Using High-Throughput Screening Data To Discriminate Compounds with Single-Target Effects from Those with Side Effects

Justin Klekota,*[,§,⊥] Erik Brauner,[‡] Frederick P. Roth,*[,||] and Stuart L. Schreiber[†,§,⊥]

Howard Hughes Medical Institute, 12 Oxford Street, Cambridge, Massachusetts 02138, Harvard Institute of Chemistry and Cell Biology, 250 Longwood Avenue, SGMB-604, Boston, Massachusetts 02115, Broad Institute of Harvard and MIT, 7 Cambridge Center, Cambridge, Massachusetts 02142, Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 250 Longwood Avenue, SGMB-322, Boston, Massachusetts 02115, and Harvard University, University Hall, Cambridge, Massachusetts 02138

Received November 14, 2005

The most desirable compound leads from high-throughput assays are those with novel biological activities resulting from their action on a single biological target. Valuable resources can be wasted on compound leads with significant 'side effects' on additional biological targets; therefore, technical refinements to identify compounds that primarily have effects resulting from a single target are needed. This study explores the use of multiple assays of a chemical library and a statistic based on entropy to identify lead compound classes that have patterns of assay activity resulting primarily from small molecule action on a single target. This statistic, called the coincidence score, discriminates with 88% accuracy compound classes known to act primarily on a single target from compound classes with significant side effects on nonhomologous targets. Furthermore, a significant number of the compound classes predicted to have primarily single-target effects contain known bioactive compounds. We also show that a compound's known biological target or mechanism of action can often be suggested by its pattern of activities in multiple assays.

INTRODUCTION

Since the advent of high-throughput assay technology and combinatorial chemistry, millions of compounds have been screened for a wide variety of biological activities,[1−3] and many successful drug leads have been discovered.[4,5] Computational tools have a central role in interpreting biological screening data to identify compound leads.[6,7] However even with current techniques, identifying good leads from phenotypic screening data is still a challenge given that the number and identity of a lead compound's targets are often unknown and that some leads have multiple targets. Generally speaking, a good lead compound identified in a high-throughput assay should have a single biological target and additional assay phenotypes that better classify its mechanism of action. Computational tools may have a role in identifying such leads.

Phenotypic assays are widely used in chemical genetics to identify compound leads but are often confounded by effects of a small molecule on multiple unknown targets. In a forward chemical-genetics experiment, a collection of compound leads identified by a high-throughput phenotypic assay is studied in subsequent rounds of follow-up assays until the biological targets of those compounds can be identified.[8−26] Because these primary assays survey phenotypes, rather than binding to a single protein, many different biological targets are possible. By contrast, in a reverse chemical genetics experiment, a compound target is known, but the small-molecule-induced phenotype is not well-characterized:[27] this also presents experimental challenges requiring numerous rounds of follow-up phenotypic assays and the possible discovery of side effects.

Compound activity may be the result of a small molecule's action on a single biological target that induces more than one phenotype or a small molecule's action on different biological targets. If a lead compound responsible for different phenotypes has multiple targets, it may be difficult to determine the target(s) inducing each phenotype. Furthermore, the compound's effects on other targets may be undesirable. Therefore, discriminating compound leads with single-target effects from those with multiple-target ('side') effects is crucial to identifying the most promising lead compounds.

To meet this challenge, computational tools must be able to model multiple classes of active compounds with multiple assay phenotypes induced by unknown targets. Certain computational tools used in drug discovery, such as decision trees and Bayesian learning, can predict activity of multiple structural classes using noisy screening data; however, they generally assume activity on a single target or a single therapeutic phenotype.[28−35] In a notable exception, decision trees were used to model activity on a small number of different targets simultaneously.[36] Alternatively, clustering algorithms have been widely used to identify distinct biologically active compound classes[37−40] and to segregate compounds into distinct structural classes that differ in their assay phenotypes.[41]

* Corresponding author e-mail: Fritz_Roth@hms.harvard.edu (F.P.R.) and Klekota@gmail.com (J.K.).
† Howard Hughes Medical Institute.
‡ Harvard Institute of Chemistry and Cell Biology.
§ Broad Institute of Harvard and MIT.
|| Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School.
⊥ Harvard University.

10.1021/ci050495h CCC: $33.50 © 2006 American Chemical Society
Published on Web 05/18/2006

This paper explores the potential of computational tools, namely clustering and an entropy-based coincidence score, to identify compound classes active in multiple assays and generate hypotheses about a compound's biological target(s) or mechanism(s) of action. Clustering appears to be an appropriate initial first step toward defining and understanding the multiplicity of compound structural classes. Here we apply an entropy-based coincidence score to structurally defined groups of compounds ('compound classes'), assessing whether the pattern of assay activities within a compound class suggests a single target or mechanism of action. We show that this score discriminates compound classes primarily known to have single-target effects (and possibly weaker undetected side effects) from those with significant side effects on nonhomologous targets. We also show that trends in the activities of a compound class confirm the known activity of many compounds and suggest the compound's known target or mechanism.

METHODS

Chemical Descriptors. Daylight fingerprints were used to represent each compound and then used to identify compound classes through clustering.[37−40,42−44] Daylight fingerprints containing 4096 bits were used to encode two-dimensional substructures (i.e. no specified stereochemistry) up to 7 bonds in length for each compound in the Chembridge Diverse Set E library.[41]

All library compounds were stored in SD files and were subsequently converted to nonisomeric canonical SMILES using Daylight's *mol2smi* algorithm. The SMILES format was edited to remove salts and charges, using a modified version of a previously described Perl module.[45] The 4096-bit Daylight fingerprints were then calculated directly from the nonisomeric canonical SMILES representations.

Clustering. The Daylight fingerprints allow each compound in the 16 320-member Chembridge Diverse Set E library to be represented by a point in a 4096-dimension chemical space with each dimension corresponding to a single bit in the fingerprint. The compound points were partitioned by K-modal clustering with Euclidean distance as a measure of similarity into $K = 2$, 5, 11, 21, 42, 85, 170, 340, 680, 1360, and 2720 clusters using modified open source software from Cluster 3.0.[46] Euclidean distance was used instead of Tanimoto coefficients,[47] because it better segregates biologically active and inactive compounds.[41] Because K-modal clustering is stochastic (random initialization), clustering for each choice of K was performed in triplicate. Each cluster represents an approximate structural class of compounds.

Class Annotation with Assay Data. Each compound class learned by clustering was annotated with data from 48 high-throughput assays performed at the Harvard Institute of Chemistry and Cell Biology (ICCB). This diverse collection of assays includes target-based fluorescence polarization assays,[27] growth-based assays,[8,9] and phenotypic assays measuring microtubule assembly, actin polymerization, endocytosis, and acetylation among others.[10−23] Given the large number of assays screened and the large number of targets that can induce a single phenotype, there is a good chance that chemical action on a single biological target will induce more than one assay phenotype.

Assays were performed once or in duplicate using compound concentrations on the order of 10 μM. The 'readout' for these assays was fluorescence, luminescence intensity, or visually assessed phenotypes from automated microscopic images. Given the large signal variation associated with high throughput assay measurements, typically only two or three states (hit/nonhit or enhancer/nonhit/suppressor) were distinguishable.

Results were filtered to remove defective plates and fluorescent compounds screened in assays with a fluorescent readout. Substantially the same collection of phenotypic assay data has previously been shown to contain information useful in describing the biological activity of small molecules[41] and has led to identification of numerous biologically active molecules.[8−23,26]

To annotate each structural class with assay data from the 48 ICCB high-throughput screens, screening results for each compound were converted to binary designations ('hit' or 'nonhit') such that the top or bottom 4% of screened compounds were considered hits in a given assay if it employed a numeric measurement (visually assessed phenotypes were textual and did not use thresholds.) This choice of thresholds was based on previous results showing that top-4% and bottom-4% thresholds yield statistically significant concentrations of hits in compound classes obtained by clustering,[41] and these measurements were filtered statistically by class scoring (described below) before their use. In most assays, only one signal (either the high or the low signal) is considered biologically meaningful, but we found that the second signal was often indicative of other compound properties of interest, such as compound toxicity or fluorescence. For example, in most cell-based assays, a high "top" signal indicates a desired phenotype, such as phosphorylation of a target protein as detected by an antibody, and a low "bottom" signal indicates lack of cell growth (compound toxicity).

Scoring Compound Classes. Next, we identified assay activities that were significantly enriched in each compound class and subsequently considered only those enriched assay activities for that class. Significance was assessed using the cumulative hypergeometric probability, used previously to predict the assay activities of compounds within a class with as much as 87% sensitivity and fewer false positives than single molecule scores.[41] The cumulative hypergeometric probability (or class score) measures the probability of obtaining the observed number of hits within a class or more by chance alone given the total number of compounds (library size) and the total number of hits for the assay type under consideration.

The class score[41] was used to evaluate each assay activity in each class. For a class of c tested compounds with h hits given a library with N tested compounds and H total hits in assay a, the probability of getting h or more hits in the class is

$$P_a+(c, h, N, H) = \Sigma_{i=h...\min(c,H)} C(H, i) * C(N-H, c-i)/C(N, c)$$

where

$$C(c, h) = c!/(c-h)!\, h!$$

For each compound class *c*, the set of assays $\{A|c\}$ for which their respective class scores P_a+ were less than 0.005 was determined. The cutoff of 0.005 was chosen because class scores with this probability were very infrequent in randomized data sets. In effect, only compound "hit" measurements that were also present in a class statistically enriched in that particular assay activity (P_a+ less than 0.005) were retained in this analysis: this reduced the actual percentage of hits in any given assay from the starting value of 4% to a smaller rate closer to 1%. The actual percentage of hits for each assay (and threshold) varied significantly between assays, ranging from 0 to 2%, reflecting the different numbers of active compounds expected in different assays as detected by class scoring (see Supporting Information E). Furthermore, the "hit" measurements that remained after statistical filtering were previously shown to be more reliable predictors of activity despite being less potent in the primary assay than many original compound "hits" excluded.[41]

The set of assay activities $\{A|c\}$ included those derived from both top and bottom assay thresholds, so a class could have assay activities obtained from both thresholds for the same assay. Many assay measurements described biological phenotypes, while others only measured compound fluorescence or compound toxicity; any measurement that had no meaningful interpretation was excluded from the analysis.

The *expected* frequency of classes with multiple assay activities was estimated by randomly permuting compounds among the class partitions, such that the size distribution of the classes and the assay measurements associated with each compound were preserved, although the specific compounds belonging to each class were different.

Coincidence Score. For each compound we define an assay outcome, a vector containing binary (0 or 1) values depending on the results of each assay considered significant. To identify compound classes that have multiple-assay phenotypes resulting from a common mechanism (such as fluorescence or a single biological target), a coincidence score was devised that measures the statistical significance of the number of unique assay outcomes of compounds within the compound class. Intuitively, assay phenotypes resulting from a common mechanism of action should be induced by the same compounds within a structural class, and therefore more compounds within the class would be expected to have the same assay outcome (Chart 1). If more than one biological target were bound by the small molecules within a class, the assay outcomes of compounds in the class would be expected to vary because the variation of structural features within that compound class should affect the activity on different biological targets (and assays) differently (Chart 1). Ideally, all compounds within any given class will have identical assay activities, but in practice this is rarely the case even for compounds thought to affect a single target. Our goal is to distinguish compound classes with highly variable assay activities from those that are more uniform.

To measure the degree to which the hits from the different assays $\{A|c\}$ coincide within a compound class, an entropy-based score called the coincidence score was calculated for hits within that class. Based on the multinomial distribution,[48] the coincidence score counts the number of ways assay data can be permuted within a compound class while generating the same observed frequency distribution of assay outcomes. Assay outcomes only attainable a few different ways are considered highly ordered, or equivalently, to have low entropy. In this case, the coincidence score measures the number of ways that hits within a class can be permuted, such that the same pattern of hits is observed within that class. A complete derivation of the coincidence score and its relationship to a more familiar measure of entropy, Shannon's entropy,[49−54] is provided in the Supporting Information sections A and B, but a synopsis is provided here.

Chart 1. Sample Compound Classes Demonstrating Assay Outcome Coincidence[a]

Class with Poorly Coinciding Assay Outcomes

	ASSAY A	ASSAY B	ASSAY C
COMPOUND 1	1	0	1
COMPOUND 2	1	0	0
COMPOUND 3	0	0	0
COMPOUND 4	0	1	0
COMPOUND 5	0	0	0
COMPOUND 6	0	1	1

Class with Highly Coinciding Assay Outcomes

	ASSAY A	ASSAY B	ASSAY C
COMPOUND 1	1	1	1
COMPOUND 2	1	1	1
COMPOUND 3	0	0	0
COMPOUND 4	0	0	0
COMPOUND 5	0	0	0
COMPOUND 6	0	0	0

[a] For purposes of illustration, here is hypothetical assay data for two compound classes. While both classes have the same number of compounds and the same number of assay hits, the degrees to which the assay outcomes coincide within those classes are vastly different. Clearly the "highly coinciding" compound class has more assay activities held in common among compounds than the "poorly coinciding" class. The highly coinciding hits suggest a common mechanism of compound action inducing assay activities A, B, and C.

For any class containing c compounds having the observed frequencies of assay outcomes $\{h_{b1\ b2...bn}\}$ obtained from $n = \#\{A|c\}$ assays (the number of significant assay activities in the class) where $b \in \{1,0\}$, the coincidence is

$$\text{configurations}(c, \{h_{b1b2...bn}\}) = c!/h_{00...0}!\ h_{10...0}!\ ...\ h_{11...1}!$$

$$\text{such that coincidence} = \log(\text{configurations})$$

The coincidence score attains a maximum value (perfectly disordered state) when each compound in the class has a unique assay outcome (every assay outcome $h_{b1\ b2...bn} = 0$ or 1) and a minimum value when all compounds have same assay outcome (some $h_{b1\ b2...bn} = c$).

In the event there are untested compounds within a class for any given assay, the coincidence score is calculated using three possible assay states: hit, nonhit, and untested or a_1, a_0, and a_U. To calculate the coincidence score in the presence of missing data, the number of configurations are first calculated for the three-state case C_3 and then are divided by the number of configurations for the two-state case C_2 having the assay states a_T *tested* and a_U *untested*, where $a_T = a_1 + a_0$. The sets of assay outcomes for a class of *c* compounds tested in *n* assays for the three-state and two-state cases are $\{a_{f1\ f2...fn}\}$ and $\{a_{d1\ d2...dn}\}$, respectively, where $f \in \{1,0,U\}$ and $d \in \{T,U\}$. Similar to the above example,

the three- and two- state configurations are calculated:

$$\text{configurations}_{\text{missing}} = \text{configurations}_{C3}/\text{configurations}_{C2}$$

and

$$\text{coincidence}_{\text{missing}} = \log(\text{configurations}_{\text{missing}}) = \text{coincidence}_{C3} - \text{coincidence}_{C2}$$

where

$$\text{configurations}_{C3}(c, \{a_{f1\ f2\ldots fn}\}) = c!/a_{00\ldots0}!\ a_{10\ldots0}!\ \ldots\ a_{UU\ldots U}!$$

and

$$\text{configurations}_{C2}(c, \{a_{d1\ d2\ldots dn}\}) = c!/a_{TT\ldots T}!a_{UT\ldots T}!\ \ldots\ a_{UU\ldots U}!$$

The coincidence score with missing data reduces to the sum of coincidence scores for complete data for compound subsets that are uniformly tested or uniformly untested against the set of assays, where $f \in \{1,0,U\}$, $d \in \{T,U\}$, and $b \in \{1,0\}$:

$$\text{configurations}_{\text{missing}} = \text{configurations}_{C3}/\text{configurations}_{C2}$$

$$= a_{TT\ldots T}!\ a_{UT\ldots T}!\ \ldots\ a_{UU\ldots U}!/a_{00\ldots0}!\ a_{10\ldots0}!\ \ldots\ a_{UU\ldots U}!$$

$$= a_{TT\ldots T}!/a_{00\ldots0}!\ a_{10\ldots0}!\ \ldots\ a_{11\ldots1}!$$
$$* a_{UT\ldots T}!/a_{U0\ldots0}!\ a_{U1\ldots0}!\ \ldots\ a_{U1\ldots1}!$$
$$* \ldots$$
$$* a_{UU\ldots U}!/a_{UU\ldots U}!$$

$$= \text{configurations}(a_{TT\ldots T}, \{a_{b1\ b2\ldots\ bn}\})$$
$$* \text{configurations}(a_{UT\ldots T}, \{a_{U\ b2\ldots\ bn}\})$$
$$* \ldots$$
$$* \text{configurations}(a_{UU\ldots U}, \{a_{UU\ldots U}\})$$

This result is intuitively appealing because the presence of untested compounds does not increase the value of any term in the coincidence score. If each compound in the class were untested in different assays (giving each compound a unique assay outcome), the $\text{coincidence}_{\text{missing}}$ score would be equal to zero, correctly ignoring the entropy added by missing data and corresponding to an empty data set.

Caution is warranted when comparing coincidence scores derived from classes with different numbers of compounds and different amounts of missing data. This type of comparison is best accomplished using *p*-values associated with the observed coincidence values P(C) to measure significance and normalized coincidence values to measure strength of the association. The raw coincidence score can be normalized by dividing it by the maximum possible entropy for that class.[55]

$$\text{coincidence}_{\text{norm}} = \text{coincidence}/\max(\text{coincidence})$$

The normalized coincidence score was calculated for each class with two or more assay annotations $\#\{A|c\} \geq 2$. While the normalized coincidence score enables direct comparison between the strength of the coincidence in classes of different sizes, it does not measure statistical significance. To accomplish this, a *p*-value was calculated for each class's coincidence as well. Because there was no practical way to calculate this *p*-value exactly (except when $\#\{A|c\}=2$), the observed coincidence score was compared to the coincidence scores obtained by randomly permuting each of the separate assay annotations within a class and then ranking the observed coincidence score to obtain the *p*-value, P(C). Random permutation was accomplished by holding missing data points stationary, so that only hits and nonhits were actually randomized. Random permutation did not preserve compound identity, but it did preserve the number of hits associated with each assay within each compound class (Chart 2).

Chart 2. Calculation of Coincidence and P(C) Scores[a]

Class with Original Assay Activities

	ASSAY A	ASSAY B	ASSAY C
COMPOUND 1	1	1	1
COMPOUND 2	1	1	1
COMPOUND 3	-	0	0
COMPOUND 4	-	0	0
COMPOUND 5	0	0	0
COMPOUND 6	0	0	0

coincidence = log (6! / 2! 2! 2!) – log(6!/ 4! 2!) = log(6)

Example of Same Class with Randomized Assay Activities, Used to Calculate P(C)

	ASSAY A	ASSAY B	ASSAY C
COMPOUND 1	0	1	0
COMPOUND 2	1	0	0
COMPOUND 3	-	0	1
COMPOUND 4	-	0	0
COMPOUND 5	1	0	1
COMPOUND 6	0	1	0

coincidence = log (6! / 2! 1! 1! 1! 1!) - log(6!/ 4! 2!) = log(24)

[a] These two classes illustrate how randomization is performed in order to calculate P(C) scores, when compounds are not tested in all the assays. Notice how both the original assay activities and randomized assay activities contain the same untested compounds indicated by dashes (−), while the distribution of hits (1) and nonhits (0) changes. Upon randomization, the entropy of the class increases, destroying the coincidence originally present among the hits.

For each class with two or more assay annotations $\#\{A|c\}$, 4000 randomizations were performed, and the frequency distribution of random coincidence scores was obtained. The *p*-value, the probability of obtaining the observed coincidence or better, was estimated by obtaining the percentile ranking of the observed coincidence score relative to the randomized scores. If the observed coincidence score was smaller than all 4000 randomized scores, then the class was assigned the default probability, $p < 1/4000$, with the understanding that more randomizations would be required for a more precise estimate. Estimated *p*-values below 0.01 and normalized coincidence scores below 0.10 were considered significant because they were observed to occur infrequently in randomized data sets. The two scores were used in combination because ranking by P(C) favors large classes of compounds with small correlations between assay activities, and the normalized coincidence score favors small classes of compounds having identical assay outcomes. The use of normalized coincidence scoring is necessary because compounds

in a perfectly uniform class, all having identical assay outcomes, will not change after randomization, yielding P(C) = 1.

For every class in each of the 33 cluster sets described above, the respective assay annotations with $\#\{A|c\} >= 2$ were scored for normalized coincidence and P(C) and flagged if they met either the P(C) < 0.01 or the normalized coincidence < 0.10 cutoff.

Robustness of Assay Annotations. A compound's assay-activity annotations must also be reproducible among different cluster sets if they are to be useful descriptions of compounds' biological activity. Because there is a random component in K-modal clustering, cluster sets will rarely match identically, and thus the resulting definitions of compound classes will differ slightly in composition and possibly in their enrichment for various assay activities. To measure the reproducibility of each compound's assay-activity annotations, the entropy-based coincidence scores and their associated *p*-values P(C) were employed. The annotations of the cluster sets containing the same number of clusters K were pooled together and compared, with particular emphasis on the $K = 1360$ cluster sets. If the same assay activities were associated with a compound in every cluster set, then the entropy of those annotations would be low and the P(C) would be significant. The "modified" coincidence score of a given compound's annotations for M assay activities scoring the presence or absence of M assay activities among 3 cluster sets is

$$\text{configurations}(M, \{a_{bbb}\}) = M!/a_{000}!\ a_{001}!\ \ldots\ a_{111}!$$

$$\text{such that coincidence} = \log(\text{configurations})$$

where a_{000} counts the number of times a given activity is seen in a compound's annotations in zero sets, and a_{111} counts the number of times a given activity is seen in a compound's annotations in all 3 sets.

Once the raw coincidence score is calculated, the probability of the observed coincidence score is also estimated by comparing it to the scores of 200 randomized class annotations in the same manner described above (yielding a minimum possible score of 1/200).

Similar comparisons were performed using cluster sets containing different numbers of clusters K (cluster sets having $K = 1360$ compared to cluster sets having $K = 2720$ for example).

Accuracy of Coincidence Score. To assess the ability of the P(C) score to discriminate between compound classes having single target effects and those with side effects, a test set of compound classes with known single target effects was compared to a test set with significant side effects on nonhomologous targets for the Chembridge Diverse Set E library.

To identify compound classes with single-target effects, the 4669 known bioactive compounds from the Chembank[26] Small Molecule Bioactives Database were compared to the 16 320 compounds in the Chembridge Diverse Set E library; 47 known bioactives (and the compound classes containing them) were identified in the Chembridge Diverse Set E library. This list was supplemented with compound classes containing 66 compounds published as bioactive in various assays of this library.[8−21] To identify compound classes with side effects (having multiple targets), compound classes scoring positively in two or more pure protein assays or scoring positively in an assay for fluorescence and another assay employing a nonfluorescent readout were identified.

The compounds identified as bioactive and the classes containing them generally only had a single reported target and were assumed to be specific since many of these compounds were previously evaluated for side effects prior to their publication or were optimized structurally: this does not preclude the possibility that secondary targets of these compounds may be observed at higher compound concentrations or in other biological assays not tested; but for the sake of this comparison, they will be labeled as "single target" classes. The compound classes labeled as having side effects were assumed not to be acting promiscuously by binding multiple targets nonspecifically (which implies a common mechanism of molecular aggregation[56]). Even if the assay activities of compound classes labeled as having "side effects" were the result of measurement error, those activities should still be uncorrelated. Additionally, even if some "single target" classes had weak side effects and some "side effect" classes had some correlated assay activities, differences in assay outcomes should be seen between the two groups as defined here, with the expectation that "single target" classes will have more highly correlated assay phenotypes in general.

The P(C) score was calculated for each of the above compound classes in the test set that had two or more assay annotations. The compound classes having a P(C) score less than the 0.01-threshold were identified, and the false positive and false negative rates of compounds with predicted single-target effects were estimated; classes with zero or one assay activity were not included in these estimates, even if those classes contained known bioactive compounds, because all possible biological activities were not necessarily measured by the assays examined here. The false negative rate is estimated as the percentage of "single target" classes having insignificant P(C) scores (greater than 0.01), P(!D|S). The false positive rate can be calculated using Bayes's Rule, given P(D) as the percentage the total classes passing the 0.01 threshold, P(D|S) as the percentage of "single target" classes passing, and P(D|!S) as the percentage of "side effect" classes passing. The false positive rate P(!S|D) is calculated using the observed values of P(D), P(D|S), and P(D|!S) and the solutions for P(S) and P(!S) (the percentage of classes with single-target effects and the percentage with side effects, respectively)

$$P(!S|D) = P(D|!S)\ [1\text{-}P(S)]/P(D)$$

where

$$P(S) \text{ is obtained by solving}$$
$$P(D) = P(S)\ P(D|S) + P(!S)\ P(D|!S)$$
$$\text{and } P(!S) = 1\text{-}P(S)$$

Compound Mechanism Hypotheses and Validation. The coincidence scores identify compound classes with likely single target relationships but do not provide information about the compound class's mechanism of action or specific biological target without further analysis. To this end, the

highly coinciding assay annotations were interpreted and used to generate hypotheses about the compounds' mechanisms of action.

Having collected the class annotations $\{\{A|c\}\}_{\text{coincidence}}$ that met the normalized coincidence or P(C) cutoffs, each class was reviewed individually using knowledge of the assay protocol and underlying biology, literature searches, and consultation with the original screeners. After review, each compound class's assay annotations were scored in six hypothesis categories: (1) fluorescence, (2) luciferase inhibition, (3) nonspecific transcriptional upregulation, (4) toxicity, (5) cell cycle arrest, and (6) specific biological mechanism. The scores for the first five categories were obtained by simply counting the number of assays that suggested each mode of activity. Toxicity is suggested by low signals in multiple cell-based assays. Compound fluorescence is suggested by classes scoring in multiple assays employing a fluorescent readout, and luciferase inhibition is similarly suggested by compound classes with low signals in multiple assays that use a luciferase reporter. Nonspecific transcriptional upregulation (most likely metal chelation) was suggested by high signals in multiple assays employing artificial promoters. "Cell cycle arrest" was suggested by activity in multiple assays that measure cell cycle arrest directly (a frequently observed phenotype that may arise via many possible targets.) Last, the "specific mechanism" hypotheses were assigned to compound classes scoring in assays that target a specific pathway (often performed on cellular extracts or purified proteins) or classes that had an interesting pattern of assay activity suggestive of a common mechanism. Among the most interesting were classes that implicated a specific biological pathway in the cell cycle arrest phenotype. A single assay annotation could often be assigned to more than one category, and there was no limit on the number of categories for which a single class could score.

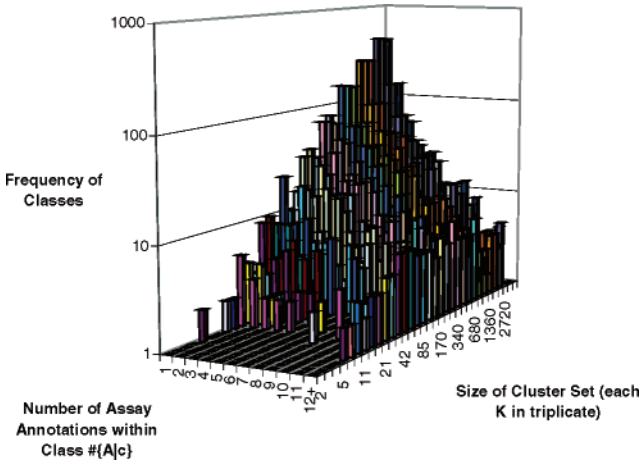
As described earlier, 47 known bioactives from the Chembank[26] Bioactives database were identified in the Chembridge Diverse Set E library, and those bioactives that belonged to classes having multiple assay activities and significant P(C) scores (less than the 0.01 threshold) were found. The known activities of these compounds were then compared to the classes' assay annotations to determine how well the known activities of these compounds can be revealed using only the classes' assay annotations. This provides critical evidence as to whether the compound classes with significant P(C) scores indicate single-target effects and also whether the assay annotations provide solid hypotheses about compound mechanism of action that can motivate specific follow-up experiments.

RESULTS AND DISCUSSION

Numerous compound classes were found to contain multiple assay activities, some of which confirmed known compound activities. In addition, the coincidence score P(C) correctly identified many compound classes as having biological activity resulting primarily from compound action on a single target.

Scoring Compound Classes. Examination of the distribution of assay activities over the compound classes $\{\#\{A|c\}\}$ revealed that multiple assays had a tendency to score positively (having $P_a + < 0.005$) in the same compound classes

Chart 3. Distribution of Assay Annotations over Compound Classes[a]



[a] Multiple assay activities are concentrated in a small number of compound classes. Each color indicates the frequency distribution of classes having multiple assay activities for a given cluster set. More than two activities are seen in compound classes of all sizes in each of the 33 cluster sets shown. Multiactivity classes could be the result of compound action on either a single target or multiple targets.

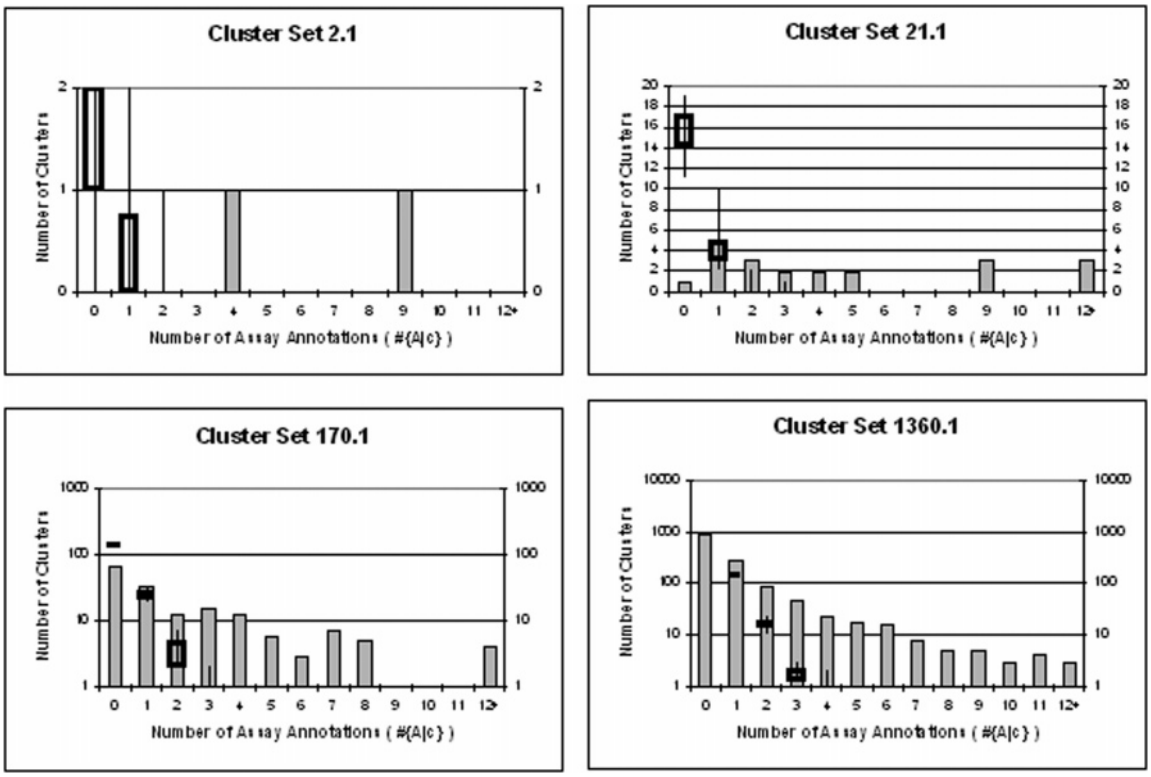
(Chart 3). In addition, after comparing the observed distribution of assay activities to distributions obtained from 10 sets of randomized class assignments (represented by box plots), it was apparent that classes with multiple assay activities occurred far more frequently than would be expected by chance alone (Chart 4). A handful of classes in each cluster set were active in seven or more assays: such an occurrence would be highly unlikely by chance (Chart 4).

The presence of compound classes with multiple assay activities may be the result of chemical modulation of a single biological target affecting each of those assays or a common mechanism for systematic error such as fluorescence. Alternatively, multiple assay activities may result from chemical action on different biological targets. The nature of these numerous compound classes with multiple assay activity annotations will be explored further.

Robustness of Assay Annotations. In addition to their prevalence, most of the compound assay annotations were reproducible and largely independent of cluster set. Assay annotations derived from cluster sets of the same size were far more reproducible than would be expected by chance alone. The following graph shows the average modified P(C) measuring the coincidence of each compound's assay annotations among the three cluster sets of the same size (K=1360 only) as a function of the compound's median number of annotations $\#\{A|c\}$ in the three cluster sets (Chart 5). Notice that the average modified P(C) score is well below the expected 50% value for two or more annotations suggesting that the compounds' class annotations were highly reproducible.

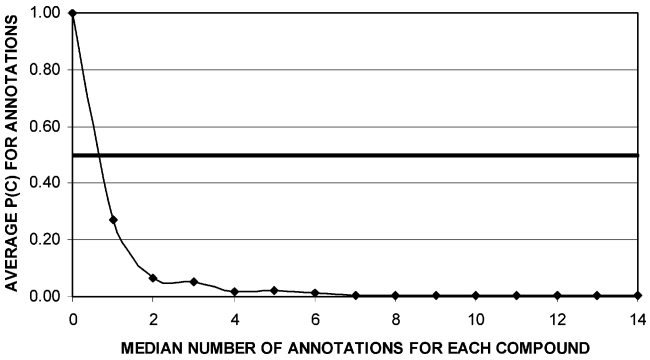
A similar graph compares the modified P(C) scores of the three pairs of cluster sets for $K = 1360$ and $K = 2720$ as a function of the median number of annotations $\#\{A|c\}$ for each compound (see Supporting Information C). The modified P(C) scores are still below the expected value of 50% but clearly show less significance than the same-sized clusters described above. The larger the size differences between cluster sets, the less reproducible the class annotations associated between them. Nonetheless, because the compound's assay annotations are reproducible between different

Chart 4. Distribution of Assay Annotations in Cluster Sets vs Box Plots of Randomized Data[a]



[a] Shown here is the distribution of assay activities from representative cluster sets with $K = 2$, 21, 170, and 1360. Comparing the observed distribution of assay activities for each cluster set (gray bar graphs) to the distribution obtained from randomized data (black box plots), it is clear that compound classes with more than two assay activities occur far more often than predicted by chance. The box plots indicate the assay activity distributions observed in 10 randomized data sets; compound classes with more than two assay activities are highly unlikely by chance alone, in stark contrast to the observed distributions indicated in bar graph form. The frequency of multiple-assay activity compound classes suggests that common mechanisms of action by these compounds may be inducing multiple assay phenotypes.

Chart 5. Median Number of Assay Annotations vs Significance of Their Coincidence P(C) for Each Compound (E=0.50)[a]



[a] The average modified P(C) score used to measure the reproducibility of a compound's assay activity annotations is plotted as a function of the median number of assay annotations assigned to a given compound among the three cluster sets $K = 1360$. A compound's assay activities among the different cluster sets coincide significantly more than expected (as indicated by P(C)<0.50) suggesting that a compound's assay activity annotations are somewhat robust to cluster set choice. Perfect reproducibility of a compound's assay activity annotations is rarely seen among the three cluster sets, but a significant number of a compound's assay annotations will be reproduced in the three cluster sets.

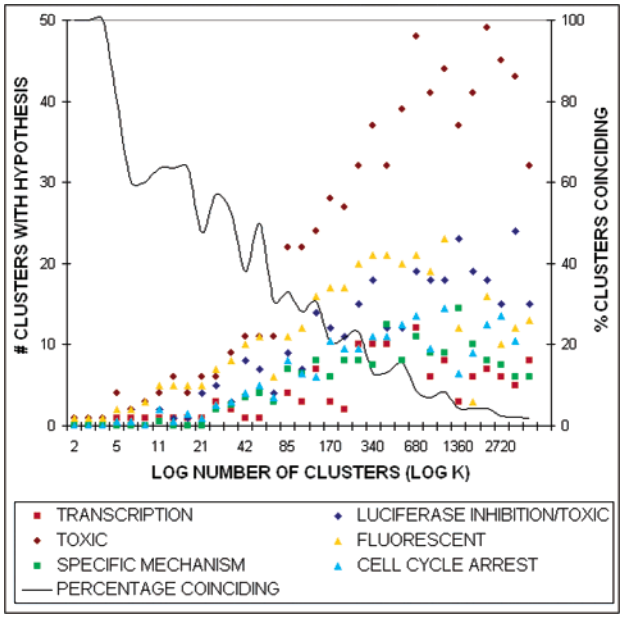
cluster sets, the assay annotations are somewhat robust to cluster choice.

Coincidence Score. The prevalence and reproducibility of the compound classes' multiple assay annotations motivated analysis of each compound class's pattern of assay activity. To this end, the degree to which multiple assay activities coincided within each compound class was measured with the coincidence score. The percentage of compound classes in each cluster set meeting the normalized coincidence (coincidence<0.10) and P(C) (P(C)<0.01) cutoffs is indicated by the black line in this chart (Chart 6). Normalized coincidence scores below 0.10 are seen in smaller classes (with fewer than 6 compounds) for the cluster sets with $K = 1360$, 2720. Significant P(C) scores were observed in both very large (100+ compounds) and small (6−12 compounds) classes suggesting that certain assay activities were correlated with each other for the entire library, not just inside small and medium-sized structural classes. Although the cutoff values for the coincidence and P(C) scores were chosen somewhat arbitrarily, few classes met both thresholds suggesting that the normalized coincidence and P(C) scores were complementary. The coincidence score compared well to a similar statistic, Shannon's Entropy, measuring the significance of the observed number of assay outcomes within a compound class (see Supporting Information B).

The percentage of classes having significant coincidence scores drops steadily with increasing K (Chart 6). This is an expected consequence of having more clusters available to segregate active from inactive compounds. Nonetheless, the prevalence of compound classes with highly coinciding assay activities suggested compound action on single targets.

Accuracy of Coincidence Score. Classes identified as primarily "single target" (because they contained bioactive compounds with single known targets) had significant P(C)

Chart 6. Frequency of Activity Hypothesis vs Number Clusters (K)[a]



[a] The frequency of various hypotheses about a compound class's mechanism is shown as a function of the number of compound classes in each cluster set. The percentage of classes with significant P(C) scores steadily decreases as a function of the number of clusters in the set. Toxicity (brown circles) is the most commonly generated hypothesis, merely indicated by low signal in multiple cell-based assays. Compound classes with potential fluorescent (orange triangles) or luciferase inhibiting (blue diamonds) activities are seen as well: the luciferase inhibition category may also include toxic compounds as discussed in the text. The least common hypothesis generated was nonspecific transcriptional upregulation (red squares), which was suggested by a high signal in multiple assays that utilize artificial promoters. An intermediate number of compound classes were given "cell cycle arrest" (blue triangle) or "specific mechanism" (green square) hypothesis indicating annotations in multiple cell cycle arrest assays or various in vitro assays, respectively. Classes that were annotated with both of these hypotheses were among the most interesting, because they implicated a specific biological pathway in the cell-cycle arrest phenotype.

scores at a far higher rate than classes identified as "side effect" classes, within the test set described above. The following charts examine classes in three cluster sets (K = 1360) and show the statistical significance of the observed number of P(C) < 0.01 in the "single target" classes relative to the "side effect" classes, the observed false negative rate P(!D|S), and the false positive rate P(!S|D) calculated using Bayes's Rule (Chart 7, 8). Note that "single target" classes had significant P(C) scores at a far greater rate than the rest of the test set, with a p-value much less than 0.01. Similarly, the "single target" classes (classes containing known bioactives) had significant P(C) scores at a far greater rate than the rest of the library (p<0.01 for each cluster set). In addition, the false negative rate (percentage of single target classes not passing the threshold) hovered around 55%, while the false positive rate was calculated around 12% using Bayes's rule (implying the predictions of compound classes with single target effects were 88% accurate). Generally speaking, a high false negative rate is tolerable if it means a lower false positive rate, because testing false positives wastes experimental resources while false negatives carry little cost when there is an abundance of active compound classes to pursue. Furthermore, a high false negative rate ("single-target" classes missing the coincidence threshold)

Chart 7. Significance and False Negative Rate of the "Single Target" Classes P(C)[a]

D,Single	Single	D,All	All	p-Value	P(!D|S): False Negative
14	24	16	84	5.56E-08	0.42
10	25	13	78	3.76E-04	0.60
9	23	12	71	1.21E-03	0.61

[a] The sample data set of "single target" and "side effect" compound classes and their P(C) scores were examined for three clusters sets (each indicated by a row in the table). The column marked "All" indicates the total number of classes in the test set having two or more assay activities, and "D,All" indicates those that also had significant P(C) scores. The columns marked "Single" and "D,Single" indicate the number of "single target" classes with two or more assay activities and the number of "single target" classes that also had significant P(C) scores, respectively. The p-values for the observed number of "single target" classes having significant P(C) scores in this test set is well below 1%. The false negative rate (i.e. the fraction of "single target" classes having poorly coinciding assay activities) is also shown.

could be indicative of compounds with previously unknown side effects.

In addition to the calculated 12% false positive rate (or equivalently, the 88% accuracy), it is also encouraging that classes predicted to be single-target contained many known bioactive compounds. Out of the ~50 compound classes predicted to have primarily single target effects in the entire library roughly a dozen (~21%) of those compound classes contained known bioactives. After removing the 12% predicted false positives, this leaves unexplored ~30 compound classes which may contain single target compounds. In some cases, the single target effects will be trivial, such as inhibition of a luciferase assay reporter or unspecified toxicity, but in other compound classes, there could be illuminating chemical genetic relationships relating multiple interesting assay phenotypes to single protein targets. Furthermore, since the cluster sets slightly differ, their compound classes' assay annotations will differ slightly as well, so among the three cluster sets, there could be many more compound classes with unexplored "single target" relationships.

Compound Mechanism Hypotheses and Validation. Each of the compound classes with highly coinciding assay activities was assigned one or more hypotheses about its mechanism of action. The following chart shows the distribution of the hypothesis types including fluorescence, luciferase inhibition, transcriptional upregulation, toxicity, cell cycle arrest, and specific biological mechanism across the classes contained in each cluster set (Chart 6). As noted earlier, the percentage of classes containing annotations passing the normalized coincidence/P(C) thresholds decreased as the number of clusters K increased (Chart 6), but the total number of such classes still increased with increasing K. This is also true for classes annotated with each of the six hypothesis types (Chart 6).

"Toxicity" was the most common hypothesis generated, followed by "fluorescence". Nonspecific "transcriptional upregulation" was observed the least. The "luciferase inhibition" hypothesis was generated frequently, but the frequency of the "luciferase inhibition" hypothesis was likely inflated by the presence of toxic compounds in the library that depressed the signal of the luciferase-based assays which were all cell-based: therefore in Chart 6 we label the "luciferase inhibition" category as "luciferase inhibition/

Chart 8. Estimated False Positive Rates[a]

D	Total	P(D)	P(D | !S)	P(D | S)	P(S)	P(!S |D): False Positive
55	222	0.25	0.03	0.58	0.39	0.08
53	221	0.24	0.06	0.40	0.53	0.11
49	226	0.22	0.06	0.39	0.47	0.15

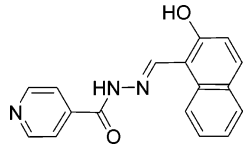
[a] The false positive rate P(!S|D) was calculated using Bayes's rule. For each cluster set, the number of clusters having two or more assay annotations is listed in the column "Total" and the number that also have P(C)<0.01 is listed in the "D" column. P(D) is the fraction of these multiple-assay classes that have significant P(C) scores, and P(D|!S) and P(D|S) indicate the fraction of "side effect" classes and "single target" classes (having multiple assay activities) receiving significant P(C) scores. Using these estimators, the false positive rate P(!S|D) was calculated between 8 and 15% using Bayes's rule. Classes not having two or more assay activities were not included in these estimates, even if those classes contained known bioactive compounds, because not all activities were necessarily assayed.

toxic". The most interesting hypotheses, "specific biological mechanism" and "cell cycle arrest" were among the least common, appearing in 28−35 compound classes in each of the cluster sets with $K = 1360$. The relative frequency observed for each hypothesis type is probably more influenced by the popularity and familiarity of certain assays, particularly those measuring toxicity and cell cycle arrest, than by the actual prevalence of each activity in the entire library. For example, most assays requiring cell growth may suggest compound toxicity if the assay signal is low for that compound. Additionally, numerous assays were developed that measured progression of the cell cycle under various conditions because of the tremendous therapeutic value of *selective* cell cycle arrest in the treatment of various cancers and other diseases. By comparison, only a handful of assays used luciferase reporters (but the frequency of the "luciferase inhibition" hypothesis is still probably inflated due to the reasons noted above.) Furthermore, in almost every case, a class annotation could be associated with two or more of the hypothesis categories, so that the frequency of each hypothesis type does not necessarily represent its actual prevalence in the library. Despite these ambiguities, it is exciting to see that potentially hundreds of compounds distributed over dozens of classes contain interesting cell cycle modulating activities or implicate specific biological mechanisms.

Below are five sample compound classes, each containing known bioactive molecules or FDA-approved drugs, with assay annotations that strongly suggest the known molecular mechanism. These examples inspire optimism about the role of coincidence scoring in future efforts to identify compounds with novel mechanisms of action using primary screening assay data.

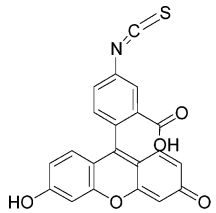
"**NIH (2-Hydroxy-1-naphthylaldehyde isonicotinoyl hydrazone), Iron Chelator**". NIH (also known as 311) (Chart 9) is a cell-permeable iron Fe^{3+} chelator.[57] The antiproliferative properties of NIH in many neoplastic cell lines have been well documented[58,59] and are believed to be a direct result of iron chelation.[59−61] A growing body of evidence suggests that the mechanism of action of NIH's antiproliferative effects may be transcriptional[59,62,63] among other possibilities.[59,64,65] By forming complexes with iron atoms Fe^{3+}, NIH disrupts iron regulatory proteins leading to the transcriptional upregulation of cell cycle inhibitors such as p21(CIP1/WAF1),[59,62] GADD45,[59] and Ndrg1.[63] While the exact relationship between transcriptional upregulation and growth arrest remains controversial, both biological activities are established for many iron chelators.[59,65]

Chart 9. NIH (311 or 2-Hydroxy-1-naphthylaldehyde Isonicotinoyl Hydrazone)[a]

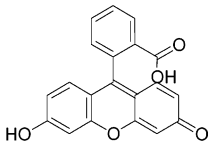


[a] NIH, the known metal chelator, was assigned to a class that appeared to induce nonspecific transcriptional upregulation and inhibit cell growth, confirming its known phenotypes and suggesting metal chelation as a common mechanism of action.

Chart 10. FITC and Fluorescein[a]



FITC



Fluorescein

[a] Fluorescein and its close structural homologue FITC are shown. Prior to the removal of fluorescent compounds from the assay data set, fluorescein was assigned to a compound class scoring in multiple assay employing a fluorescent readout including some using FITC, suggesting compound fluorescence as a common mechanism of action.

Class 1195 of cluster set 1360.2 contained NIH and scored positively in eight assays that suggest nonspecific transcriptional upregulation and low cell growth. Two of these assays indicate upregulation of transcription by an artificial gene reporter (four assays out of the original 48 employ artificial gene reporters that suggest possible transcriptional upregulation as indicated by elevated assay signals). Therefore, a "transcriptional upregulation" hypothesis was assigned to this compound class. Six other assays indicated low cell growth, suggesting a "toxic" or possibly a "cell cycle arrest" hypothesis. Both "transcriptional upregulation" and "toxicity/cell cycle arrest" corroborate the known biological activity of NIH as an iron chelator; other compound classes having these same two phenotypes may also be iron chelators. In total, class 1195 containing 27 compounds scored in eight highly coinciding assays ($P(C) < 0.00025$), suggesting that these phenotypes were induced by a common mechanism, namely iron chelation. Ten of the 27 compounds in this class were not hits in any of the 8 assays, with the activity significantly concentrated in the rest of the compound class as detected by the coincidence score. Classes in $K = 1360$ or 2720 cluster sets containing NIH reliably reproduced the toxic phenotype and less reliably reproduced the transcriptional upregulation phenotype (probably due to the smaller number of assays directly utilizing transcription) with a probability $p < 0.005$.

Fluorescein/FITC, Fluorescence. Fluorescein (Chart 10) is an FDA-approved drug used in fluorescein angiography, a technique used to visualize the blood vessels at the back of the eye to diagnose retinal abnormalities.[66] Fluorescein isothiocyanate (FITC) (Chart 10) is a closely related molecule used to label antibodies and proteins in many biological assays, such as fluorescence polarization assays which measure protein binding.[67,68] Six out of the 48 ICCB assays studied employ FITC or GFP reporters, both of which emit

at the 520 nm peak wavelength, one control assay directly measured compound fluorescence at 520 nm, and two other assays employ reporters that emit at the nearby 580 nm peak wavelength. With this in mind, it was not surprising that a class of compounds containing fluorescein emerged from the above analysis with a "fluorescent" hypothesis (prior to the removal of fluorescent compounds from the assay data set).

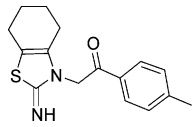
Class 1076 of cluster set 1360.1 containing fluorescein scored in 10 highly coinciding assays. Eight of these assays employ a fluorescent readout and suggest that these compounds score as false positives due to compound fluorescence (this phenomenon was observed before fluorescent compounds were deleted from the assays as described above.) The two other assays, not shown, suggested other activity. Although two other assays scored in this class, the "fluorescent" hypothesis was much more compelling because it was inspired by 8 assays. Additionally as noted above, no class would be predicted to have 8 or more assay activities by chance alone. In total, the 10 assay activities in class 1076 containing 10 compounds coincided with a P(C) estimated below 0.00025, suggesting a single mechanism (in this case compound fluorescence).

The presence of two other assay activities not obviously connected to compound fluorescence did not significantly diminish the significance of the coincidence score. One of these assay activities is directly attributable to inhibition of protein arginine methyltransferase, since its known inhibitors AMI-4, AMI-5, and AMI-6[20] are present in this compound class. However, removal of the most fluorescent compounds from all but one of the assays employing a fluorescent readout left the remaining assay activities (including one assay directly measuring compound fluorescence) with insignificant P(C) scores correctly suggesting that the arginine methyltransferase and fluorescence activities were uncorrelated and the result of different mechanisms.

This class of fluorescein homologues illustrates both the strength and weakness of the coincidence score. While the significant P(C) score seems to verify the presence of compound action by a single mechanism inducing many of the assay activities, it does not guarantee that all observed assay activities arise via that mechanism given the protein arginine methyltransferase inhibitors present. Nonetheless, the presence of multiple correlated assay phenotypes does significantly inform the search for the single target or mechanism, as it does for this case of compound fluorescence, and this could help focus experimental efforts that seek to optimize the compounds in this class structurally in order to remove undesired side effects. At some point, too many side effects will render the P(C) score insignificant, deprioritizing those compound classes in follow-up studies. Moreover, when a given assay activity (fluorescence) is not so over-represented, this problem would generally not occur (which is why the analysis was repeated after removing fluorescent compounds.) Other classes containing fluorescein in other large-K cluster sets (K=1360 or 2720) showed similar assay coincidences and had significantly reproducible class annotations according to the modified coincidence scores comparing cluster sets with the same K and different K ($p<0.005$).

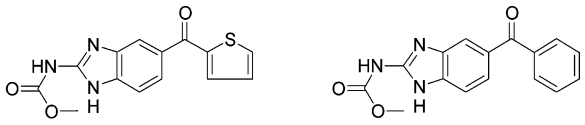
Pifithrin-α, p53, and Luciferase Inhibitor. Pifithrin-α (Chart 11) is a known apoptosis inhibitor, which acts by inhibiting p53, a known tumor suppressor.[69,70] Pifithrin-α has previously been evaluated as a potential treatment for side effects of cancer radiation treatment and chemotherapy, in hopes that it could prevent the death of healthy cells.[69,70] It was also recently reported that pifithrin-α could inhibit firefly luciferase and protein constructs containing firefly luciferase which are commonly used as assay reporters.[71] The recent report cautioned about the possibility of pifithrin-α causing false positives in biological assays using firefly luciferase.

Chart 11. Pifithrin-α[a]



[a] Pifithrin-α is a known p53 inhibitor that has also been reported to inhibit firefly luciferase. This compound was assigned to a class that appeared to inhibit luciferase reporter-based assays and cause growth defects. Its coincidence score P(C) did not meet the 0.01 threshold, correctly suggesting that these two phenotypes were induced by different biological targets.

Chart 12. Nocodazole and Mebendazole[a]

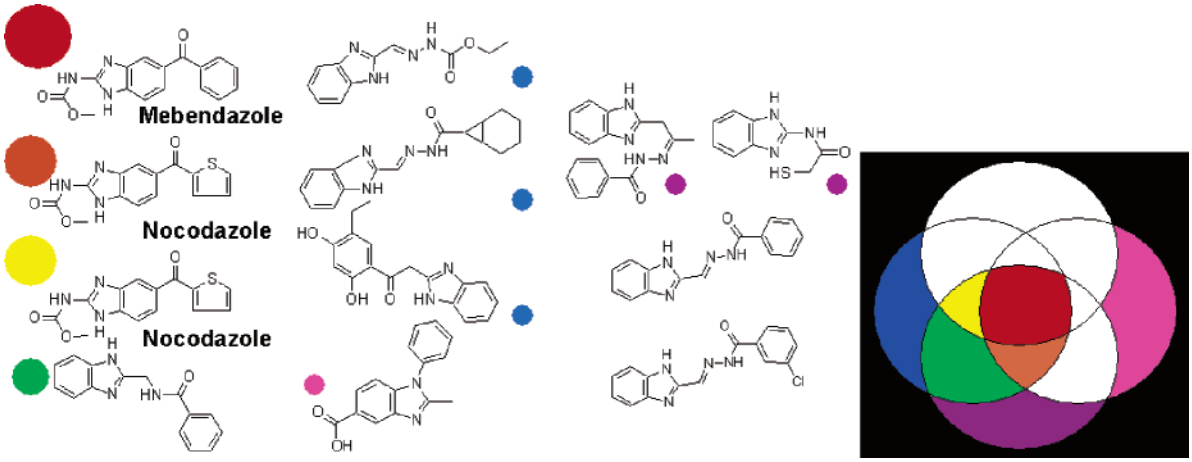


Nocodazole Mebendazole

[a] The known drug mebendazole and the bioactive nocodazole both are known microtubule destabilizers, and both compounds were assigned to a class that appeared to induce cell cycle arrest (specifically in mitosis) and disrupt endosome trafficking. The coincidence score suggested a common mechanism of action inducing these assay phenotypes. The comparison of this class's assay annotations to another class containing known microtubule destabilizers suggests that this pattern of assay annotations may be characteristic of microtubule destabilizers in general, arguing for the discernability of a specific hypothesis about a compound's mechanism of action from its assay activities alone.

Class 830 of cluster set 1360.2 contains pifithrin-α and scored positively in two assays that used a luciferase assay reporter and two assays that indicate growth defects. The low signal in two of these assays suggested either toxicity or direct inhibition of luciferase, while one indicated cellular toxicity and another identified an observed growth defect. The compound class was ultimately assigned both a "luciferase inhibition" hypothesis and a "toxic" hypothesis. In total, class 830 containing 12 compounds scored in four assays that coincided with a probability P(C) = 0.2742, failing to meet the 0.01 threshold and suggesting the presence of side effects. Appropriately, the hits from the two luciferase reporter assays, considered independently of the other two assays, coincided well with a P(C) = 0.00125, consistent with their common mechanism. In this case, both activity hypotheses appear to be corroborated by the two known activities of this compound class, luciferase inhibition and growth irregularity (p53 inhibition). Furthermore, the hypotheses were assigned an insignificant P(C) score, correctly suggesting that pifithrin-α activities were the result of action on different biological targets. The luciferase inhibition and toxicity phenotypes were also significantly reproducible ($p<0.005$) in classes from the K = 1360 or 2720 cluster sets containing pifithrin-α. It is worth noting that while pifithrin-α has at least one side effect, the luciferase inhibition side effect is a tolerable one with the benefit of hindsight, but one could imagine the confusion created by pursuing

Chart 13. Venn Diagram Showing the Highly Coinciding Assay Activities of Compound Class Containing Nocodazole and Mebendazole[a]



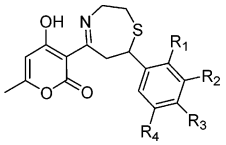
[a] The compounds in this class scored in four assays measuring mitotic arrest, inhibition of endocytosis, BrdU−incorporation (cell growth), and mitotic arrest (listed clockwise starting at the top). The hits across these four assays are highly coinciding with a significant P(C) score estimated around 0.014. Larger circles next to a compound structure indicate a compound that scored in more assays, and smaller circles indicate a compound that scored in fewer assays: Mebendazole has the largest circle because it scored in all four assays, and both copies of nocodazole have medium-sized circles because they scored in three assays. The colors of the circles correspond to the colored regions of the 4-way Venn diagram indicating the hit compounds held in common by each assay (each assay is represented by a ring). Notice how some compounds have multiple assay activities, while others (including nine not shown) have no observed activities, suggesting correlation among the assay hits. If the hits within this compound class were randomly assigned, the corresponding chart would have more shaded regions in the perimeter rather than the center as seen here.

pifithrin-α as lead with only the benefit of two sets of unrelated assay phenotypes.

Nocodazole and Mebendazole, Microtubule Destabilizers. Nocodazole and mebendazole (Chart 12) are both known microtubule destabilizers. Mebendazole is an FDA-approved drug used to kill gastrointestinal parasitic worms.[72,73] By destabilizing microtubules, both compounds arrest cells in mitosis and yield relevant therapeutic value.[74−78] Nocodazole and mebendazole's ability to block vesicle trafficking inside the cell has also been noted.[76,79]

Class 380 of cluster set 2720.3 contains both nocodazole and mebendazole and scored positively in two assays that suggest mitotic arrest, one that suggested growth inhibition and another assay that suggested disruption of endosome trafficking (Chart 13). The "cell-cycle arrest" phenotype (specifically in mitosis) corroborates the known arresting properties of nocodazole and mebendazole. The disruption of endosome trafficking also corroborates nocodazole's known ability to disrupt vesicle trafficking. Arguably, the combination of mitotic arrest and the disruption of endosome trafficking, which relies on the microtubule network, strongly suggests the "specific biological mechanism" hypothesis that the spindle checkpoint is being triggered by possible microtubule destabilization;[77] therefore, the known target of these compounds could potentially be predicted by interpreting the assay phenotypes alone. In total, class 380 containing 21 compounds (nocodazole in duplicate plus one copy of mebendazole) scored in 4 assays that coincided with a probability of P(C) = 0.014, suggesting a single mechanism (in this case microtubule destabilization.) Eleven out of the 21 compounds in this class contained no hits in any of the four assays with all the activity significantly concentrated in the remainder of the class, as detected by the significant P(C) score. These assay annotations, particularly of the mitotic arrest phenotype, were also significantly reproduced in other classes containing either nocodazole or mebendazole according to the modified coincidence scores discussed above ($p < 0.005$) for the large-K cluster sets.

Chart 14. Consensus Structure of Microtubule Destabilizer "Class 2"[a]



2a R_1=OMe, R_2=OMe, R_3=OMe, R_4=H
2b R_1=OMe, R_2=H, R_3=OMe, R_4=H
2c R_1=OMe, R_2=H, R_3=H, R_4=OMe
2d R_1=OMe, R_2=H, R_3=H, R_4=H
2e R_1=OH, R_2=OMe, R_3=H, R_4=H
2f R_1=H, R_2=OCH_2O-, R_3=OCH_2O-, R_4=H
2g R_1=H, R_2=OMe, R_3=OMe, R_4=OMe
2h R_1=OMe, R_2=H, R_3=OMe, R_4=OMe
2i R_1=OMe, R_2=OMe, R_3=H, R_4=H
2j R_1=H, R_2=H, R_3=OMe, R_4=H
2k R_1=H, R_2=OMe, R_3=OH, R_4=H
2l R_1=H, R_2=OH, R_3=H, R_4=H
2m R_1=OH, R_2=Cl, R_3=H, R_4=Cl
2n R_1=H, R_2=H, R_3=Cl, R_4=H
2o R_1=H, R_2=Br, R_3=H, R_4=H
2p R_1=H, R_2=H, R_3=isopropyl, R_4=H
2q R_1=H, R_2=H, R_3=Me, R_4=H

[a] "Class 2" has published activity inducing mitotic arrest by destabilizing microtubules. These compounds were assigned to a compound class that scored in assays measuring mitotic arrest, disruption of endosome trafficking, and even one assay measuring involvement of the spindle checkpoint. These assay activities strongly suggest class 2's known microtubule destabilizing activity and closely parallel the assay activities assigned to another class containing other known microtubule destabilizers, nocodazole and mebendazole.

Microtubule Destabilizer "Class 2". A recent publication identified a class of 17 microtubule destabilizers[8] (Chart 14). This class was originally identified by a high-throughput TG-3 cytoblot assay of A549 lung epithelial cells for a protein nucleolin that is specifically phosphorylated during mitosis. In addition to inducing mitotic arrest, this class was subsequently classified as microtubule destabilizing in a follow-up assay using purified bovine brain tubulin, supporting the conclusion that this class of compounds targeted tubulin α/β directly, triggering the spindle checkpoint and inducing mitotic arrest.[8]

Class 229 of cluster set 1360.2 contains 14 of the ICCB microtubule destabilizers listed above, which score in 5 assays suggesting mitotic arrest, 1 assay suggesting endosome trafficking disruption, and another suggesting involvement of the spindle checkpoint. This class also scored positively in two assays that suggested possible "toxicity" and one other assay with no mechanistic interpretation. In total, class 229, containing 29 compounds scored in 10 assays that coincided

with an estimated P(C) = 0.00025, suggesting a single mechanism (microtubule destabilization). These annotations were significantly reproduced in other classes from cluster sets with $K = 1360$ or 2720 that contained any of the 14 ICCB microtubule destabilizers ($p < 0.005$) according to the modified coincidence scores.

Impressively, this class scored in all the assays in which class 380 (containing nocodazole and mebendazole) scored. Not only does this suggest the same "cell-cycle (mitotic) arrest" hypothesis due to the microtubule destabilization as for nocodazole and mebendazole, but also these two examples argue strongly for the generality of a very specific hypothesis about a compound's biological mechanism obtained by only using primary assay data. Even if the pattern of assay activities is not an intuitive one, such that no obvious biological mechanism is suggested, if one class's assay annotations match the annotations of a class with a known mechanism, then one could infer the same potential biological mechanism for each class.

CONCLUSIONS

Selection of leads for follow-up from phenotypic assays can be enhanced by examination of those leads' activity in other primary assays. Using the coincidence score to evaluate the pattern of assay activities within a compound class can identify the best leads—those primarily having single target effects—for follow-up studies. Additionally, the pattern of assay activities of a given compound class can inform the search for that class's compound target or mechanism and better classify its phenotype.

It is important to note that generating specific target or mechanism hypotheses from compound classes' assay annotations still requires the same biological intuition and literature review used in conventional low-throughput chemical genetics experiments. However, hypothesis generation may be enhanced by coincidence scoring because it can save time otherwise spent testing intractable leads, particularly those with side effects on nonhomologous targets, and the assay outcomes of a compound class can aid the selection of follow-up assays even when a specific target is not suggested for the compound class.

The use of coincidence scoring may be further refined as the quality of high-throughput assay data improves (through the use of compound IC50s for example). Coincidence scoring could also be used to evaluate the assay activities of small combinatorial libraries to identify side effects of selected leads and their structural homologues. Furthermore, coincidence scoring may have applications in the high-content screening of individual compounds in order to measure the degree to which a compound's dose−response curves in different assays coincide as part of efforts to structurally optimize compounds and remove side effects.

Target identification remains a daunting problem, and target identification through assay annotations alone appears to be limited to special cases. Underneath the goal of target identification is the need to identify compounds with interesting combinations of assay phenotypes resulting from chemical action on a single target. To this end, coincidence scoring may have a role.

ACKNOWLEDGMENT

We thank Jason McIntosh and Caroline Shamu of the Broad Institute Chemical Biology Program (BCB) and the Harvard Institute of Chemistry and Cell Biology, Jeremy Muhlich of the BCB and the Massachusetts Institute of Technology, Nicola Tolliday of the BCB, Gabriel Berriz of the Harvard Medical School, and Tudor Oprea of the University of New Mexico for their consultations. We thank the National Cancer Institute's Initiative for Chemical Genetics, which supported all screening and informatic (ChemBank) activity that served as the basis for the current study. S.L.S. is an Investigator at the Howard Hughes Medical Institute at the Department of Chemistry and Chemical Biology, Harvard University. F.P.R was supported in part by the Keck Foundation and by NIH Grants R01 HG0017115, R01 HG003224, and U01 HL81341.

Supporting Information Available: (A) Derivation of the coincidence score, (B) comparison of coincidence scoring and Shannon's entropy, (C) median number of assay annotations vs significance of the coincidence of paired annotations (E=0.50), (D) references, and (E) percentage of assay hits before and after filtering by class scores. This information is available free of charge via the Internet at http://pubs.acs.org.′

REFERENCES AND NOTES

(1) Patel, D. V.; Gordon, E. M. Applications of small molecule combinatorial chemistry to drug discovery. *Drug Discovery Today* **1996**, *1* (4), 134−144.
(2) Baum, R. M. Combinatorial approaches provide fresh leads for medicinal chemistry. *Chem. Eng. News.* **1994**, *72* (6), 20−26.
(3) Schreiber, S. L. Target-oriented and diversity-oriented organic synthesis in drug discovery. *Science* **2000**, *287* (5460), 1964−9.
(4) Caron, P. R.; Mullican, M. D.; Mashal, R. D.; Wilson, K. P.; Su, M. S.; Murcko, M. A. Chemogenomic approaches to drug discovery. *Curr. Opin. Chem. Biol.* **2001**, *5* (4), 464−70.
(5) Bredel, M.; Jacoby, E. Chemogenomics: an emerging strategy for rapid target and drug discovery. *Nat. Rev. Genet.* **2004**, *5* (4), 262−75.
(6) Agrafiotis, D. K.; Myslik, J. C.; Salemme, F. R. Advances in diversity profiling and combinatorial series design. *Mol. Divers.* **1998**, *4* (1), 1−22.
(7) Agrafiotis, D. K.; Lobanov, V. S.; Salemme, F. R. Combinatorial informatics in the post-genomics ERA. *Nat. Rev. Drug Discovery* **2002**, *1* (5), 337−46.
(8) Haggarty, S. J.; Mayer, T. U.; Miyamoto, D. T.; Fathi, R.; King, R. W.; Mitchison, T. J.; Schreiber, S. L. Dissecting cellular processes using small molecules: identification of colchicine-like, taxol-like and other small molecules that perturb mitosis. *Chem. Biol.* **2000**, *7* (4), 275−86.
(9) Koeller, K. M.; Haggarty, S. J.; Perkins, B. D.; Leykin, I.; Wong, J. C.; Kao, M. C.; Schreiber, S. L. Chemical genetic modifier screens: small molecule trichostatin suppressors as probes of intracellular histone and tubulin acetylation. *Chem. Biol.* **2003**, *10* (5), 397−410.
(10) Mayer, T. U.; Kapoor, T. M.; Haggarty, S. J.; King, R. W.; Schreiber, S. L.; Mitchison, T. J. Small molecule inhibitor of mitotic spindle bipolarity identified in a phenotype-based screen. *Science* **1999**, *286* (5441), 971−4.
(11) Stockwell, B. R.; Haggarty, S. J.; Schreiber, S. L. High-throughput screening of small molecules in miniaturized mammalian cell-based assays involving post-translational modifications. *Chem. Biol.* **1999**, *6* (2), 71−83.
(12) Kim, T.; Kim, T. Y.; Lee, W. G.; Yim, J.; Kim, T. K. Signaling pathways to the assembly of an interferon-beta enhanceosome. Chemical genetic studies with a small molecule. *J. Biol. Chem.* **2000**, *275* (22), 16910−7.
(13) Nieland, T. J.; Penman, M.; Dori, L.; Krieger, M.; Kirchhausen, T. Discovery of chemical inhibitors of the selective transfer of lipids mediated by the HDL receptor SR-BI. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (24), 15422−7.
(14) Fantin, V. R.; Berardi, M. J.; Scorrano, L.; Korsmeyer, S. J.; Leder, P. A novel mitochondriotoxic small molecule that selectively inhibits tumor cell growth. *Cancer Cell* **2002**, *2* (1), 29−42.
(15) Kao, R. Y.; Jenkins, J. L.; Olson, K. A.; Key, M. E.; Fett, J. W.; Shapiro, R. A small-molecule inhibitor of the ribonucleolytic activity

of human angiogenin that possesses antitumor activity. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (15), 10066−71.
(16) Yarrow, J. C.; Feng, Y.; Perlman, Z. E.; Kirchhausen, T.; Mitchison, T. J. Phenotypic screening of small molecule libraries by high throughput cell imaging. *Comb. Chem. High Throughput Screening* **2003**, *6* (4), 279−86.
(17) Feng, Y.; Yu, S.; Lasell, T. K.; Jadhav, A. P.; Macia, E.; Chardin, P.; Melancon, P.; Roth, M.; Mitchison, T.; Kirchhausen, T. Exo1: a new chemical inhibitor of the exocytic pathway. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (11), 6469−74.
(18) Straight, A. F.; Cheung, A.; Limouze, J.; Chen, I.; Westwood, N. J.; Sellers, J. R.; Mitchison, T. J. Dissecting temporal and spatial control of cytokinesis with a myosin II Inhibitor. *Science* **2003**, *299* (5613), 1743−7.
(19) Peterson, J. R.; Bickford, L. C.; Morgan, D.; Kim, A. S.; Ouerfelli, O.; Kirschner, M. W.; Rosen, M. K. Chemical inhibition of N-WASP by stabilization of a native autoinhibited conformation. *Nat. Struct. Mol. Biol*. **2004**, *11* (8), 747−55.
(20) Cheng, D.; Yadav, N.; King, R. W.; Swanson, M. S.; Weinstein, E. J.; Bedford, M. T. Small molecule regulators of protein arginine methyltransferases. *J. Biol. Chem.* **2004**, *279* (23), 23892−9.
(21) Haggarty, S. J.; Koeller, K. M.; Wong, J. C.; Grozinger, C. M.; Schreiber, S. L. Domain-selective small-molecule inhibitor of histone deacetylase 6 (HDAC6)-mediated tubulin deacetylation. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (8), 4389−94.
(22) Yarrow, J. C.; Totsukawa, G.; Charras, G. T.; Mitchison, T. J. Screening for cell migration inhibitors via automated microscopy reveals a Rho-kinase inhibitor. *Chem. Biol.* **2005**, *12* (3), 385−95.
(23) Boyce, M.; Bryant, K. F.; Jousse, C.; Long, K.; Harding, H. P.; Scheuner, D.; Kaufman, R. J.; Ma, D.; Coen, D. M.; Ron, D.; Yuan, J. A selective inhibitor of eIF2alpha dephosphorylation protects cells from ER stress. *Science* **2005**, *307* (5711), 935−9.
(24) Tochtrop, G. P.; King, R. W. Target identification strategies in chemical genetics. *Comb. Chem. High Throughput Screening* **2004**, *7* (7), 677−88.
(25) Luesch, H.; Wu, T. Y.; Ren, P.; Gray, N. S.; Schultz, P. G.; Supek, F. A genome-wide overexpression screen in yeast for small-molecule target identification. *Chem. Biol.* **2005**, *12* (1), 55−63.
(26) Strausberg, R. L.; Schreiber, S. L. From knowing to controlling: a path from genomics to drugs using small molecule probes. *Science* **2003**, *300* (5617), 294−5.
(27) Degterev, A.; Lugovskoy, A.; Cardone, M.; Mulley, B.; Wagner, G.; Mitchison, T.; Yuan, J. Identification of small-molecule inhibitors of interaction between the BH3 domain and Bcl-xL. *Nat. Cell Biol.* **2001**, *3* (2), 173−82.
(28) Young, S. S.; Hawkins, D. M. Analysis of a 2(9) full factorial chemical library. *J. Med. Chem.* **1995**, *38* (14), 2784−8.
(29) Rusinko, A., III; Farmen, M. W.; Lambert, C. G.; Brown, P. L.; Young, S. S. Analysis of a large structure/biological activity data set using recursive partitioning. *J. Chem. Inf. Comput. Sci.* **1999**, *39* (6), 1017−26.
(30) Rusinko, A., III; Young, S. S.; Drewry, D. H.; Gerritz, S. W. Optimization of focused chemical libraries using recursive partitioning. *Comb. Chem. High Throughput Screening* **2002**, *5* (2), 125−33.
(31) van Rhee, A. M. Use of recursion forests in the sequential screening process: consensus selection by multiple recursion trees. *J. Chem. Inf. Comput. Sci.* **2003**, *43* (3), 941−8.
(32) Lagunin, A.; Stepanchikova, A.; Filimonov, D.; Poroikov, V. PASS: prediction of activity spectra for biologically active substances. *Bioinformatics* **2000**, *16* (8), 747−8.
(33) Labute, P. Binary QSAR: a new method for the determination of quantitative structure activity relationships. *Pac. Symp. Biocomput.* **1999**, 444−55.
(34) Labute, P.; Nilar, S.; Williams, C. A probabilistic approach to high throughput drug discovery. *Comb. Chem. High Throughput Screening* **2002**, *5* (2), 135−45.
(35) Glick, M.; Klon, A. E.; Acklin, P.; Davies, J. W. Enrichment of extremely noisy high-throughput screening data using a naive Bayes classifier. *J. Biomol. Screen* **2004**, *9* (1), 32−6.
(36) Stockfisch, T. P. Partially unified multiple property recursive partitioning (PUMP−RP): a new method for predicting and understanding drug selectivity. *J. Chem. Inf. Comput. Sci.* **2003**, *43* (5), 1608−13.
(37) Brown, R. D.; Martin, Y. C. Use of Structure−Activity Data To Compare Structure-Based Clustering Methods and Descriptors for Use in Compound Selection. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, (3), 572−584.
(38) Brown, R. D.; Martin, Y. C. The Information Content of 2D and 3D Structural Descriptors Relevant to Ligand−Receptor Binding. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 1−9.
(39) Menard, P. R.; Lewis, R. A.; Mason, J. S. Rational Screening Set Design and Compound Selection: Cascaded Clustering. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 497−505.
(40) Wild, D. J.; Blankley, C. J. Comparison of 2D fingerprint types and hierarchy level selection methods for structural grouping using Ward's clustering. *J. Chem. Inf. Comput. Sci.* **2000**, *40* (1), 155−62.
(41) Klekota, J.; Brauner, E.; Schreiber, S. L. Identifying Biologically Active Compound Classes Using Phenotypic Screening Data and Sampling Statistics. *J. Chem. Inf. Model.* **2005**, *45* (6), 1824−1836.
(42) Doman, T. N.; Cibulskis, J. M.; Cibulskis, M. J.; McCray, P. D.; Spangler, D. P. Algorithm5: A Technique for Fuzzy Similarity Clustering of Chemical Inventories. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 1195−1204.
(43) MacCuish, J.; Nicolaou, C.; MacCuish, N. E. Ties in proximity and clustering compounds. *J. Chem. Inf. Comput. Sci.* **2001**, *41* (1), 134−46.
(44) Martin, Y. C.; Kofron, J. L.; Traphagen, L. M. Do structurally similar molecules have similar biological activity? *J. Med. Chem.* **2002**, *45* (19), 4350−8.
(45) Chemistry-Smiles-0.13: Smile Parser, Cpan.org: 2003.
(46) The C Clustering Library, The University of Tokyo, Institute of Medical Science, Human Genome Center, 2003.
(47) Tanimoto, T. T. *An Elementary Mathematical Theory of Classification and Prediction*; IBM Internal Report; 1958.
(48) Butler, R. W.; Sutton, R. K. Saddlepoint Approximation for Multivariate Cumulative Distribution Functions and Probability Computations in Sampling Theory and Outlier Testing. *J. Am. Stat. Assoc.* **1998**, *93* (442), 596−604.
(49) Shannon, C. E. The mathematical theory of communication. 1963. *MD Comput.* **1997**, *14* (4), 306−17.
(50) Cover, T. M.; Thomas, J. A. *Elements of Information Theory*; John Wiley & Sons: New York, 1991; pp 12−49, 268−270.
(51) Nemenman, I. *Inference of Entropies of Discrete Random Variables with Unknown Cardinalities*; Technical Report 2002; NSF-ITP-02-52; KITP, UCSB.
(52) Nemenman, I.; Shafee, F.; Bialek, W. Entropy and inference, revisited. In *Advances in Neural Information Processing Systems 14*; MIT Press: Cambridge, MA, 2002.
(53) Steck, H.; Jaakkola, T. *Bias-Corrected Bootstrap and Model Uncertainty. Advances in Neural Information Processing Systems*; 2003; p 16.
(54) Paninski, L. Estimation of Entropy and Mutual Information. *Neural Comput.* **2003**, *15*, 1191−1253.
(55) Valdar, W. S. Scoring residue conservation. *Proteins* **2002**, *48* (2), 227−41.
(56) McGovern, S. L.; Caselli, E.; Grigorieff, N.; Shoichet, B. K. A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *J. Med. Chem.* **2002**, *45* (8), 1712−22.
(57) Richardson, D. R. Mobilization of iron from neoplastic cells by some iron chelators is an energy-dependent process. *Biochim. Biophys. Acta* **1997**, *1320* (1), 45−57.
(58) Richardson, D. R. Iron chelators as therapeutic agents for the treatment of cancer. *Crit. Rev. Oncol Hematol.* **2002**, *42* (3), 267−81.
(59) Darnell, G.; Richardson, D. R. The potential of iron chelators of the pyridoxal isonicotinoyl hydrazone class as effective antiproliferative agents III: the effect of the ligands on molecular targets involved in proliferation. *Blood* **1999**, *94* ((2), 781−92.
(60) Buss, J. L.; Neuzil, J.; Gellert, N.; Weber, C.; Ponka, P. Pyridoxal isonicotinoyl hydrazone analogs induce apoptosis in hematopoietic cells due to their iron-chelating properties. *Biochem. Pharmacol.* **2003**, *65* (2), 161−72.
(61) Chaston, T. B.; Lovejoy, D. B.; Watts, R. N.; Richardson, D. R. Examination of the antiproliferative activity of iron chelators: multiple cellular targets and the different mechanism of action of triapine compared with desferrioxamine and the potent pyridoxal isonicotinoyl hydrazone analogue 311. *Clin. Cancer Res.* **2003**, *9* (1), 402−14.
(62) Liang, S. X.; Richardson, D. R. 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* **2003**, *24* (10), 1601−14.
(63) Le, N. T.; Richardson, D. R. 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* **2004**, *104* (9), 2967−75.
(64) Green, D. A.; Antholine, W. E.; Wong, S. J.; Richardson, D. R.; Chitambar, C. R. Inhibition of malignant cell growth by 311, a novel iron chelator of the pyridoxal isonicotinoyl hydrazone class: effect on the R2 subunit of ribonucleotide reductase. *Clin. Cancer Res.* **2001**, *7* (11), 3574−9.
(65) Le, N. T.; Richardson, D. R. Potent iron chelators increase the mRNA levels of the universal cyclin-dependent kinase inhibitor p21(CIP1/WAF1), but paradoxically inhibit its translation: a potential mechanism of cell cycle dysregulation. *Carcinogenesis* **2003**, *24* (6), 1045−58.
(66) McMahon, R. T.; Tso, M. O.; McLean, I. W. Histologic localization of sodium fluorescein in human ocular tissues. *Am. J. Ophthalmol.* **1975**, *80* (6), 1058−65.

(67) Cars, O.; Forsum, U.; Hjelm, E. New immunofluorescence method for the identification of group a, b, c, e, and g streptococci. *Acta Pathol. Microbiol. Scand. B* **1975**, *83* (2), 145−52.
(68) Freeman, D. A.; Crissman, H. A. Evaluation of six fluorescent protein stains for use in flow microfluorometry. *Stain Technol.* **1975**, *50* (4), 279−84.
(69) Komarov, P. G.; Komarova, E. A.; Kondratov, R. V.; Christov-Tselkov, K.; Coon, J. S.; Chernov, M. V.; Gudkov, A. V. A chemical inhibitor of p53 that protects mice from the side effects of cancer therapy. *Science* **1999**, *285* (5434), 1733−7.
(70) Komarova, E. A.; Gudkov, A. V. Suppression of p53: a new approach to overcome side effects of antitumor therapy. *Biochemistry (Moscow)* **2000**, *65* (1), 41−8.
(71) Rocha, S.; Campbell, K. J.; Roche, K. C.; Perkins, N. D. The p53-inhibitor pifithrin-alpha inhibits firefly luciferase activity in vivo and in vitro. *BMC Mol. Biol.* **2003**, *4* (1), 9.
(72) Seo, B. S.; Cho, S. Y.; Kang, S. Y.; Chai, J. Y. Anthelmintic Efficacy Of Methyl-5-Benzoylbenzimidazole-2-Carbamate(Mebendazole) Against Multiple Helminthic Infections. *Kisaengchunghak Chapchi* **1977**, *15* (1), 11−16.
(73) Cabrera, B. D.; Valdez, E. V.; Go, T. G. Clinical trials of broad spectrum anthelmintics against soil-transmitted helminthiasis. *Southeast Asian J. Trop. Med. Public Health* **1980**, *11* (4), 502−6.
(74) Holden, H. E.; Crider, P. A.; Wahrenberg, M. G. Mitotic arrest by benzimidazole analogues in human lymphocyte cultures. *Environ. Mutagen.* **1980**, *2* (1), 67−73.
(75) Howells, R. E.; Delves, C. J. A simple method for the identification of compounds which inhibit tubulin polymerization in filarial worms. *Ann. Trop. Med. Parasitol.* **1985**, *79* ((5), 507−12.
(76) Rennison, M. E.; Handel, S. E.; Wilde, C. J.; Burgoyne, R. D. Investigation of the role of microtubules in protein secretion from lactating mouse mammary epithelial cells. *J. Cell Sci.* **1992**, *102* (Pt 2), 239−47.
(77) Vasquez, R. J.; Howell, B.; Yvon, A. M.; Wadsworth, P.; Cassimeris, L. Nanomolar concentrations of nocodazole alter microtubule dynamic instability in vivo and in vitro. *Mol. Biol. Cell* **1997**, *8* (6), 973−85.
(78) Mikhailov, A.; Gundersen, G. G. Relationship between microtubule dynamics and lamellipodium formation revealed by direct imaging of microtubules in cells treated with nocodazole or taxol. *Cell Motil. Cytoskeleton* **1998**, *41* (4), 325−40.
(79) Borgers, M.; De Nollin, S.; Verheyen, A.; Vanparijs, O.; Thienpont, D. Morphological changes in cysticerci of Taenia taeniaeformis after mebendazole treatment. *J. Parasitol.* **1975**, *61* (5), 830−43.

CI050495H