly accessible.[36] The *Verrucosispora* spp. and *Micromonospora* spp. appeared identical on isolation plates and provided a real world example of distinguishing strains cultivated from marine invertebrates. In some cases these bacteria were identical by morphology and 16S rDNA sequences. After LC/MS and processing, a total of 26 027 buckets and 74 principal components were generated and explained 98% of the variation in the data set. Analysis of the resulting scores plot (PC1 versus PC2) showed identifiable groups. As stated in the Experimental Section, the loadings plot was dynamically linked to the scores plot to facilitate identification of compounds that caused variance. After analysis of the scores plot, 7 clustered groups were identified (Figure 2a). To address the question of whether selection by PCA resulted in low chemical redundancy, we selected one representative strain from each of the 7 clustered groups and generated a heat map (Figure 2b). The heat map of metabolite profiles showed little overlap and complementary chemical space based on detected ions. Overall, this would greatly reduce chemical redundancy in a screening library when strains were selected on the basis of PCA.

We also compared and contrasted the use of secondary metabolomics with phylogenetic trees generated from 16S rDNA sequences (Figure 2c). In some cases, the grouping observed in the PCA scores plot (Figure 2a) paralleled those observed in the phylogenetic tree. For example, the G1 and G3 groups were dominated by *Verrucosispora* spp. On the other hand, *Verrucosispora* spp. found in both G1 and G3 had identical 16S sequences indicating that PCA provided finer resolution to distinguish strains compared to morphology and 16S comparisons. The results indicated that strains with nearly identical 16S gene sequences do not necessarily produce the same natural products and supported our approach for strain selection based on natural product production. These results are consistent with other published results where strains that

dx.doi.org/10.1021/ac202623g | *Anal. Chem.* 2012, 84, 4277–4283

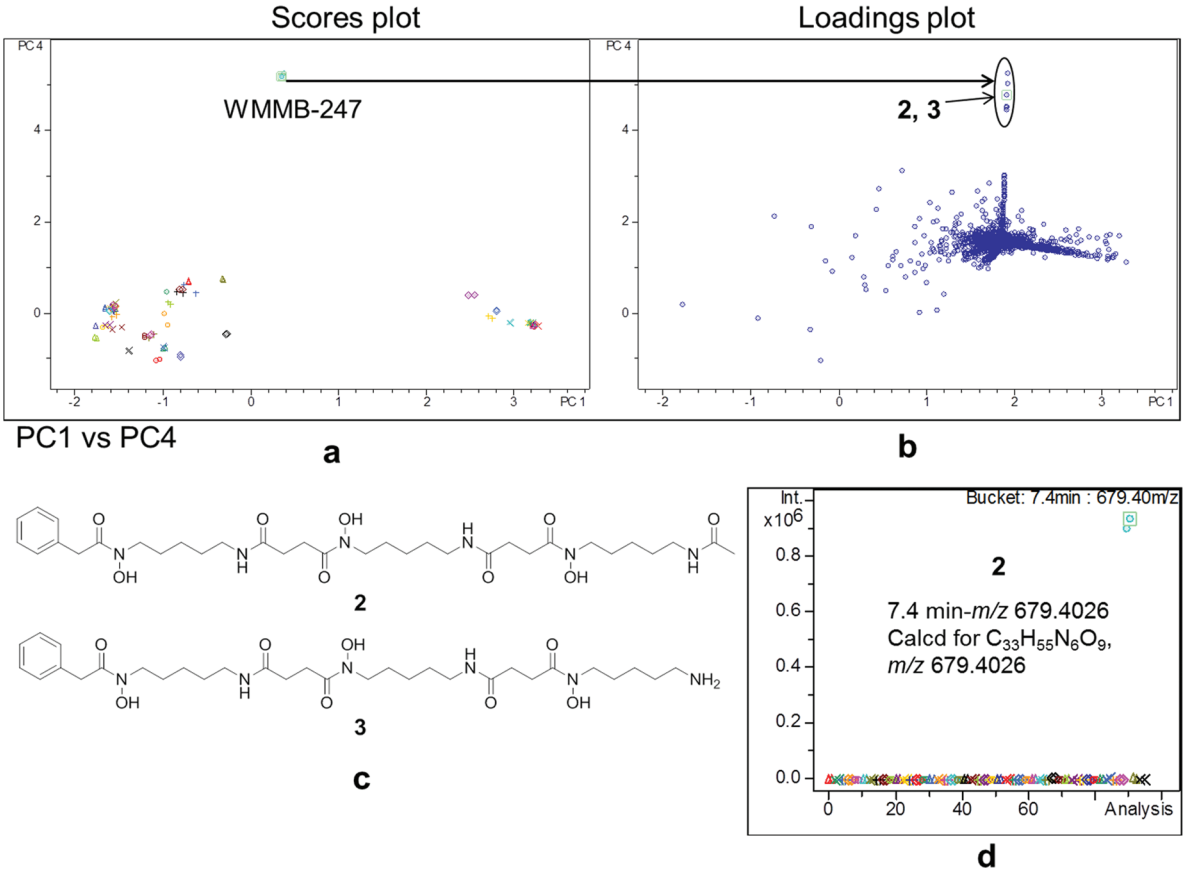


Figure 4. Discovery of natural products unique to strain WMMB-247. (a) PCA scores plot (PC1 vs PC4). The PC planes were adjusted to separate WMMB-247. (b) PCA loadings plot. The loadings plot shows compounds (**2**–**3**) that were unique to WMMB-247. (c) Structures of **2** and **3**. (d) The bucket statistic plot of **2** (RT 7.4 min *m/z* 679.4026, $[M + H]^+$). As shown, compound **2** was only detected in WMMB-247.

appear nearly identical by morphology and 16S rDNA sequences (>99% identity) can produce different natural products.[21] LC/MS-PCA proved useful for distinguishing strains on the basis of laboratory production of secondary metabolites. With respect to strain dereplication, LC/MS-PCA could provide a selection criterion prior to 16S sequencing to reduce the number of bacteria that need to be sequenced in a drug discovery program.

A limiting factor to using PCA was that supporting approaches were necessary to increase the scale. For example, PCA would not be useful for analyzing 1000 strains at a time. We have found that analysis of between 20 and 50 strains was practical. For strain selection from a cultivated collection, we use a combination of gross morphology and source organism to classify groups to be analyzed by PCA. Additionally, analysis can be performed on overlapping groups such that each successive analysis contains some data from the previous analysis. Combined with gross morphology, we have now successfully analyzed over 500 strains with great success.

Discovery of Natural Products Unique to Specific Strains. Previous work showed that natural products produced by *Myxococcus* spp. could be rapidly mapped to producing strains and novel natural products could be identified using secondary metabolomics.[33,37] Therefore, we evaluated the process for marine invertebrate associated bacteria. The loadings plot was used to identify compounds that caused a clustered group to separate. Again, Profile Analysis allowed a point and click in the loadings plot to observe the presence of the compound in all analyzed strains. Additionally, an accurate mass was obtained through the dynamically linked table. For selected putative new natural products, isolation and structure elucidation of those compounds was undertaken as described in Supporting Information. In addition to conventional approaches, an LC-MS-SPE-NMR system[38] was used in conjunction with $^1H\{^{13}C,^{15}N\}$ 1.7 mm NMR cryogenic probe to facilitate rapid structure elucidation.

As an example, compound **1** (RT 5.5 min-*m/z* 600.7804, $[M + 2H]^{2+}$) (Figure 3b) was a significant contributor to group 7 (Figure 3a) because **1** and group 7 were located in the same quadrant in the loadings plot and the scores plot, respectively. From the bucket statistics plot (Figure 3c), compound **1** was only detected in the strains from group 7, which consisted of *Nocardia* spp. along with two *Micromonospora* spp. After rapid structure dereplication based on high resolution MS data of **1** (Figure 3), compound **1** was a new natural product that has been targeted for structure determination but was beyond the scope of this work.

Identifying compounds responsible for groups near the center of the scores plot, such as group 4 (Strain WMMB-247, Figure 3a), was not straightforward. However, observing different PC planes could separate these groups. As shown in Figure 4a, WMMB-247 was separated in the PC1 vs PC4 plane. Compounds unique to WMMB-247 were identified in the corresponding loadings plot (Figure 4b). The major compounds unique to WMMB-247 were elucidated as two rare phenylacetyl-desferrioxamines (**2**–**3**) (Figure 4c). The structures of **2** and **3** were confirmed using MS/MS and NMR (Figures S3 and S4). The bucket statistics plot for compound **2**

dx.doi.org/10.1021/ac202623g | *Anal. Chem.* 2012, 84, 4277–4283

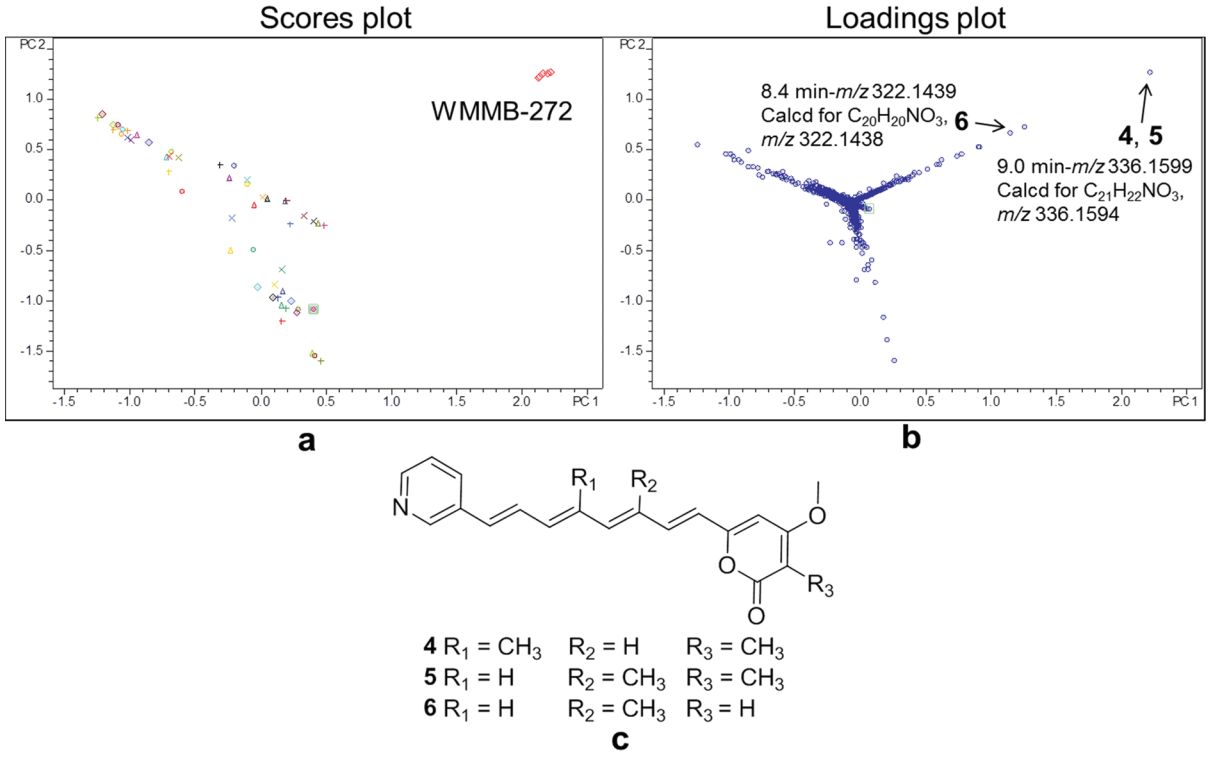


Figure 5. Discovery of polyenepyrones from a marine-derived *Streptomyces* sp. (WMMB-272). (a) PCA scores plot indicated separation of strain WMMB-272 from a collection of 53 marine-derived *Streptomyces* spp. (b) PCA loadings plot displayed compounds responsible for the separation. Compounds (**4**–**6**) were isolated after scale-up fermentation. (c) Structures of compounds (**4**–**6**) were determined by NMR and MS. Compounds **4**–**6** belong to a rare class of polyenepyrones.

(RT 7.4 min *m*/*z* 679.4026, $[M + H]^+$) indicated that the compound was only detected in strain WMMB-247 (Figure 4d). Desferrioxamines **2** and **3** were recently reported and promoted the growth of previously uncultivable bacteria when released from neighboring strains.[39]

Having success with strains grown in liquid, we tested LC/MS-PCA to discover novel natural products from marine-derived *Streptomyces* spp. grown on solid media. Fifty-three marine-derived *Streptomyces* spp. cultivated from tropical ascidians, *Tridemnum orbiculatum* and *Didemnum psammathode*, were analyzed using LC/MS-PCA (Figure 5). In the scores plot, strain WMMB-272 was clearly distinct from the other strains, which indicated unique natural products. Molecular formulas were identified through interactive analysis of the loadings plot, and searches in databases indicated that the compounds responsible for the variance observed in PCA were new natural products. Subsequently, we purified the compounds, elucidated their structures, and determined that they were novel natural products (see 1D and 2D NMR assignment in Table S2, Figure S5–S8). Although **4** and **5** were published by Fenical and co-workers[40] during the preparation of this paper, we independently discovered them using PCA. Compound **6** was a new natural product, and this is the first report of the structure. In addition, we identified 20 out of the 53 *Streptomyces* spp. isolates produced a number of unique and putative new natural products. The isolation and structure elucidation of these compounds will be reported in due course. Compared to traditional approaches, this method represented a remarkably high discovery rate from *Streptomyces* spp.

Assessment of Microbial Natural Product Production by Secondary Metabolomics. Secondary metabolomics was also evaluated as a method to assess natural product production. We hypothesized that LC/MS-PCA would be a valuable tool to investigate regulation of biosynthesis. Therefore, we selected a system where regulation was well understood, desferrioxamine (iron siderophores) production and regulation. Bacteria have developed sophisticated mechanisms to regulate iron metabolism, which have been extensively investigated.[41] A marine-derived *Micromonospora* sp. (WMMB-224) was selected because it produced a diverse group of desferrioxamine. Since desferrioxamine biosynthesis was shown to be highly responsive to iron concentrations, the effects of iron in three growth media (ASW-A, ASW-K, and DI-A) (see recipe in Table S3 in Supporting Information) on desferrioxamine production was investigated. Using LC/MS-PCA, a total of 30 conditions (Table S4 in Supporting Information) were simultaneously analyzed. The grouping pattern in the PCA scores plot was primarily based on the concentrations of iron added and not on the type of iron added or the three media types (Figures S1a and S2a,b in Supporting Information). After examining the loadings plot (Figure S1b), a major difference between group 1 and group 3 was the absence and presence of compound **7** (putative new natural product RT 10.5 min *m*/*z* 431.2767, $[M + H]^+$), respectively. As predicted, the other major difference between group 1 and group 3 was primarily due to the presence and absence of desferrioxamines, respectively. However, desferrioxamine B (**8**) and desferrioxamine D1 (**9**) in group 1 were observed as Fe^{3+}-chelated forms in group 3 (Figure S1b). This proof of concept suggests that LC/MS-PCA would be a valuable tool for analyzing regulation of biosynthesis.

In conclusion, LC/MS-PCA ("secondary metabolomics") was effective for selecting strains to yield the most chemically diverse and novel natural products for drug discovery.

dx.doi.org/10.1021/ac202623g | *Anal. Chem.* 2012, 84, 4277–4283

Importantly, we showed that chemistry did not always correlate with 16S phylogenetic trees, and those differences could easily be observed using PCA. Overall, analysis of secondary metabolites provided finer resolution of strain differentiation than 16S. Additionally, heat maps indicated that little chemical redundancy was observed when selecting representative strains for drug discovery. In the end, we have validated LC/MS-PCA as a rapid method to select strains on the basis of laboratory production of secondary metabolites. In addition to strain selection, secondary metabolomics greatly assisted with discovery of novel natural products and will be a useful tool to study regulation of natural product biosynthesis. Overall, these methods could enable a wide range of studies in the field of microbial natural products by drastically decreasing analysis time compared to manual comparison of LC/MS chromatograms.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Additional Experimental Section, tables, and figures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: tbugni@pharmacy.wisc.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by funding from School of Pharmacy, the Graduate School, and College of Agriculture and Life Sciences at the University of Wisconsin-Madison. This work was also funded by the NIH, NIGMS Grant R01 GM092009, NIH Grant RC4GM096347, and the Natural Sciences and Engineering Research Council of Canada Postdoctoral Research Fellowship. We would like to thank the Analytical Instrumentation Center at the University of Wisconsin—Madison for the facilities to acquire spectroscopic data, especially NMR and MS. This study made use of the National Magnetic Resonance Facility at Madison (NMRFAM), which is supported by NIH Grants P41RR02301 (BRTP/NCRR) and P41GM66326 (NIGMS). Additional equipment was purchased with funds from the University of Wisconsin, the NIH (RR02781, RR08438), the NSF (DMB-8415048, OIA-9977486, BIR-9214394), and the USDA. We would like to thank Dr. Eldon Ulrich for making part of our LC/MS data publicly accessible through the Biological Magnetic Resonance Data Bank, Dr. Aiko Barsch and Dr. Sandy Yates (Bruker Daltonics) for helpful discussions about metabolomics, and Dr. Kaifeng Hu (NMRFAM) for his assistance with NMR data acquisition. We would also like to thank the reviewers of this manuscript for their invaluable suggestions and comments.

■ REFERENCES

(1) Walsh, C. T.; Fischbach, M. A. *J. Am. Chem. Soc.* **2010**, *132*, 2469−93.
(2) Newman, D. J.; Cragg, G. M. *J. Nat. Prod.* **2012**, *75*, 311−335.
(3) Li, J.; Vederas, J. C. *Science* **2009**, *325*, 161−5.
(4) Wender, P. A.; Miller, B. L. *Nature* **2009**, *460*, 197−201.
(5) Wullschleger, S.; Loewith, R.; Hall, M. N. *Cell* **2006**, *124*, 471−84.
(6) Zerikly, M.; Challis, G. L. *ChemBioChem* **2009**, *10*, 625−33.
(7) Brady, S. F.; Simmons, L.; Kim, J. H.; Schmidt, E. W. *Nat. Prod. Rep.* **2009**, *26*, 1488−503.
(8) Li, X.; Guo, J.; Dai, S.; Ouyang, Y.; Wu, H.; Sun, W.; Wang, G. *Curr. Top. Med. Chem.* **2009**, *9*, 1525−35.
(9) Amy, L. L.; Bradley, S. M. *Nat. Prod. Rep.* **2011**, *28*, 411−28.
(10) Walsh, C. T. *Acc. Chem. Res.* **2008**, *41*, 4−10.
(11) Bérdy, J. *J. Antibiot.* **2005**, *58*, 1−26.
(12) Davies, J. *Curr. Opin. Chem. Biol.* **2011**, *15*, 5−10.
(13) Davies, J.; Ryan, K. S. *ACS Chem. Biol.* **2012**, *7*, 252−9.
(14) Fenical, W.; Jensen, P. R. *Nat. Chem. Biol.* **2006**, *2*, 666−73.
(15) Bugni, T. S.; Richards, B.; Bhoite, L.; Cimbora, D.; Harper, M. K.; Ireland, C. M. *J. Nat. Prod.* **2008**, *71*, 1095−8.
(16) Johnson, T. A.; Sohn, J.; Inman, W. D.; Estee, S. A.; Loveridge, S. T.; Vervoort, H. C.; Tenney, K.; Liu, J.; Ang, K. K.-H.; Ratnam, J.; Bray, W. M.; Gassner, N. C.; Shen, Y. Y.; Lokey, R. S.; McKerrow, J. H.; Boundy-Mills, K.; Nukanto, A.; Kanti, A.; Julistiono, H.; Kardono, L. B. S.; Bjeldanes, L. F.; Crews, P. *J. Nat. Prod.* **2011**, *74*, 2545−55.
(17) Harvey, A. L. *Curr. Opin. Chem. Biol.* **2007**, *11*, 480−484.
(18) Handelsman, J.; Rondon, M. R.; Brady, S.; Clardy, J.; Goodman, R. M. *Chem. Biol.* **1998**, *5*, R245−9.
(19) Ritacco, F. V.; Haltli, B.; Janso, J. E.; Greenstein, M.; Bernan, V. S. *J. Ind. Microbiol. Biotechnol.* **2003**, *30*, 472−9.
(20) Ayuso, A.; Clark, D.; Gonzalez, I.; Salazar, O.; Anderson, A.; Genilloud, O. *Appl. Microbiol. Biotechnol.* **2005**, *67*, 795−806.
(21) Jensen, P. R.; Williams, P. G.; Oh, D. C.; Zeigler, L.; Fenical, W. *Appl. Environ. Microbiol.* **2007**, *73*, 1146−52.
(22) Raamsdonk, L. M.; Teusink, B.; Broadhurst, D.; Zhang, N.; Hayes, A.; Walsh, M. C.; Berden, J. A.; Brindle, K. M.; Kell, D. B.; Rowland, J. J. *Nat. Biotechnol.* **2001**, *19*, 45−50.
(23) Buchholz, A.; Hurlebaus, J.; Wandrey, C.; Takors, R. *Biomol. Eng.* **2002**, *19*, 5905.
(24) Oliver, F. *Plant Mol. Biol.* **2002**, *48*, 155−71.
(25) Kersten, R. D.; Dorrestein, P. C. *ACS Chem. Biol.* **2009**, *4*, 599−601.
(26) Abbasi, E.; Layne, J.; Behr, H.; Campbell, J.; Countryman, S. *Drug Discovery Dev.* **2009**, *10*, 24−6.
(27) Koerner, P.; Mathews, T. *LC/GC European Applied Notebook*; March 2, 2010; http://chromatographyonline.findanalytichem.com/.
(28) Gritti, F.; Leonardis, I.; Shock, D.; Stevenson, P.; Shalliker, A.; Guiochon, G. *J. Chromatogr., A* **2010**, *1217*, 1589−603.
(29) Nielsen, K. F.; Månsson, M.; Rank, C.; Frisvad, J. C.; Larsen, T. O. *J. Nat. Prod.* **2011**, *74*, 2338−48.
(30) Stacey, C.; Goetz, S.; Baessmann, C. *Technical Note TN-23*; Bruker Daltonik GmbH, 2008, www.bdal.com.
(31) Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. *Anal. Chem.* **2006**, *78*, 779−87.
(32) Katajamaa, M.; Miettinen, J.; Oresic, M. *Bioinformatics* **2006**, *22*, 634−6.
(33) Krug, D.; Zurek, G.; Schneider, B.; Garcia, R.; Müller, R. *Anal. Chim. Acta* **2008**, *624*, 97−106.
(34) Genilloud, O.; González, I.; Salazar, O.; Martín, J.; Tormo, R. J. *J. Ind. Microbiol. Biotechnol.* **2011**, *38*, 375−89.
(35) Liu, X.; Ashforth, E.; Ren, B.; Song, F.; Dai, H.; Liu, M.; Wang, J.; Xie, Q.; Zhang, L. *J. Antibiot.* **2010**, *63*, 415−22.
(36) The data can be freely accessed through the following link: http://www.bmrb.wisc.edu/secondary_metabolomics/marine_bacteria_profiles.html.
(37) Krug, D.; Zurek, G.; Revermann, O.; Vos, M.; Velicer, G. J.; Müller, R. *Appl. Environ. Microbiol.* **2008**, *74*, 3058−68.
(38) Seger, C.; Godejohann, M.; Tseng, L.; Spraul, M.; Girtler, A.; Sturm, S.; Stuppner, H. *Anal. Chem.* **2005**, *77*, 878−85.
(39) D'Onofrio, A.; Crawford, J. M.; Stewart, E. J.; Witt, K.; Gavrish, E.; Epstein, S.; Clardy, J.; Lewis, K. *Chem. Biol.* **2010**, *17*, 254−64.
(40) Fukuda, T.; Miller, E. D.; Clark, B. R.; Alnauman, A.; Murphy, C. D.; Jensen, P. R.; Fenical, W. *J. Nat. Prod.* **2011**, *74*, 1773−8.
(41) Massé, E.; Salvail, H.; Desnoyers, G.; Arguin, M. *Curr. Opin. Microbiol.* **2007**, *10*, 140−5.