Chemistry & Biology
Resource

"Function-First" Lead Discovery: Mode of Action Profiling of Natural Product Libraries Using Image-Based Screening

Christopher J. Schulze,[1] Walter M. Bray,[2] Marcos H. Woerhmann,[3] Joshua Stuart,[3] R. Scott Lokey,[2] and Roger G. Linington[1,*]
[1]Department of Chemistry and Biochemistry
[2]Chemical Screening Center
[3]Department of Biomolecular Engineering
University of California Santa Cruz, Santa Cruz, CA 95064, USA
*Correspondence: rliningt@ucsc.edu
http://dx.doi.org/10.1016/j.chembiol.2012.12.007

SUMMARY

Cytological profiling is a high-content image-based screening technology that provides insight into the mode of action (MOA) for test compounds by directly measuring hundreds of phenotypic cellular features. We have extended this recently reported technology to the mechanistic characterization of unknown natural products libraries for the direct prediction of compound MOAs at the primary screening stage. By analyzing a training set of commercial compounds of known mechanism and comparing these profiles to those obtained from natural product library members, we have successfully annotated extracts based on MOA, dereplicated known compounds based on biological similarity to the training set, and identified and predicted the MOA of a unique family of iron siderophores. Coupled with traditional analytical techniques, cytological profiling provides an avenue for the creation of "function-first" approaches to natural products discovery.

INTRODUCTION

Historically, natural products have been a rich source of lead compounds for drug development against a wide array of biological targets. Currently, over half of all Food and Drug Administration (FDA)-approved drugs are derived either directly or indirectly from natural sources. However, the role of natural products in the pharmaceutical industry has been in steady decline over the past two decades, due in large part to the shift toward high-throughput screening (HTS) of synthetic libraries because of the difficulties associated with deconvoluting biologically active components from complex natural product extracts (Koehn and Carter, 2005). One area where natural products continue to be of high value is in the discovery of antitumor drugs. Since 1940, almost 75% of approved small molecules for the treatment of cancer have been either natural products, semisynthetic derivatives of natural product scaffolds, or synthetic compounds inspired by natural products pharmacophores (Newman and Cragg, 2012). The field of marine natural products has found recent success in this area, with 15 marine natural product-inspired compounds either approved by the FDA or in clinical trials for the treatment of cancer (Gerwick and Moore, 2012). These data suggest that as the field continues to develop, marine-derived natural products will be a growing resource for the discovery of novel anticancer agents.

Broadly speaking, all drug discovery screening campaigns can be categorized into one of two approaches: target-based screening against specific molecular targets, or whole cell-based screening to reach defined biological end points (Koehn and Carter, 2005). Although natural products occupy a diverse array of chemical space compared to synthetic combinatorial libraries (Lachance et al., 2012), natural product extracts are not particularly well suited to target-based HTS. They pose many challenges to HTS platforms, the most significant of which is that extracts contain mixtures of compounds of unknown concentration, which can significantly complicate biological readouts (Harvey, 2007). In addition, hits from target-based HTS often fail to elicit the desired biological response in whole-cell systems due to poor pharmacokinetic properties. By contrast, efficacy can be readily confirmed in whole-cell screens, but these provide no information about mode of action (MOA), making hit prioritization and optimization challenging (Swinney and Anthony, 2011).

Several new advances in screening technologies have been developed with the goal of unifying these two approaches. One such technique, cytological profiling (CP), uses automated fluorescence microscopy to interrogate the physiological and phenotypic effects of test compounds on cellular development (Perlman et al., 2004; Mitchison, 2005; Young et al., 2008; Sutherland et al., 2011). By using biologically relevant fluorescent markers, MOA predictions can be made based on the phenotypic "fingerprint" of the cells. In addition, it is possible to cluster compounds by MOA based on their cytological profiles, given that compounds with similar modes of action elicit similar physiological responses (Young et al., 2008). A conceptually similar technique, the Connectivity Map, has been developed by the Broad Institute to predict the MOAs for small molecules based on transcriptional analysis, rather than fluorescent markers (Lamb et al., 2006).

We have applied a modified cytological profiling platform to our library of marine natural product prefractions with the objectives of (1) annotating cytotoxic compounds in the library based

on their MOAs, (2) dereplicating prefractions based on comparison of MOAs to a library of known compounds, and (3) probing prefractions with unique CP fingerprints for the presence of novel antitumor compounds. Although this technology has previously been used to profile pure natural products (Young et al., 2008) and to probe extracts for specific biological activities (Sumiya et al., 2011), this study shows that cytological profiling can be implemented in screening large libraries of natural product mixtures of unknown chemical composition. Overall, the results of this study show that natural product prefractions can be accurately clustered based on mode of action, and that prefractions containing the same compounds both cluster together, and cluster with matching reference compounds from the training set. We have shown that prefractions containing different compounds that act by similar mechanisms share diagnostic and readily identifiable phenotypic fingerprints, and that this platform can therefore be used to characterize screening library members by MOA. We can also identify clusters of prefractions that do not share significant similarity to any of our reference compounds, which can be selected as high priority for future study. Finally, this technology has been used to discover and predict the mode of action of a family of previously undescribed siderophores, microferrioxamines A–D (1–4).

RESULTS

Cytological Profiling of a Known Compound Library with Annotated Modes of Action

In order to verify the efficacy of our cytological profiling platform, we first screened a library of 480 small molecules from the Harvard Institute of Chemistry and Cell Biology (ICCB) collection. The library contains compounds that affect a variety of known cellular targets, including the cytoskeleton, protein synthesis, kinases, nuclear hormone receptors, ion channels, and G protein coupled receptors, among others. Automated image analysis generated a score for 250 cellular features (Table S1 available online) corresponding to shape, intensity, and size for both individual cells and well averages, and was performed as previously described (Young et al., 2008). Taken together, these scores generate a cytological profile for each compound screened (Supplemental Experimental Procedures). These data identified 21 discrete mechanistic/target classes from the 480-member library that displayed intraclass Pearson correlations that were significantly higher than extra-class pairwise similarities ($p < 0.0001$). These data validated the ability of our cytological profiling platform to cluster pure compounds by mechanistic class.

Cytological Profiling of Marine Natural Product Prefractions

To evaluate the suitability of cytological profiling for the characterization of natural products mixtures, a library of 624 marine-derived bacterial prefractions were evaluated using the approach described above. To prepare this library, six prefractions were generated from each crude extract using C_{18} solid phase extraction cartridges to give screening materials of reduced complexity, containing an estimated 2–20 compounds per prefraction. Following screening and imaging, feature scores were calculated by divergence from internal control wells in each plate. The data acquired from the ICCB library and the 624 natural product prefractions were combined and clustered using Cluster 3.0. Clustered data were visualized as a dendrogram using Java TreeView (Figure 1A).

From this dendrogram, seven prominent clusters were selected for further investigation based on the internal similarity of their cytological profiles by Pearson correlation (Figure 1A). In most instances, clusters contained both natural product prefractions and known compounds from the ICCB library, suggesting that active constituents from the natural product prefractions should target similar cellular structures or pathways to the ICCB compounds in the same cluster. To test this hypothesis, select prefractions from each cluster were analyzed by liquid chromatography-tandem mass spectrometry (LC-MS) and nuclear magnetic resonance (NMR) spectroscopy and the structures of active constituents determined as outlined below.

Prediction of Compound Modes of Action Using Cytological Profiling

One of the most readily identifiable clusters is Cluster A (Figure 1B). This cluster has a distinctive phenotype that results from DNA modulation and/or arrest of the cell cycle. The set of fluorescent stains employed in this study is particularly well suited to detecting effects on cell cycle progression. These probes include 5-ethynyl-2′-deoxyuridine (EdU), which is incorporated into the nascent DNA strand in S phase and subsequently visualized by copper-mediated Click coupling to rhodamine azide, and a mouse antiphospho-histone H3 (pHH3) antibody visualized with anti-mouse Cy5-conjugated rabbit antibody to identify cells actively undergoing mitosis. Strongly negative deviations from the control wells in Cluster A for both EdU and pHH3 resulted in the formation of a cluster of compounds that cause cell-cycle arrest. In addition, this cluster showed positive deviations from control wells in nuclear shape and size features, which is attributable to the lack of mitotic cells with condensed chromatin in these wells. Many of the ICCB compounds that are expected to affect DNA synthesis or transcription, such as actinomycin D, aphidicolin, mitomycin C, and camptothecin are tightly correlated in Cluster A.

Preliminary analysis of the natural product prefractions in this cluster sought to determine if they contained compounds that were consistent with this phenotype, but not represented in the ICCB library. An automated fraction collection strategy, which uses LC-MS to simultaneously separate and collect the eluent directly into 96-well plates at 1 min intervals, was employed to examine a subset of these prefractions. This fractionation strategy simultaneously acquires retention time, UV absorbance, and mass spectrometric data for all components in the prefraction, providing a detailed chemical annotation to compliment the biological annotation. These ''peak libraries'' were rescreened using the standard CP platform, with recapitulation of the pertinent phenotypes in specific wells providing direct identification of the active constituents from the LC-MS trace.

The peak library of prefraction 1516D showed a single active peak (Figure 2A) with a UV absorbance profile consistent with a chromophore containing an extended conjugation system. Purification of this peak by C_{18} reverse-phase high-performance

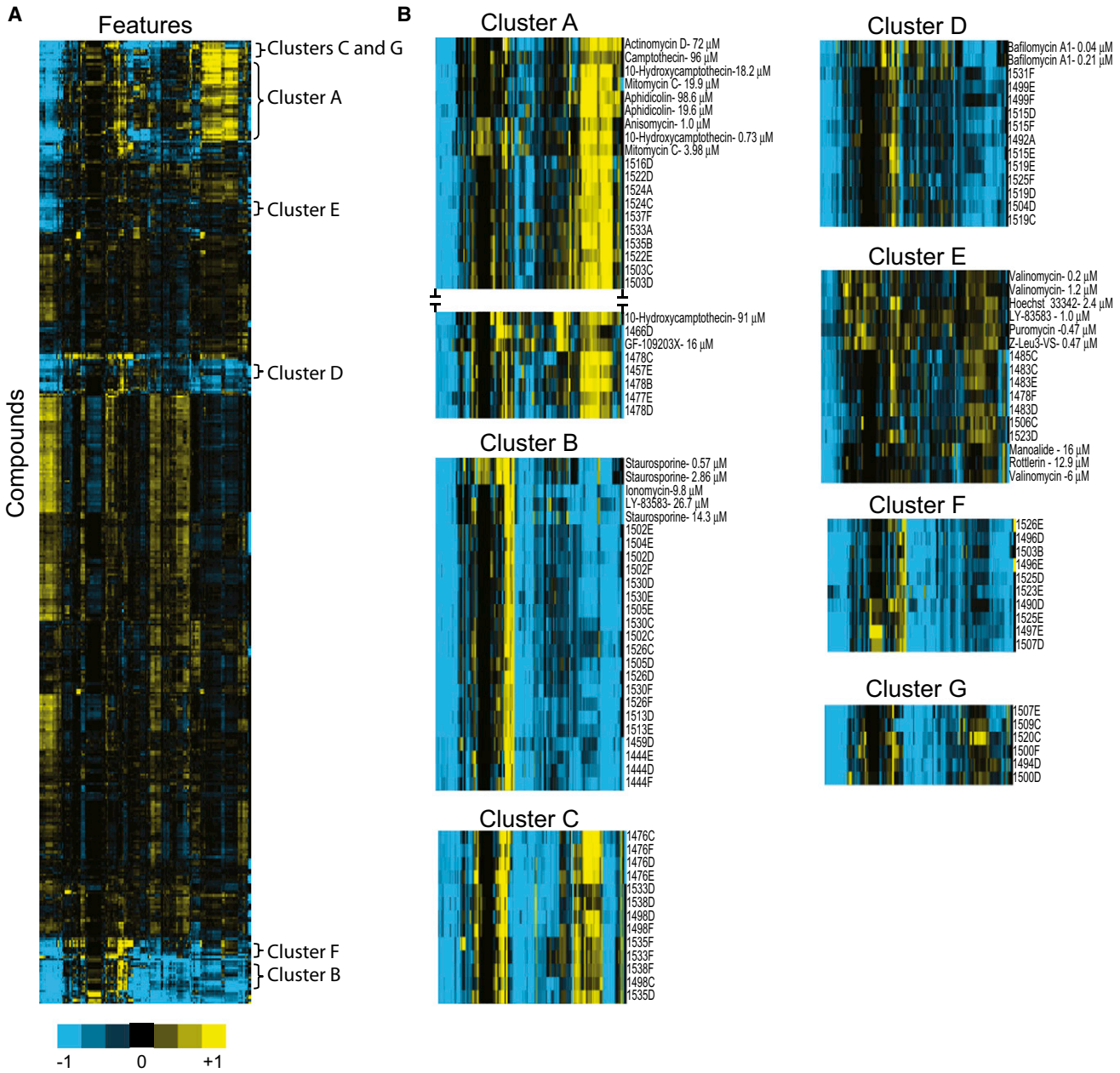

Figure 1. Heat Map Depicting Cytological Profiling Dendrogram with Pertinent Clusters Highlighted

(A) Complete data set of 624 marine natural product prefractions and 1,920 discrete samples from the ICCB collection (480 compounds at four dilutions). Individual compounds are presented on the y axis with individual features on the x axis.

(B) Detailed view of clusters A–G. In most instances, clusters contained both natural product prefractions and reference compounds from the ICCB library. Cluster A has been truncated in order to highlight the specific prefractions that were analyzed.

See also Figure S1and Table S1.

liquid chromatography (RP-HPLC) yielded a yellow amorphous solid, which exhibited a similar CP fingerprint to the prefraction (Figure S2). Analysis by (+)-FTICR MS (observed $[M + Na]^+$ at m/z 1107.4664) and interpretation of one- and two-dimensional NMR experiments (Supplemental Experimental Procedures) identified this compound as the known antitumor agent, mithramycin (Grundy et al., 1953; Wohlert et al., 1999), which was confirmed via LC-MS by coinjection with an authentic sample of mithramycin A (Figure S2). In order to corroborate the cytological profiling fingerprint exhibited by 1516D, images from this well were examined and found to exhibit very low levels of EdU and pHH3 staining, which is characteristic of cells in mitotic arrest (Figure 2B). Mithramycin is known to reversibly bind to the minor groove of GC-rich regions of DNA in the presence of Mg^{2+}, inhibiting DNA and RNA polymerase activity (Mansilla et al., 2010). In addition, the specificity for GC-rich

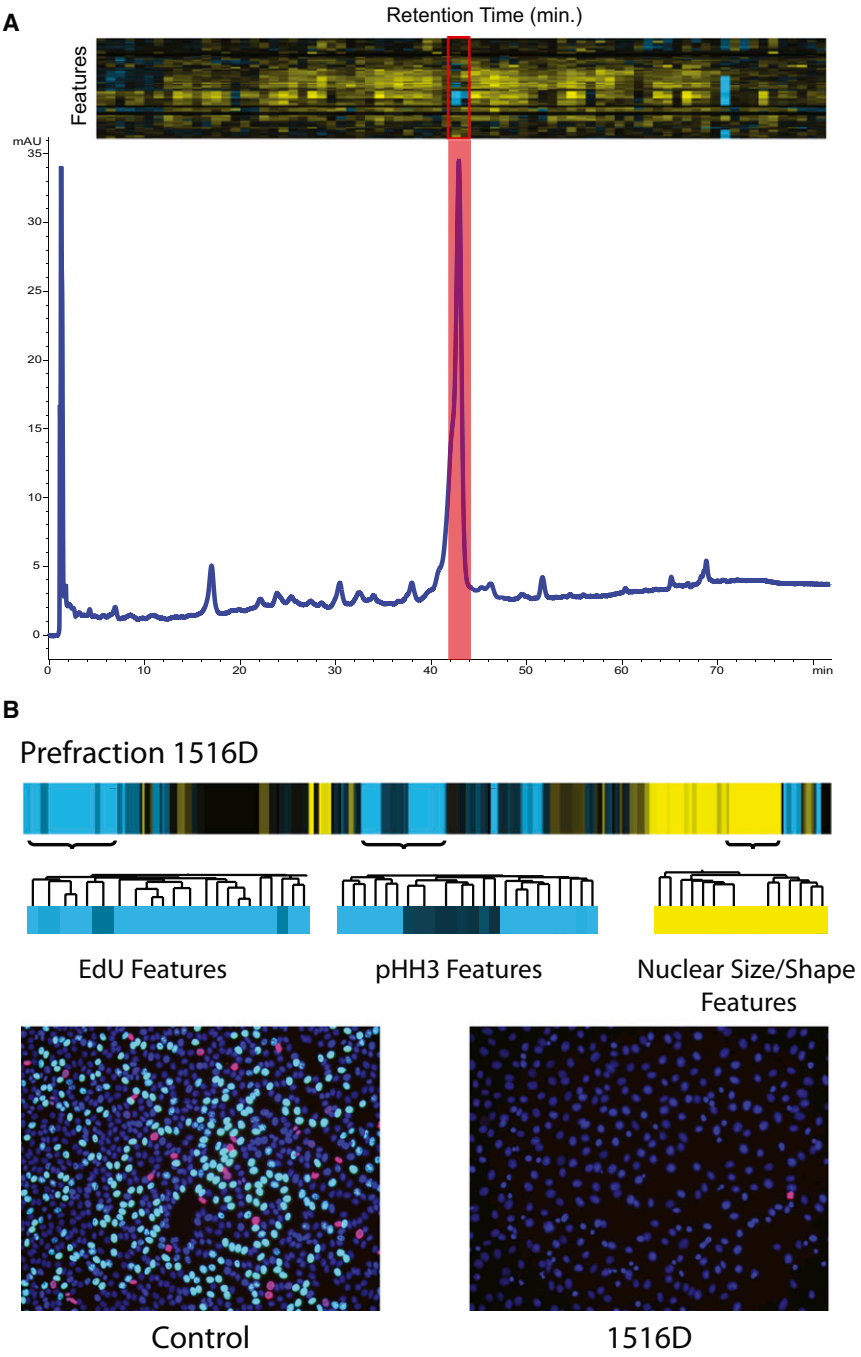

Figure 2. Identification and MOA Characterization of Prefraction 1516D from Cluster A

(A) HPLC trace (280 nm) and CP profiles for each time slice for the peak library of prefraction 1516D. CP dendrogram presented with individual features on the y axis, and time slices on the x axis. The active wells (minutes 42–44) showed a decrease in both EdU and pHH3 staining (highlighted with red box), that matched the CP fingerprint for the original prefraction. The blue features at minute 70 correspond to an unidentified media component contaminant that elutes at the end of the LC-MS trace.

(B) Full CP fingerprint and expansions of key individual features for prefraction 1516D. These expansions reveal strong negative deviations from control wells for features corresponding to DNA synthesis (EdU) and mitotic cells (pHH3). In addition, positive deviations from control wells are observed for features corresponding to nuclear size and shape, which is characteristic of cell-cycle exit at G1/S. Analysis of the fluorescent images confirmed the phenotype predicted by the CP profile (blue, Hoechst; cyan, EdU; pink, pHH3). See also Figure S2.

regions prevents the transcription of numerous regulatory proteins, including Sp1 family transcription factors, which modulates the transcription of many genes necessary for cell survival and differentiation (Remsing et al., 2003). The inclusion of this extract in Cluster A is therefore supported by the published molecular targets of mithramycin, all of which lead to cell-cycle arrest.

Further analysis of prefractions in Cluster A led to the isolation of resistomycin (Rosenbrook, 1967), resistoflavin (Eckardt et al., 1979), and tetracenomycin D (Yue et al., 1986) from prefraction 1478D. Analysis of LC-MS data revealed that resistomycin was present in an additional two extracts (1477E and 1494E), which are also present in Cluster A. The identity of each compound was verified by (−)-FTICR MS and analysis of NMR spectra (Figure S1). All three of these compounds have reported cytotoxicity against mammalian cells (Martin et al., 2002; Gorajana et al., 2007; Vijayabharathi et al., 2011). Furthermore, resistomycin has recently been shown to bind DNA by fluorescence competition experiments and circular dichroism studies (Vijayabharathi et al., 2012), and to prevent transcription by inhibition of RNA polymerase (Arora, 1985). Cytological profiling of a dilution series of the pure compounds verified that they cause cell-cycle arrest (Figure S1). The DNA binding and RNA polymerase inhibition properties of resistomycin are in accordance with the MOAs exhibited by mithramycin and the ICCB compounds in Cluster A, which supports the hypothesis that this cluster contains reference compounds and prefractions that have similar MOAs.

Discovery of Microferrioxamines A–D

Investigation of the prefractions in Cluster A has resulted in the isolation of compounds that have been previously reported to interact with DNA, and therefore prevent cells from completing the cell cycle. During this annotation process, prefraction 1522E was identified as a prefraction that displayed a similar phenotype, yet did not match known compounds from this cluster by LC-MS analysis. Secondary screening of the peak library showed the

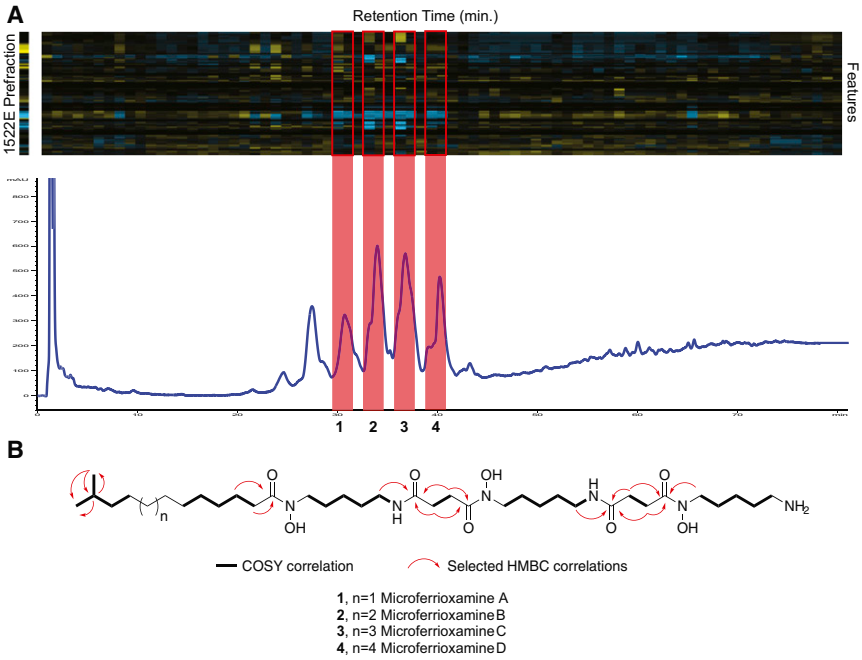

Figure 3. Identification and Characterization of Microferrioxamines A–D (1–4) from Prefraction 1522E

(A) HPLC trace (210 nm) and CP profiles for each time slice for the peak library of prefraction 1522E. CP dendrogram presented with individual features on the y axis, and time slices on the x axis. The original 1522E prefraction CP fingerprint is presented on the left for reference.

(B) Structures of microferrioxamines A–D (1–4) and key COSY and HMBC correlations for the connection of the alternating cadaverine and succinyl units and the identification of the iso-branched alkyl terminus. Structures were determined by one- and two-dimensional NMR, and MS^n experiments.

See also Figure S3 and Table S2.

presence of four major compounds responsible for the biological activity (Figure 3A). Isolation of these compounds by C_{18} RP-HPLC, followed by (+)-FTICR MS, revealed that the four compounds were structurally related, differing by 14 atomic mass units (amu). In addition to the $[M + H]^+$ signal, each of these compounds exhibited a second mass adduct that differed by an increase of 52.9 amu. Analysis of the isotope distributions for these adducts, coupled with interpretation of the mass differences, identified this mass as $[M + Fe - 2H]^+$, suggesting that these compounds had the capacity to chelate iron, which was also verified by UV spectroscopy (Figure S3). Extensive one- and two-dimensional NMR experiments (Table S2), coupled with MS^n analysis (Figure S3) led to the discovery of four iron siderophores, microferrioxamines A–D (1–4) (Figure 3B). These compounds are structurally similar to the recently reported siderophore promicroferrioxamine (Yang et al., 2011), with the C terminus being iso-branched as opposed to the anteiso-branched alkyl chain reported for promicroferrioxamine.

Iron siderophores, such as desferrioxamine (DFO), have been previously reported to have antitumor activity by causing G1/S arrest (Richardson et al., 2009). Iron depletion can have a variety of effects in tumor cells such as inhibition of ribonucleotide reductase, which is critical for DNA synthesis, as well as alteration of the expression of multiple proteins that regulate the cell cycle, such as p53, GADD45, cyclin D1, and $p21^{CIP1/WAF1}$ (Liang and Richardson, 2003; Le and Richardson, 2004; Fu and Richardson, 2007; Nurtjahja-Tjendraputra et al., 2007; Saletta et al., 2011). Although structurally dissimilar to the DNA binding agents uncovered in Cluster A, the phenotype observed in the cytological profiles of microferrioxamines A–D is consistent with cells in G1/S arrest. Discovery of this family of iron siderophores, and prediction of their mode of action validates that CP can effectively cluster compounds by MOA and demonstrates that this approach can be successfully applied to the discovery of new compounds without previously knowledge of their biological targets.

Dereplication of Known Compounds Based on Biological Fingerprints

The ICCB library contains many examples of molecules with cellular targets that are independent of cell-cycle regulation. It was expected that active compounds in our prefractions that shared similar modes of action would cluster closely with these known compounds, even if those targets were not directly examined by the fluorescent stain set. Therefore, we evaluated the most prominent clusters that contained multiple natural product prefractions and ICCB compounds to test the hypothesis that CP could be used for the rapid dereplication of prefractions containing known compounds.

Cluster B contains 19 natural product prefractions, and further examination of the dendrogram in Figure 1 shows the closest match to the ICCB training set is staurosporine, a pan-kinase inhibitor (Omura et al., 1977). LC-MS analysis of these 19 prefractions revealed that each also contained staurosporine, a commonly encountered secondary metabolite produced by Actinobacteria. Despite containing staurosporine at a range of concentrations, the cytological fingerprint of these extracts is sufficiently diagnostic for all of these prefractions to cluster tightly together, indicating that this methodology can correctly predict the presence of specific compound classes, even in the absence of defined compound concentrations or fluorescent stains that directly report on the target(s) of the active compound.

Cluster C is phenotypically similar to Cluster A (cell-cycle arrest), with the exception of the features that describe nuclear shape. Cluster C has a strongly negative deviation from the mean for these values, which signifies smaller condensed nuclei. This suggests that many of the cells are apoptotic, which is supported by evaluation of the fluorescent images showing low cell count and a phenotype typical of apoptotic cells. Analysis of the 13 prefractions in this cluster revealed that they all contain complex mixtures of anthracyclines in high concentration, as determined by their diagnostic UV profiles (Cowan et al., 1997). As a representative example, cosmomycin D (Ando et al., 1985) was isolated from prefraction 1498D, and its structure confirmed by NMR experiments (Figure S1) and

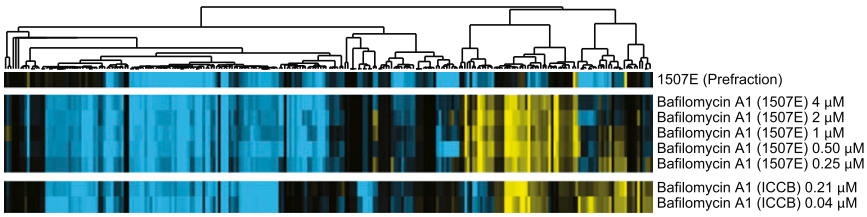

Figure 4. Cytological Profiles of Dilution Series for Bafilomycin A1 from Both Prefraction 1507E and ICCB Library and Comparison to Cytological Profile for Original 1507E Prefraction

Bafilomycin A1 was isolated and characterized from prefraction 1507E. However, this prefraction did not cluster with the references standards of bafilomycin A1 (Figure 1B). Analysis of the images revealed few viable cells in the well, which precluded the generation of an accurate CP profile. Pure bafilomycin A1, isolated from 1507E, was screened in a dilution series and compared to the reference standard concentrations of bafilomycin A1 in cluster D and prefraction 1507E. Purified bafilomycin A1 showed higher correlation to the reference standard (Pearson correlation = 0.67) than to the original prefraction (Pearson correlation = 0.46). The clustered heat map can be viewed in Figure S4.

(+)-HRESITOFMS (observed $[M + 2H]^{2+}$ at m/z 595.2996). The exact mechanism of anthracycline cytotoxicity is complicated, involving DNA intercalation and topoisomerase II poisoning. It results in apoptosis, which is in line with the phenotype observed through CP screening (Cipollone et al., 2002; Temperini et al., 2005). Despite showing a phenotype consistent with apoptotic cells, this cluster only contains anthracyclines, suggesting that cell death associated with anthracycline toxicity results in a diagnostic phenotype that is identified by our cytological profiling platform.

Cluster D exhibits multiple prefractions that cluster closely with bafilomycin A1, a vacuolar ATPase-inhibitor (Werner et al., 1984; Bowman et al., 1988). Analysis of these prefractions confirmed the presence of bafilomycin A1 (observed $[M + Na]^+$ at m/z 645.3993) in a subset of these samples (Supplemental Experimental Procedures). Other extracts in this cluster appeared to contain a biologically active compound that had a similar UV profile and mode of action to bafilomycin A1. Isolation of this compound and analysis by NMR and (+)-FTICR MS (observed $[M + Na]^+$ at m/z 838.4383) revealed that these extracts contained bafilomycin B1 (Werner et al., 1984) (Figure S1). Although this analog has substantial additional structural complexity compared to bafilomycin A1, it is has also been reported to target vacuolar ATPases (Bowman et al., 1988), validating the presence of these prefractions in this cluster and demonstrating that CP analysis can characterize compounds by mode of action regardless of structural variation among compounds that target the same cellular processes.

A final example of dereplication of known compounds based on biological mode of action involved the isolation of valinomycin (Brockmann and Schmidt-Kastner, 1955), a potassium channel ionophore (Tosteson et al., 1967) from prefractions 1483D and 1483E, which are contained in Cluster E. This molecule was isolated and characterized by (+)-FTICR MS (observed $[M + Na]^+$ at m/z 1133.6209) and one- and two-dimensional NMR experiments (Figure S1). Although the cytological fingerprint of valinomycin is weaker than other clusters, close examination of the dendrogram (Figure 1B) reveals that three concentrations of valinomycin from the ICCB library cluster closely to both of these prefractions, highlighting that CP screening can identify compound classes, even when the mode of action for these compounds is not related to the suite of fluorescent stains used to obtain the structural features. This last result is significant, as it demonstrates that the resolution of CP analysis extends well beyond the small number of processes directly quantified by the fluorescent stain set, and indicates that CP screening is capable of classifying compounds at significantly higher resolution than cell-cycle progression alone.

Analysis of Cluster F and G: Low Cell Count and Deconvolution of Multiple Compounds per Prefraction

One of the inherent obstacles in screening of natural product mixtures is that the concentration of the active constituent is unknown. All of our prefractions are diluted to a set concentration prior to screening, based on empirical observations of compound yields from liquid culture. However, although standardization of dilution volumes is a practical solution to the issue of variable metabolite yields from bacterial cultures, it does result in the creation of stock solutions containing mixed concentrations of test compounds. In cases where the actual concentration of test compounds is high, prefractions can cause complete cell death, resulting in the absence of discernible phenotypes in screening images. Fortunately, these clusters can be readily identified by strongly negative deviations from the mean for nearly all features, as represented by Clusters F and G. This "death phenotype" can be verified by examination of the nuclear count for each well, with prefractions that cause substantial cell death having very low nuclear count values compared with the average for control wells.

Closer examination of cluster G revealed the presence of bafilomycin A1 from prefraction 1507E. This was initially in contrast to our hypothesis that all prefractions containing bafilomycin A1 should be present in Cluster D. However, the low cell count in this well precluded the generation of a diagnostic cytological fingerprint. Serial dilution of the bafilomycin A1 sample isolated from prefraction 1507E, followed by cytological profiling, showed that this compound exhibited a phenotypic dose-dependent response that recapitulated the phenotype originally seen in Cluster D when compared to the relevant concentrations of bafilomycin A1 from the ICCB library (Figures 4 and S4).

Another frequently encountered challenge with natural product screening programs is the presence of multiple biologically active compounds in a single prefraction. In cytological profiling, this can result in CP fingerprints that do not represent the MOA of either compound, complicating the accuracy of the clustering approach. In the course of analyzing Cluster F, staurosporine was found to be present in prefraction 1526E. However, 1526E is separated from the staurosporine cluster (Cluster B), suggesting that the CP fingerprint for 1526E could be the result of the additive effects of two or more bioactive components. To test this hypothesis, a peak library of this prefraction was generated to afford a one-compound-one-well library for secondary

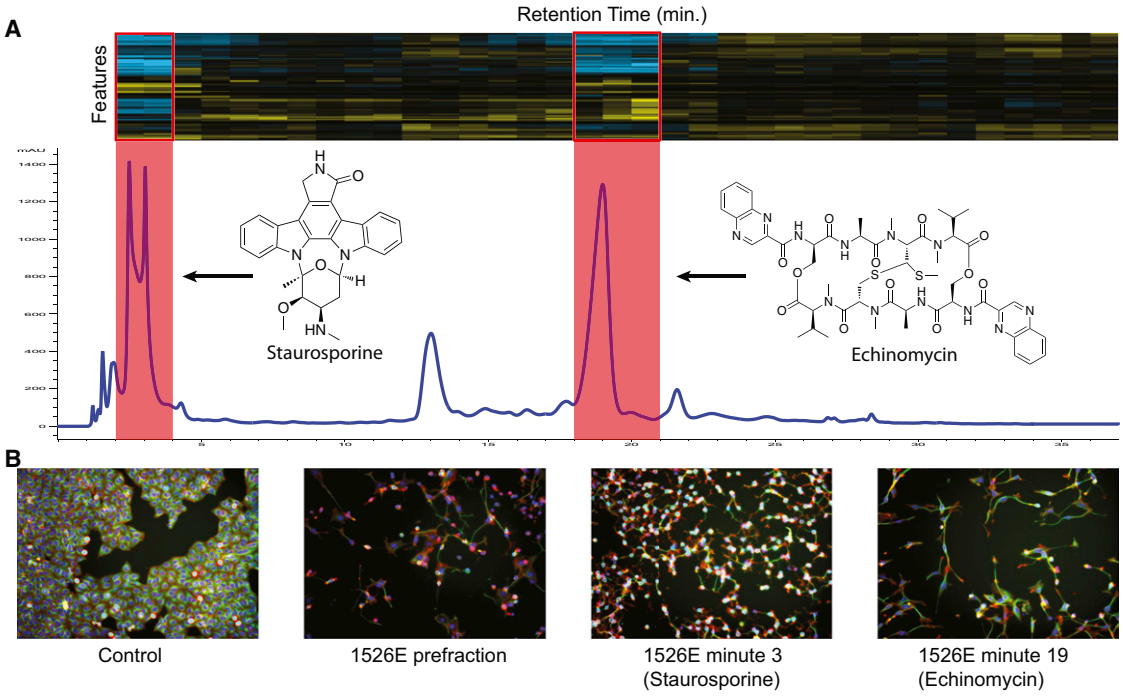

Figure 5. Delineation of Individual Biological Fingerprints from Extracts Containing Multiple Bioactive Constituents

(A) HPLC trace (230 nm) and CP profiles for each time slice for the peak library of prefraction 1526E. CP dendrogram presented with individual features on the y axis, and time slices on the x axis. Combined CP and LC-MS analyses reveal the presence of both staurosporine (minutes 2–4) and echinomycin (minutes 18–21), as two active constituents with differing modes of action in the same prefraction. Examination of the individual cytological profiles for these two compounds predicts that these compounds do not share the same MOA, and justifies the unique CP fingerprint observed for prefraction 1526E.

(B) Fluorescence microscopy screening images from primary screening and peak library plates. Images for staurosporine and echinomycin show marked differences in cell morphology consistent with different MOAs. Both pure compound images differ from the phenotypes seen for both the original prefraction (1526E) and the corresponding control well.

See also Figure S5.

screening. The strategy of employing peak library separation as a secondary screen is advantageous, as it provides parallel determination of the CP fingerprints for each individual component of the prefraction in one screening event. Screening of the peak library for 1526E revealed the presence of two bioactive components, each of which possessed a discrete phenotype (Figure 5A). One of these compounds was readily identified as staurosporine by (+)-HRESITOFMS (observed $[M + H]^+$ at m/z 467.2076) and LC-MS, as well as the similarity of the peak library CP fingerprint to the reference standard of staurosporine from the ICCB library. Purification of the unknown compound, characterization by (+)-HRESITOFMS (observed $[M + Na]^+$ at m/z 1123.4117), and assignment of one- and two-dimensional NMR spectra revealed the presence of the known bis-intercalator, echinomycin (Dell et al., 1975; Gilbert and Feigon, 1991) (Figures 5A and S5). The peak library shows a distinction between the phenotypes of staurosporine and echinomycin as expected by their differing modes of action. These results are further corroborated by examination of the fluorescent images (Figure 5B), which clearly show phenotypic divergence upon treatment with these two compounds. The combination of CP and peak library screening is therefore capable of deconvoluting complex natural product extracts containing multiple biologically active compounds, and accurately determining the precise phenotypic effect of individual constituents, even as components of complex mixtures of unknown concentration.

DISCUSSION

This study demonstrates the power of cytological profiling for natural products drug discovery. In designing this initial study, we sought to determine the suitability of CP for predicting MOAs from complex natural product mixtures, and in doing so to create a platform that would serve as a valuable tool for the natural products discovery community. Although image-based screening has previously been used to identify actin poisons from sponge extracts (Sumiya et al., 2011), this example highlights the application of CP screening for the unbiased characterization of natural products libraries. Using CP we have been able to generate detailed biological annotation of the prefractions in our library at the primary screening stage. This approach has allowed us to identify groups of prefractions that act by similar MOAs, to dereplicate the library based on profile similarity to a training set of known compounds, and has led to the discovery and MOA prediction of the iron siderophores, microferrioxamines A–D. This strategy improves upon standard cytotoxicity assays, where the output is generally percent growth inhibition, by allowing crucial, biology-driven triage decisions to be made early in the discovery process. Although MOA information is

classically acquired at a late stage in the project, CP gives an insight into the potential MOA of lead compounds at the primary screening stage. This provides an opportunity to prioritize subsets of prefractions from large libraries, and permits explicit inclusion or exclusion of particular MOAs based on programmatic screening priorities.

Arrest of the cell cycle (Cluster A) is one of the most readily identifiable clusters in the dendrogram (Figure 1A). This is, in part, due to the fact that our set of fluorescent probes is directly observing the extent of DNA synthesis, the progression into mitosis, and the size and shape of nuclei. However, both our prefractions and the ICCB library contain compounds that act by other distinct and diverse mechanisms. We are able to profile compounds that act as kinase inhibitors (staurosporine), vacuolar ATPase inhibitors (bafilomycin A1 and B1), and K^+ ionophores (valinomycin), despite not probing any of these features directly. For example, mitochondrial swelling is one of the hallmark features of treatment of cells with valinomycin (Klein et al., 2011). Although our stain set does not contain probes that are specific for mitochondrial shape, the overall cellular phenotype is sufficient for clustering prefractions containing valinomycin with pure compound reference samples. In addition to the aforementioned mechanisms, we can also identify other clusters with clinically relevant MOAs, including actin and microtubule poisons. However, none of the prefractions tested here contained compounds that act by either of these mechanisms and these phenotypes were therefore not investigated further. As an extension to this technology, we can envision the creation of customized stain sets that visualize the effects on specific organelles, such as the mitochondria, golgi apparatus, or endoplasmic reticulum, which would enhance sensitivity for particular features of interest.

In extending the CP methodology to natural products libraries, we were particularly interested to explore whether this technology could predict the modes of action for previously uncharacterized compounds. To this end, four structurally related iron siderophores were isolated and characterized from prefraction 1522E in Cluster A. The discovery of this family of natural products highlights the ability of CP to aid in the discovery of novel chemistry, and supports the hypothesis that CP clustering is driven by biological MOA rather than structural similarity. In many cases, we have observed clusters of prefractions that contain the same molecule clustering closely with the respective pure compound from the training set. This is the expected result; however, it does not address the hypothesis that clustering is based solely on compound MOA, regardless of similarity in chemical structure. Although other compounds isolated from prefractions in Cluster A, mithramycin and resistomycin, have substantial structural differences, they both share flat aromatic moieties that are known to interact with DNA. On the other hand, cell-cycle arrest caused by the microferrioxamines is presumably caused by intracellular iron chelation, which has downstream effects on numerous proteins that regulate the cell cycle. These molecules share little structural resemblance to any other compounds found in Cluster A and are not thought to interact directly with DNA. Yet their inclusion in Cluster A is driven by the same cellular phenotype that is observed from other compounds that have classically been known to cause cell-cycle arrest.

Although CP is able to give an accurate picture of the general mode of action for most compounds, it is unable to provide a detailed output of the precise molecular target(s). Rather, it allows for a MOA hypothesis to be formed at the outset, requiring additional downstream biochemical experiments in order to elucidate the molecular target for a given compound. This level of detail represents a significant improvement over the live/dead assays that are classically used to screen large natural product libraries, and provides an opportunity to consider alternative approaches to natural products-based drug discovery. Among these, prioritization of prefractions with low CP fingerprint similarity to the training set of known drugs provides an avenue for discovering compounds that hit targets discrete from those covered by current therapeutics. In addition, CP screening has the potential to be used to identify compounds that cause phenotypes similar to those generated by genetic manipulation or siRNA treatment for specific pathways of interest. The ability to garner MOA information from primary screens therefore represents a valuable tool for natural products drug discovery research programs.

It is pertinent to recognize the limitations of using cytological profiling as a primary screen in natural products discovery, and how we have attempted to overcome some of the inherent obstacles that exist in using this strategy. One of the major challenges in any natural products screening platform is that the concentrations of compounds in the library are largely unknown. As CP profiles can be affected by widely different concentrations of active compounds, four concentrations of each compound in the ICCB library were included to improve clustering with natural product prefractions containing active compounds of unknown titer. We generally observe two or three of these concentrations clustering together, with the highest and lowest concentrations resulting in cell death, or inactivity, respectively. This demonstrates that CP fingerprints typically remain constant over a wide concentration range, and that most compounds possess broad concentration ranges that afford recognizable CP fingerprints. When potent compounds exist in high titer, cell counts become low, precluding the creation of informative cytological profiles. However, these clusters can be readily identified and screened again at appropriate dilutions by examining nuclear count values. Serial dilution of prefractions of interest affords dose-dependent phenotypes, which can be incorporated into the data set to determine accurate positioning within the dendrogram.

Another potential problem with CP screening concerns the presence of multiple biologically active compounds in a single prefraction. If the prefraction contains two compounds with distinct MOAs, as seen in 1526E (Figure 5A), the cytological profile will show divergence from both of the expected individual profiles. Our approach allows the delineation of the activities of multiple bioactive compounds by secondary screening of peak libraries. Using this strategy, the CP profiles of each individual compound in the prefraction can be obtained independently. These independent profiles can then be reincorporated into the main CP data set, and clustered with the reference library to determine the accurate position in the dendrogram for each compound.

Validation of CP for profiling natural products libraries sets the stage for further applications in this area. Identification of

compounds with defined MOA-specific fingerprints can be used to expand chemical space in target areas where current therapeutic options are limited, or where current lead compounds suffer from poor pharmacokinetic properties. This would best be achieved through the inclusion of additional positive controls in the training set to generate phenotypes for specific targets, and comparison of these phenotypes to the CP results for screening libraries. Alternatively, CP screening can be used to identify prefractions with fingerprints that do not match compounds from the training set. The presence of unique CP fingerprints within the natural products library suggests that either the prefraction contains multiple bioactive components, in which case the true CP fingerprints can be determined using peak library screening or, more excitingly, that the active constituent does not act by any of the commonly encountered MOAs. Determining this information confers a significant advantage in selecting active prefractions for follow-up studies. We have recently expanded this platform to our entire library of 3,168 marine natural product prefractions and have identified multiple clusters of compounds that show divergence from the MOAs of compounds in the ICCB training set. Although a unique fingerprint does not give an unequivocal indication of an unknown MOA, it allows for hit prioritization that would be impossible by considering cell viability alone. Follow-up studies on these prefractions hold promise for the discovery of new antitumor natural products that could have significance as chemical probes to identify new targets or lead scaffolds for development as next-generation therapeutics.

SIGNIFICANCE

Recent advances in image-based screening technologies and data analysis strategies provide an opportunity to expand the repertoire of methods used in natural products discovery. We have applied a modified version of cytological profiling to our library of marine-derived bacterial prefractions in order to gain biological information about their modes of action at the primary screening stage. These data have allowed for the prediction of modes of action for specific compounds directly from prefractions, which have been corroborated by the isolation of known compounds that are consistent with these predictions. The use of CP to interrogate unknown natural product prefractions is an important step in the evolution of this technology, as it opens the door for the discovery of new chemical scaffolds while simultaneously offering evidence of their biological significance. Highlighting this aspect, we have isolated a structurally unique family of iron siderophores that are predicted to cause mitotic arrest by analysis of their cytological profile. This prediction has been verified by profiling of the pure compounds, and is supported by literature precedence.

As natural products continue to produce structurally diverse molecules with unique biological functions, we envision that this platform will play an increasing role in "function-first" approaches to lead discovery. The number of published natural product structures now stands at over 200,000, but the biological annotation of many of these structures is limited. Therefore, there is currently an unexplored opportunity to ascribe detailed biological functions to many of these scaffolds and, in the process, to discover the next generation of tools for both chemical genetic studies, and lead compounds for future drug development.

EXPERIMENTAL PROCEDURES

Preparation of Peak Libraries

A 45 μl aliquot of prefraction DMSO stock was lyophilized and fractionated by C_{18} reversed- phase HPLC (Phenomenex Synergi Fusion-RP, 10 × 250 mm column, 2 ml min^{-1} flow rate) using a MeOH/H_2O (0.02% formic acid) solvent system. Each prefraction was run on a gradient specifically tailored to produce the most highly resolved chromatography. Eluent was collected into deep well 96-well plates using an automated time-based fraction collection method consisting of 1 min time slices, and subsequently concentrated to dryness in vacuo. Dried plates were resolubilized (10 μl DMSO per well), sonicated to ensure homogeneity, reformatted to 384-well format and submitted for biological screening.

Compound Transfer and Cell Culture

HeLa cells were plated into two clear-bottom 384-well plates (Corning) at a density of 2,500 cells per well in 25 μl Dulbecco's modified Eagle's medium (DMEM). The cells were incubated for 24 hr at 37°C under 5% CO_2. After this incubation period, 150 nl of natural product prefractions in DMSO were transferred to the 384-well culture plates using the pin tool attachment on the Janus MDT (Perkin-Elmer). The plates were transferred to 37°C for 19 hr.

Cell Staining

After 19 hr of incubation, the two 384-well plates were treated with either nuclear or cytoskeletal stain sets. For the nuclear stain set, 20 μM of 5-ethynyl-2′-deoxyuridine (EdU) in DMEM was added to cells and incubated for 1 hr at 37°C. Cells were fixed with 4% formaldehyde for 20 min, washed with PBS, treated with 0.5% Triton X-100 for 10 min, and washed with PBS. Plates were blocked with a 2% BSA PBS solution for 20 min, washed with PBS, and stained with rhodamine-azide using click chemistry by incubation with 4 mM $CuSO_4$, 2 mg/ml sodium ascorbate, and 1 μg/ml rhodamine azide in 100 mM TRIS buffer for 30 min at 25°C in the dark. After washing with PBS, rabbit antiphospho-histone H3 (Millipore) in 2% BSA PBS was added and the plate was incubated overnight at 4°C. Plates were rinsed with PBS, and Cy5 conjugated goat anti-rabbit antibody and Hoechst stain were added for 1 hr in 2% BSA in PBS at 25°C. The plates were then left in 0.1% azide in PBS. For the cytoskeletal stain set, cells were fixed with 4% formaldehyde for 20 min, then washed with PBS. After treatment with 0.5% Triton X-100 for 10 min, the plates were washed with PBS and blocked with a 2% BSA PBS solution. After washing with PBS, plates were incubated overnight at 4°C with FITC-conjugated mouse anti-α-tubulin and rabbit antiphospho-histone H3 antibodies in 2% BSA PBS. After washing with PBS, Cy5-conjugated goat anti-rabbit antibody, rhodamine-labeled phalloidin, and Hoechst stain in 2% BSA PBS were added and plates were incubated for 1 hr at 25°C. The plates were then left in 0.1% azide in PBS.

Imaging and Image Analysis

The plates were imaged using an ImageXpress Micro epifluorescent microscope (Molecular Devices) with a 10× Nikon objective lens. We obtained four images per well for each wavelength in a plate resulting in 4,608 images for the nuclear stain set and 6,144 images for the cytoskeletal stain set. The plates were analyzed using MetaXpress (Molecular Devices). Measurements were taken using built-in morphometry metrics, the multiwavelength cell scoring, transfluor, and micronuclei modules. These measurements were converted to feature scores, clustered using the Cluster 3.0 software, and analyzed using Java TreeView.

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures, two tables, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.chembiol.2012.12.007.

ACKNOWLEDGMENTS

This work is supported in part by the California Institute for Quantitative Biosciences (to R.G.L.), the US National Institutes of Health grant R01GM084530 (to R.S.L.), and the US National Science Foundation CAREER award NSF DBI:0845783 (to J.S.). Support for the UCSC Chemical Screening Center was provided through US National Institutes of Health grant 1-S10-RR022455. We thank J. Davies for assistance with automation of data analysis, B. Dickey and M. Morales for assistance with natural product library generation, and Y.-L. Yang and P. Dorrestein for provision of authentic NMR spectra for promicroferrioxamine.

Received: September 20, 2012
Revised: November 27, 2012
Accepted: December 13, 2012
Published: February 21, 2013

REFERENCES

Ando, T., Hirayama, K., Takahashi, R., Horino, I., Etoh, Y., Morioka, H., Shibai, H., and Murai, A. (1985). Cosmomycin D, a new anthracycline antibiotic. Agric. Biol. Chem. *49*, 259–262.

Arora, S.K. (1985). Molecular structure of heliomycin, an inhibitor of RNA synthesis. J. Antibiot. (Tokyo) *38*, 113–115.

Bowman, E.J., Siebers, A., and Altendorf, K. (1988). Bafilomycins: a class of inhibitors of membrane ATPases from microorganisms, animal cells, and plant cells. Proc. Natl. Acad. Sci. USA *85*, 7972–7976.

Brockmann, H., and Schmidt-Kastner, G. (1955). Valinomycin I, XXVII. Mitteil. über Antibiotica aus Actinomyceten. Chem. Ber. *88*, 57–61.

Cipollone, A., Berettoni, M., Bigioni, M., Binaschi, M., Cermele, C., Monteagudo, E., Olivieri, L., Palomba, D., Animati, F., Goso, C., et al. (2002). Novel anthracycline oligosaccharides: influence of chemical modifications of the carbohydrate moiety on biological activity. Bioorg. Med. Chem. *10*, 1459–1470.

Cowan, W.M., Jessell, T.M., and Zipursky, S.L. (1997). Molecular and cellular approaches to neural development (New York: Oxford University Press).

Dell, A., Williams, D.H., Morris, H.R., Smith, G.A., Feeney, J., and Roberts, G.C.K. (1975). Structure revision of the antibiotic echinomycin. J. Am. Chem. Soc. *97*, 2497–2502.

Eckardt, K., Fritzsche, H., and Tresselt, D. (1979). Composition of antibiotic resistoflavin. Tetrahedron *35*, 1621–1624.

Fu, D., and Richardson, D.R. (2007). Iron chelation and regulation of the cell cycle: 2 mechanisms of posttranscriptional regulation of the universal cyclin-dependent kinase inhibitor p21CIP1/WAF1 by iron depletion. Blood *110*, 752–761.

Gerwick, W.H., and Moore, B.S. (2012). Lessons from the past and charting the future of marine natural products drug discovery and chemical biology. Chem. Biol. *19*, 85–98.

Gilbert, D.E., and Feigon, J. (1991). The DNA sequence at echinomycin binding sites determines the structural changes induced by drug binding: NMR studies of echinomycin binding to [d(ACGTACGT)]2 and [d(TCGATCGA)]2. Biochemistry *30*, 2483–2494.

Gorajana, A., Venkatesan, M., Vinjamuri, S., Kurada, B.V.V.S.N., Peela, S., Jangam, P., Poluri, E., and Zeeck, A. (2007). Resistoflavine, cytotoxic compound from a marine actinomycete, Streptomyces chibaensis AUBN1/7. Microbiol. Res. *162*, 322–327.

Grundy, W., Goldstein, A., Rickher, C., Hanes, M., Warren, H., and Sylvester, J. (1953). Aureolic acid, a new antibiotic. 1. Microbiologic studies. Antibiot. Chemother. *3*, 1215–1217.

Harvey, A.L. (2007). Natural products as a screening resource. Curr. Opin. Chem. Biol. *11*, 480–484.

Klein, B., Wörndl, K., Lütz-Meindl, U., and Kerschbaum, H.H. (2011). Perturbation of intracellular K(+) homeostasis with valinomycin promotes cell death by mitochondrial swelling and autophagic processes. Apoptosis *16*, 1101–1117.

Koehn, F.E., and Carter, G.T. (2005). The evolving role of natural products in drug discovery. Nat. Rev. Drug Discov. *4*, 206–220.

Lachance, H., Wetzel, S., Kumar, K., and Waldmann, H. (2012). Charting, navigating, and populating natural product chemical space for drug discovery. J. Med. Chem. *55*, 5989–6001.

Lamb, J., Crawford, E.D., Peck, D., Modell, J.W., Blat, I.C., Wrobel, M.J., Lerner, J., Brunet, J.-P., Subramanian, A., Ross, K.N., et al. (2006). The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. Science *313*, 1929–1935.

Le, N.T.V., and Richardson, D.R. (2004). Iron chelators with high antiproliferative activity up-regulate the expression of a growth inhibitory and metastasis suppressor gene: a link between iron metabolism and proliferation. Blood *104*, 2967–2975.

Liang, S.X., and Richardson, D.R. (2003). The effect of potent iron chelators on the regulation of p53: examination of the expression, localization and DNA-binding activity of p53 and the transactivation of WAF1. Carcinogenesis *24*, 1601–1614.

Mansilla, S., Garcia-Ferrer, I., Méndez, C., Salas, J.A., and Portugal, J. (2010). Differential inhibition of restriction enzyme cleavage by chromophore-modified analogues of the antitumour antibiotics mithramycin and chromomycin reveals structure-activity relationships. Biochem. Pharmacol. *79*, 1418–1427.

Martin, P., Rodier, S., Mondon, M., Renoux, B., Pfeiffer, B., Renard, P., Pierré, A., and Gesson, J.P. (2002). Synthesis and cytotoxic activity of tetracenomycin D and of saintopin analogues. Bioorg. Med. Chem. *10*, 253–260.

Mitchison, T.J. (2005). Small-molecule screening and profiling by using automated microscopy. ChemBioChem *6*, 33–39.

Newman, D.J., and Cragg, G.M. (2012). Natural products as sources of new drugs over the 30 years from 1981 to 2010. J. Nat. Prod. *75*, 311–335.

Nurtjahja-Tjendraputra, E., Fu, D., Phang, J.M., and Richardson, D.R. (2007). Iron chelation regulates cyclin D1 expression via the proteasome: a link to iron deficiency-mediated growth suppression. Blood *109*, 4045–4054.

Omura, S., Iwai, Y., Hirano, A., Nakagawa, A., Awaya, J., Tsuchya, H., Takahashi, Y., and Masuma, R. (1977). A new alkaloid AM-2282 OF Streptomyces origin. Taxonomy, fermentation, isolation and preliminary characterization. J. Antibiot. (Tokyo) *30*, 275–282.

Perlman, Z.E., Slack, M.D., Feng, Y., Mitchison, T.J., Wu, L.F., and Altschuler, S.J. (2004). Multidimensional drug profiling by automated microscopy. Science *306*, 1194–1198.

Remsing, L.L., Bahadori, H.R., Carbone, G.M., McGuffie, E.M., Catapano, C.V., and Rohr, J. (2003). Inhibition of c-src transcription by mithramycin: structure-activity relationships of biosynthetically produced mithramycin analogues using the c-src promoter as target. Biochemistry *42*, 8313–8324.

Richardson, D.R., Kalinowski, D.S., Lau, S., Jansson, P.J., and Lovejoy, D.B. (2009). Cancer cell iron metabolism and the development of potent iron chelators as anti-tumour agents. Biochim. Biophys. Acta *1790*, 702–717.

Rosenbrook, W., Jr. (1967). The structure of resistomycin. J. Org. Chem. *32*, 2924–2925.

Saletta, F., Suryo Rahmanto, Y., Siafakas, A.R., and Richardson, D.R. (2011). Cellular iron depletion and the mechanisms involved in the iron-dependent regulation of the growth arrest and DNA damage family of genes. J. Biol. Chem. *286*, 35396–35406.

Sumiya, E., Shimogawa, H., Sasaki, H., Tsutsumi, M., Yoshita, K., Ojika, M., Suenaga, K., and Uesugi, M. (2011). Cell-morphology profiling of a natural product library identifies bisebromoamide and miuraenamide A as actin filament stabilizers. ACS Chem. Biol. *6*, 425–431.

Sutherland, J.J., Low, J., Blosser, W., Dowless, M., Engler, T.A., and Stancato, L.F. (2011). A robust high-content imaging approach for probing the mechanism of action and phenotypic outcomes of cell-cycle modulators. Mol. Cancer Ther. *10*, 242–254.

Swinney, D.C., and Anthony, J. (2011). How were new medicines discovered? Nat. Rev. Drug Discov. *10*, 507–519.

Temperini, C., Cirilli, M., Aschi, M., and Ughetto, G. (2005). Role of the amino sugar in the DNA binding of disaccharide anthracyclines: crystal structure of the complex MAR70/d(CGATCG). Bioorg. Med. Chem. *13*, 1673–1679.

Tosteson, D.C., Cook, P., Andreoli, T., and Tieffenberg, M. (1967). The effect of valinomycin on potassium and sodium permeability of HK and LK sheep red cells. J. Gen. Physiol. *50*, 2513–2525.

Vijayabharathi, R., Bruheim, P., Andreassen, T., Raja, D.S., Devi, P.B., Sathyabama, S., and Priyadarisini, V.B. (2011). Assessment of resistomycin, as an anticancer compound isolated and characterized from Streptomyces aurantiacus AAA5. J. Microbiol. *49*, 920–926.

Vijayabharathi, R., Sathyadevi, P., Krishnamoorthy, P., Senthilraja, D., Brunthadevi, P., Sathyabama, S., and Priyadarisini, V.B. (2012). Interaction studies of resistomycin from Streptomyces aurantiacus AAA5 with calf thymus DNA and bovine serum albumin. Spectrochim. Acta A Mol. Biomol. Spectrosc. *89*, 294–300.

Werner, G., Hagenmaier, H., Drautz, H., Baumgartner, A., and Zähner, H. (1984). Metabolic products of microorganisms. 224. Bafilomycins, a new group of macrolide antibiotics. Production, isolation, chemical structure and biological activity. J. Antibiot. (Tokyo) *37*, 110–117.

Wohlert, S.E., Künzel, E., Machinek, R., Méndez, C., Salas, J.A., and Rohr, J. (1999). The structure of mithramycin reinvestigated. J. Nat. Prod. *62*, 119–121.

Yang, Y.L., Xu, Y., Kersten, R.D., Liu, W.T., Meehan, M.J., Moore, B.S., Bandeira, N., and Dorrestein, P.C. (2011). Connecting chemotypes and phenotypes of cultured marine microbial assemblages by imaging mass spectrometry. Angew. Chem. Int. Ed. Engl. *50*, 5839–5842.

Young, D.W., Bender, A., Hoyt, J., McWhinnie, E., Chirn, G.W., Tao, C.Y., Tallarico, J.A., Labow, M., Jenkins, J.L., Mitchison, T.J., and Feng, Y. (2008). Integrating high-content screening and ligand-target prediction to identify mechanism of action. Nat. Chem. Biol. *4*, 59–68.

Yue, S., Motamedi, H., Wendt-Pienkowski, E., and Hutchinson, C.R. (1986). Anthracycline metabolites of tetracenomycin C-nonproducing Streptomyces glaucescens mutants. J. Bacteriol. *167*, 581–586.