JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

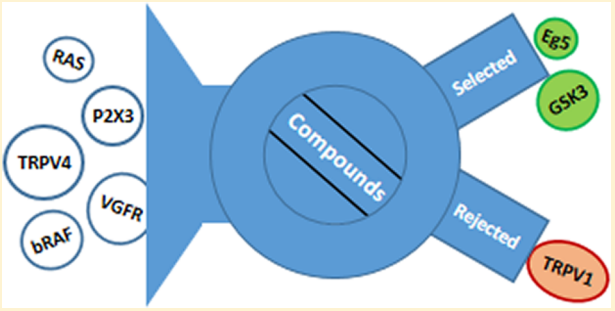
Enrichment Analysis for Discovering Biological Associations in Phenotypic Screens

Valery R. Polyakov, Neil D. Moorcroft, and Amar Drawid*

Sanofi, 55 Corporate Drive, Bridgewater, New Jersey 08807, United States

Supporting Information

ABSTRACT: A phenotypic screen (PS) is used to identify compounds causing a desired phenotype in a complex biological system where mechanisms and targets are largely unknown. Deconvoluting the mechanism of action of actives and identification of relevant targets and pathways remains a formidable challenge. Current methods fail to use the rich information available regarding compounds and their targets in a systematic way for this deconvolution. We have developed an enrichment analysis algorithm to identify targets associated with the desired phenotype in a rigorous data-driven manner using actives and hundreds of thousands of inactives in a PS, as well as results of thousands of available legacy target-based screens in an institution. Our method quantifies association between the PS and targets while reducing sampling bias, which leads to identification of novel targets, additional chemical matter, and appropriate assays. Its use is illustrated using two examples from our laboratories: TRAIL and DNA fragmentation. Enrichment analysis of these PSs is discussed using both biological pathway analysis and known cell biology to demonstrate the value of our method. We believe this enrichment analysis method is an indispensable tool for the analysis of PSs.



INTRODUCTION

The purpose of a phenotypic screen (PS) is to identify compounds causing a desired phenotype in a complex biological system where mechanisms of phenotypic change and most targets are unknown.[1−3] While PSs have generated quite a few drugs,[4] deconvolution of their results to determine mechanism of action (MOA) of the actives and identification of associated targets and pathways poses a significant challenge. Most approaches for deconvolution focus on a particular active compound and then try to identify its target directly through binding experiments or indirectly by comparing its genome-wide activity profile with those of compounds with known targets.[1] These approaches, however, fail to capitalize on the information available on all other actives in the PS.

An institution usually has a legacy of target-based screening results, which, in conjunction with all actives in the PS, can be used to tease out biological association of the desired phenotype. One intuitive way is to determine how many actives in the PS are also active in each of the "historical" target-based screens. Compound interaction at a target can be hypothesized to cause the desired phenotype if the screen for that target shares several common actives with the PS. This simple counting method, however, has sampling bias. A larger number of compounds may have been tested for that particular target compared with others, and hence the observance of several common actives with the PS may be just an artifact with no real biological association.

This sampling bias can be overcome by using enrichment analysis. Actives in a PS are "enriched" for a particular characteristic if the PS contains more actives corresponding to that characteristic than what would be expected by sheer chance. Enrichment analysis can also be referred to as over-representation analysis. Our enrichment analysis provides a systematic way to determine how enriched the actives in the PS are with the actives in a target-based screen. Statistically significant enrichment indicates that compound interaction at the target may possibly cause the desired phenotype. Enrichment analysis has been commonly performed in biology for comparing two gene sets, for example, to identify pathways associated with genes differentially expressed in a transcript profiling experiment.[5−7] While such "biological" enrichment analysis of genes typically uses all genes in the genome, formulation of an enrichment analysis problem in the chemical space is not obvious. In this publication, we have formulated and developed an enrichment analysis algorithm that takes into account actives and inactives in a PS and target-based screens to identify targets associated with the desired phenotype systematically. As long as compounds tested in a screen can be discriminated as actives and inactives, the algorithm determines which of the actives and inactives in the PS are active or inactive in a target-based screen to find the association with the target. The algorithm can be applied independently of structural

Received: April 25, 2013
Published: January 18, 2014

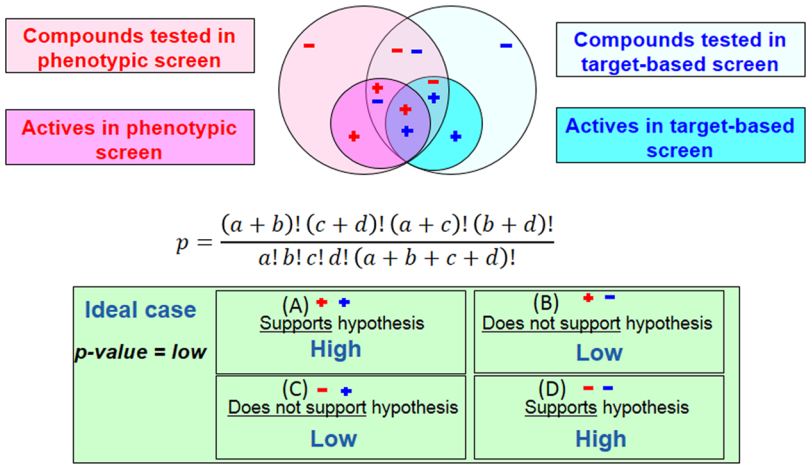


Figure 1. Description of the enrichment analysis method for a PS. Top part shows a Venn diagram consisting of compounds tested in the PS (large faint pink circle), actives in the PS (small dark pink circle), compounds tested in a target-based screen (large faint blue circle) and actives in the target-based screen (small dark blue circle). Actives and inactives in a screen are designated as + and −, respectively. Red and blue are for PS and target-based screen, respectively. Only those compounds tested in both screens are considered (++, +−, −+, −−). Bottom part shows that (A) actives in both screens (++) and (D) inactives in both screens (−−) support the hypothesis that compound interaction at the target results in the desired phenotype, whereas B and C, that is, actives in one screen and inactives in the other (+− and −+) do not support it. These four values form a 2 × 2 contingency table for the Fisher exact test used in the enrichment analysis. In the ideal case, diagonal elements of the contingency table are high and nondiagonal elements are low, resulting in a low *p*-value and providing evidence to the hypothesis that compound interaction at the target results in the desired phenotype. While the illustration is provided for one target, actual analysis is performed by considering all targets one at a time.

clustering or molecular binding mode. We have applied this method to multiple projects, illustrated by the following two PSs: TRAIL and DNA fragmentation. The TRAIL PS was performed contemporaneously with development of the algorithm, while the DNA fragmentation PS was performed prior to development of the algorithm and was used for validation purposes.

TNF-related apoptosis-inducing ligand (TRAIL) is a promising target for the treatment of cancer because of its ability to induce apoptosis in a range of oncogenic cells while sparing normal cells.[8] This selective apoptosis induction has led to TRAIL agonistic antibodies and recombinant TRAIL therapies that are currently undergoing clinical trials.[9] Unfortunately, their efficacy is limited by drug resistance that can be intrinsic to the cell type or can be acquired by cells that were initially TRAIL sensitive.[10,11] For this reason, there has been interest in identifying compounds that will sensitize tumor cells to TRAIL induced apoptosis.[12,13] In this context, a phenotypic high-throughput screen (HTS) was performed to identify TRAIL potentiators with clinical utility for the treatment of cancer.

In a broader sense, malignant cells commonly have a decreased ability to die under normal physiological conditions, and therefore drug-mediated restoration or enhancement of apoptosis remains a valid approach to treat oncogenic diseases. Indeed, the primary outcome for most chemotherapeutic therapies is apoptosis, albeit via a variety of mechanisms.[14] Apoptosis can be induced through extrinsic death-receptor engagement such as TRAIL but can also arise from intrinsic pathway activation, down-regulation of cell survival mechanisms, activation of caspases, or other known or unknown mechanisms. However, all apoptosis pathways lead to DNA fragmentation.[15] Therefore, a cell-based phenotypic HTS campaign was undertaken specifically measuring DNA fragmentation as a surrogate for apoptosis induction irrespective of MOA. The DNA fragmentation PS was performed prior to development of the enrichment analysis algorithm, which was applied retrospectively.

METHODS

Hypothesis Construction and Testing. For enrichment analysis of the PS, targets were considered one at a time. Enrichment analysis determines statistical significance for the hypothesis that actives in the PS are enriched with the actives for the target, which implies that compound interaction at that target results in the desired phenotype. All compounds, whether active or inactive, screened in both the PS and the target-based screen were considered for this analysis (Figure 1). They were used to create a 2 × 2 contingency table as follows:

A. Number of compounds active in both the PS and target-based screen. These compounds support the hypothesis that compound interaction at the target results in the desired phenotype.
B. Number of compounds active in the PS but inactive in the target-based screen. These compounds do not support the above hypothesis.
C. Number of compounds inactive in the PS but active in the target-based screen. These compounds do not support the above hypothesis either.
D. Number of compounds inactive in both the PS and target-based screen. These compounds support the above hypothesis.

Compounds evaluated in only one screen were ignored because they did not provide any evidence for or against this hypothesis.

Enrichment analysis employed the Fisher exact test to determine whether the PS actives were enriched with the actives for the target. Fisher exact test for over-representation is based on hypergeometric probability distribution that describes sampling without replacement and is used to calculate one-tailed *p*-value.[16] This *p*-value gives the probability that the observed or extreme screening output occurs "by chance" if the

dx.doi.org/10.1021/ci400245c | J. Chem. Inf. Model. 2014, 54, 377−386

target is assumed not to be responsible for the desired phenotype. Low *p*-value indicates possible association between the target and phenotype, which, however, does not necessarily imply that the interaction between compounds and targets results in the desired phenotype.

Implementation. Construction of contingency tables for enrichment analysis required retrieval of all historical screen data for all compounds, all actives as well as all inactives, tested in the PS. This computationally intensive procedure required fast retrieval from the database of compound activities in historical screens and storage and manipulation of the large amount of retrieved data.

Classification scheme for designating a compound active or inactive from historical screening data was prespecified in the software based on the standard screening process in the organization. Specifically, a compound was considered active for a given target if a dose–response experiment was performed and the corresponding measurement (e.g., IC_{50} or EC_{50}) was less than 10 μM. A compound was considered inactive if the dose response measurement was greater than this threshold or if only single point measurements (e.g., percent inhibition or activation) were recorded. The logic behind such classification, in addition to single point measurements usually not being very accurate, is that if single point activity were promising, the compound would have been selected for dose–response studies according to our standard screening process. If multiple activity values were present, an average value was calculated. Alternative classification schemes can exist. For example, if compounds are annotated in the database as being active or inactive in a particular screen (as is the case in many PubChem data sets), that classification can be retrieved directly. Classification can also be based on single point inhibition or activation values.

Input to the algorithm (Figure 2) was a comma-separated text file of all compounds active in the PS containing one compound identifier (CID) per line. PS activities defined by HTS dose–response data were used as input for both TRAIL and DNA fragmentation without supplementary refinement through further characterization. Any additional available data (for example, project team or counter screen data) can be considered during the evaluation of each target-based hypothesis generated by the algorithm. Optionally, one or more chemical series names to which a compound belongs to can be added after its CID, separated by commas. Manually curated chemical series information can be useful at later stages of the project when active compounds are used to identify possible chemotypes responsible for activity against a particular target.

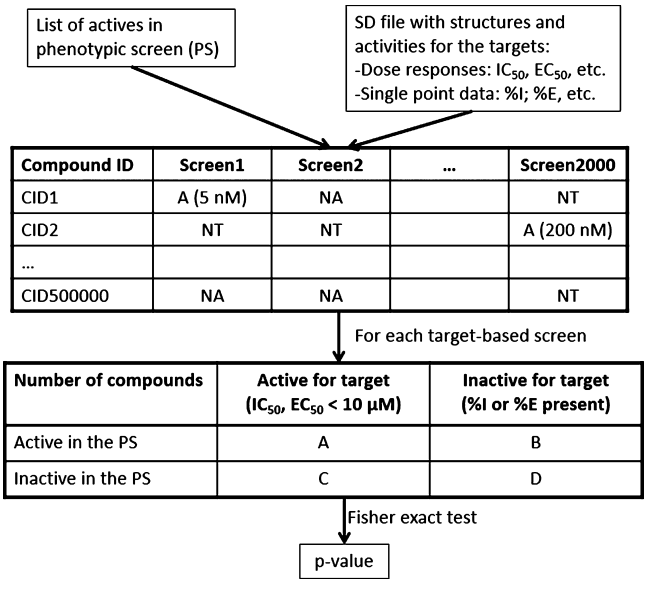


Figure 2. Method workflow illustration. In the top table, A = active, with activity value given in the parentheses; NA = not active; NT = not tested.

SD (structure data, a wrap of the standard Molecular Design Limited, Inc. molfile allowing inclusion of associated data) files corresponding to each CID were retrieved from the in-house compound activity database. SD files containing single point and dose–response data were parsed to identify screens in which the compound was tested and retrieve the compound's activities in those screens. The above classification scheme was used to determine whether the compound was (i) active, (ii) inactive, or (iii) not tested in each screen. SD files can also optionally contain compound structure and biological annotation (e.g., mode of action of the compound on the target), which can be useful at later stages of the project.

In the next step, the 2 × 2 contingency table described above was generated for each target-based screen. Specifically, all active and inactive compounds in the PS were tabulated based on their activity or inactivity in that particular screen using either HTS dose–response data or other information about the compounds. Specifically, in the TRAIL project, the PS actives were determined by manual inspection of potentiation curves by the screener. In the DNA fragmentation screen, a 1 μM activity cutoff was used. Compounds not tested in that screen were ignored.

Finally, enrichment analysis was performed for each target-based screen using the Fisher exact test described above, and the *p*-value was recorded.

Hardware and Software. Client-side calculations were performed from a Lenovo ThinkStation C20x[17] computer running Windows[18] 7 Service Pack 1 64-bit operating system that was equipped with Intel Xeon X5650 CPU[19] (two processors), 2.66 GHz, and 12 GB of RAM. Microsoft .NET framework 4.0[20] was installed on the computer. All software programming utilized Microsoft Development Environment (Microsoft Visual Studio 2010 Ultimate) 2010 in C#.NET language.[21]

RESULTS

Enrichment analysis was performed for the TRAIL and DNA fragmentation PS campaigns. While these two screens used ~500 000 compounds each with some overlap, enrichment analysis included millions of compounds tested in over 1300 target-based screens.

Enrichment analysis of 1264 confirmed hits from the TRAIL PS identified over 100 potential molecular targets with a *p*-value (Fisher test) below the threshold of 0.05. Select examples are illustrated in Table 1. Results of pathway analysis performed using MetaCore pathway analysis software (GeneGo)[22] and literature review are shown in Figure 3, with an emphasis on related molecular pathways with multiple identified targets. Existing literature and known cell biology were used to prioritize each molecular target or pathway based on plausibility and novelty. Finally, qualitative correlations between target potency and PS efficacy, distribution of PS activity across distinct chemical series potent on a given target, and apparent compound selectivity using historical data were all used in the

dx.doi.org/10.1021/ci400245c | J. Chem. Inf. Model. 2014, 54, 377–386

Table 1. Selected Output from Enrichment Analysis of 1264 Confirmed TRAIL PS Actives[a]

target	number of actives	number of inactives	% of actives	*p*-value of over-representation
β-catenin	20	1	95	0[b]
CDK2	64	110	37	0[b]
CDK4	13	72	15	0[b]
CDK5	33	92	26	0[b]
VEGFR1	19	179	10	0[b]
VEGFR2	17	122	12	0[b]
SRC	23	50	32	0[b]
JNK1	6	26	19	3.6×10^{-11}
GSK3b	117	109	52	1.9×10^{-10}
Gquartet DNA	6	69	8	5.9×10^{-8}
MEKK1	13	35	27	3.1×10^{-6}
PAK1	4	190	2	9.9×10^{-6}
FAK	5	149	3	1.6×10^{-4}
INSRK	3	121	2	3.3×10^{-4}
IGF1R	4	151	3	8.7×10^{-4}
tubulin	4	5	44	3.0×10^{-3}
VEGFR	3	2	60	4.8×10^{-3}
ABL	11	30	27	1.1×10^{-2}

[a]"Number of actives" column represents term A of Figure 2, while "number of inactives" column represents term C. [b]*p*-value of 0 is too small to detect.

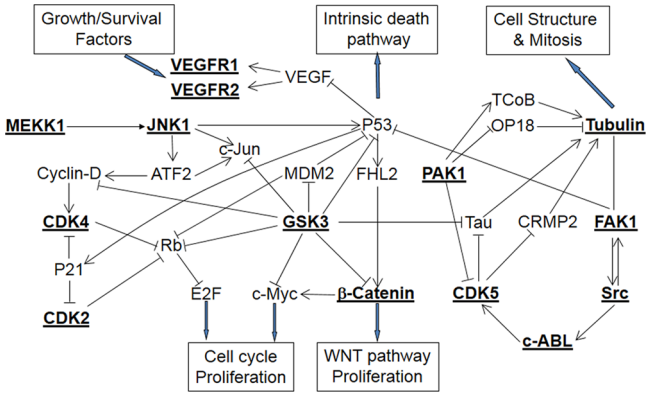


Figure 3. Results of pathway analysis based on a subset of targets enriched in TRAIL PS. Enriched targets are in bold and underlined. Relevant pathways were identified using MetaCore pathway analysis software (GeneGo)[22] and merged to form a cohesive picture using literature review.

evaluation. Chemical series present in the PS were detected by automatic clustering of all active compounds followed by the manual curation of the results by a medicinal or a computational chemist. The output of this prioritization is discussed below.

Glycogen synthase kinase 3 (GSK3) stood out from the analysis with a strong association (*p*-value of 1.89×10^{-10}; the authors consider an association to be strong if the *p*-value is less than 0.05) and a large number of active compounds. This chemical matter spanned several structurally distinct series specifically optimized for GSK3 inhibition with established cell-based efficacy, together with compounds from several series primarily inhibiting other kinases but whose historical kinase selectivity profiling demonstrated GSK3 inhibition. This breadth of chemical matter across six series confirmed a strong and robust association between GSK3 inhibition and activity in the TRAIL PS. Further validating the hypothesis initially proved challenging because GSK3 inhibitors often inhibit cell cycle dependent kinases (CDKs)[23] as well, and CDKs were also strongly implicated in the enrichment analysis (Table 1). We evaluated the available kinase profiling data for each compound (data not shown), and this informatics approach supported the separation of CDK inhibition from GSK3 pharmacology using proprietary compounds discovered in the HTS such as indazole inhibitor **1**[24] (Figure 4). In this case, the 7-position halogen is responsible for selectivity against CDKs.[25]

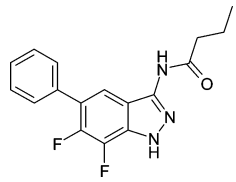


Figure 4. GSK3 inhibitor compound **1**.[24]

Both TRAIL sensitization[26,27] and attenuation[28] by GSK3 inhibition has been reported previously. This apparent paradox reflects the context-dependent nature of GSK3 signaling in both the extrinsic and intrinsic death pathways.[29] The intrinsic pathway is relevant because extrinsic-initiated caspase activation alone is insufficient for apoptosis in most cells (known as type-II cells[30]) without signal amplification through the intrinsic pathway. Mitochondrial membrane disruption is a key aspect of intrinsic apoptosis, and GSK3 facilitates this through activation of Bax[31] and inactivation of Mcl-1.[32] Outside the mitochondria, GSK3 is intimately involved with various pro- and antiapoptotic transcription factors such as p53, c-Myc, HSF-1, and *β*-catenin, plus structural proteins including tubulin, tau, and kinesin.[29] Interestingly, several of these pathways and proteins were identified in our pathway analysis and link to other molecular targets from the enrichment analysis (Figure 3). GSK3 also controls receptor activation in extrinsic apoptosis via a multiprotein complex proximal to the TRAIL receptor, where GSK3 inhibition directly amplifies death receptor signaling.[33,34] Overall GSK3 is a strong candidate MOA for chemical triage.

β-catenin is an integral WNT pathway component and was a strong mechanistic candidate for TRAIL potentiation (Table 1). Indeed, abnormal WNT signaling is associated with cancer[35] and *β*-catenin down-regulation can sensitize cancer cells to TRAIL.[36] However, almost all active compounds contributing to the low *p* value also inhibit GSK3, a key WNT pathway regulator.[37,38] GSK3 inhibition stabilizes *β*-catenin resulting in nuclear translocation and consequent transcriptional activation of genes including c-Myc, E-cadherin, and cyclin D1.[39] In this pathway, GSK3 inhibition is associated with cell growth and survival rather than apoptosis. Therefore, *β*-catenin over-representation appears to be a bystander artifact and surrogate of GSK3 inhibition; it was discounted from further consideration.

G-quartet DNA was also identified as a possible target (Table 1). G-quadruplex binders and intercalators inhibit DNA translation and initiate a rapid DNA damage response[40] and cell death.[41] As such, these compounds likely sensitize TRAIL through intrinsic pathway activation or up-regulation of death receptors. Indeed, a previously reported TRAIL enhancer screen[12] identified many intrinsic pathway activators such as DNA or RNA synthesis inhibitors, and DNA binders and proteosome inhibitors. We identified similar compounds

dx.doi.org/10.1021/ci400245c | J. Chem. Inf. Model. 2014, 54, 377−386

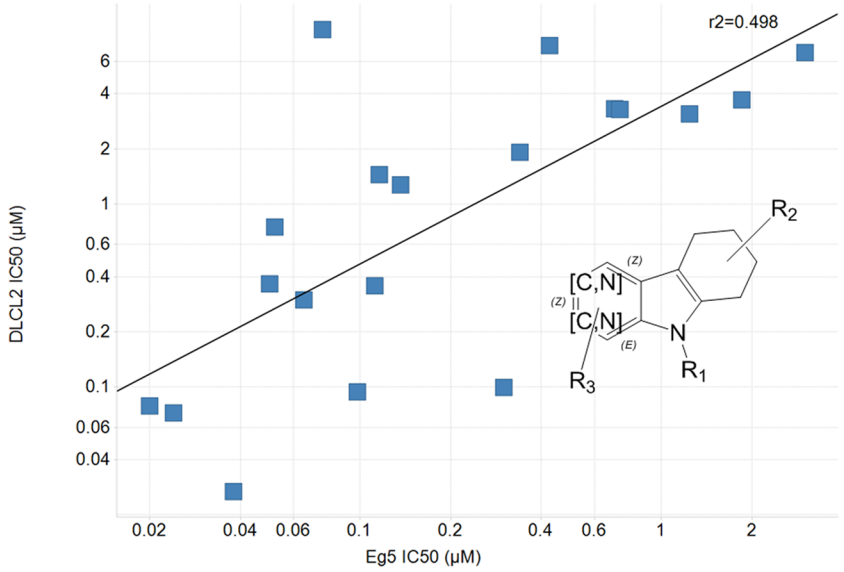


Figure 5. A plot of log(Eg5 IC_{50}) vs log(DLCL2 IC_{50}) showing a positive correlation between the two measurements in the tetrahydro-γ-carboline series. Inactive compounds in DLCL2 cell line were excluded from the plot.

including gemcitabine, decitabine, 6-thioguanine, mercaptopurine, 6-azauridine, toyocamycin, and mycophenolic acid, along with protein synthesis inhibitors streptovitacin A and cycloheximide. Interestingly, both TRAIL and its receptors are transcriptional targets of p53,[42] and this may partially explain why many chemotherapy drugs sensitize cancer cells to TRAIL.[43,11] Tumor suppressor p53 plays a central role in intrinsic apoptosis, with inactive, mutated, or insufficient p53 contributing significantly to chemotherapy resistance.[44] TF-1 cells used in the TRAIL PS are hemizygous for wild-type and mutant p53, but wild-type apoptotic function is reported for the resulting protein.[45] Absence of a p53-based screen among our historical biological screening data precluded p53 from being highlighted directly in the enrichment analysis. Nonetheless, G-quartet DNA and other identified targets closely associated with p53 from our pathway analysis (Figure 3) argue strongly for assays evaluating p53 or intrinsic pathway activation as part of follow-up MOA studies.

Enrichment analysis also identified several growth factors including vascular endothelial growth factor (VEGF; receptor isoforms, VEGFR, VEGFR1, VEGFR2), insulin receptor kinase (INSRK), and insulin-like growth factor 1 receptor (IGF1R) (Table 1). TF-1 cells are growth factor-dependent for survival, with deprivation causing intrinsic mitochondria-mediated apoptosis.[46,47] Therefore, growth factor antagonism could amplify extrinsic TRAIL-mediated apoptotic signaling through the intrinsic pathway. Supporting this, VEGF inhibition has been implicated in TRAIL sensitization based on the antiangiogenic drugs sorafenib[48−51] and sunitinib.[52] However, these drugs inhibit several kinase targets beyond VEGFR including PDGFR, Flt3, and RET, and therefore VEGF signaling is not definitively implicated. IGF1R inhibitors have shown mixed effects, both sensitizing[53] and protecting against extrinsic apoptosis.[54] Overall, there is sufficient evidence that growth factor receptor inhibition could be responsible for sensitizing TF-1 cells to TRAIL, but additional studies are needed to confirm this hypothesis, including a rigorous selectivity evaluation of the various growth factor inhibitory compounds identified out of the PS.

While only few compounds contributed to tubulin enrichment (Table 1), the HTS actives included antimitotic agent grisofulvin and several colchicine derivatives, suggesting that this correlation is authentic. Indeed, microtubule-targeted agents nocodazole and paclitaxel are reported to sensitize cancer cells to TRAIL.[55,67,56] The enrichment analysis also identified ABL, Src, FAK, PAK1, and CDK5, kinase targets linked to cytoskeleton arrangements and tubulin stability (Figure 3), but these targets proved less fruitful. The Src, ABL, and CDK5 inhibitors primarily inhibited GSK3, and GSK3 could not be discounted adequately. Inspecting the PAK1 actives revealed TRAIL activity restricted to a single chemotype, questioning the authenticity of PAK1. FAK was also deprioritized because the five TRAIL active FAK inhibitors were disproportionately weak against FAK compared with the 149 FAK inhibitors that did not potentiate TRAIL. Overall, the mitotic pathway appears to contribute to TRAIL potentiation, but the associated kinase targets illustrate the importance of reviewing contributing compound profiles when evaluating enrichment analysis hypotheses.

MAPK superfamily members MEKK1 and JNK1 were also highlighted by our algorithm (Table 1). Functionally, the MAPKs play a complex role in apoptosis in both the intrinsic[57] and extrinsic signaling pathways,[58] with JNK inhibition known to reduce[59] as well as enhance[60] apoptosis in cancer cells. Of particular note, JNK inhibitors can synergize with GSK3 inhibitor mediated TRAIL potentiation[61] but can prevent TRAIL sensitization caused by mitotic arrest.[62] An informatics analysis showed that many of the MAPK inhibitors identified in the TRAIL PS also inhibited GSK3 and several other kinases. Despite the compound selectivity challenges, MAPK pathway contribution to efficacy appears worthy of further studies.

CDKs also featured prominently in the enriched targets (Table 1). CDK activity is often deregulated in tumors; however, the interplay between cell cycle control and apoptosis is complex.[63] Cell cycle arrest induced by CDK inhibitors roscovitine and olomoucine sensitizes various cancer cells to TRAIL[64−67] but is protective in others.[68] In growth factor-dependent cells such as TF-1, conflicting signals for cell proliferation and cell cycle arrest may be sufficient to move the

dx.doi.org/10.1021/ci400245c | J. Chem. Inf. Model. 2014, 54, 377−386

survival−apoptosis balance toward death. Interestingly, CDK inhibitors can activate the p53 pathway,[69] with intrinsic pathway activation also a candidate MOA. Given that GSK3 inhibitors tend to also inhibit CDKs, the reciprocal challenge applies to validating CDKs. While mixed GSK3−CDK inhibitors contributed to enrichment, data mining also identified compounds with CDK inhibition as primary MOA (data not shown). While this supports but not definitively validates CDKs in TRAIL potentiation, deconvoluting individual CDK members proved very challenging. Nonetheless, cell-cycle analysis appears valuable for MOA studies.

Overall, enrichment analysis of the TRAIL PS identified several targets feeding into the extrinsic and intrinsic death pathways. The strong context dependence and contradictory roles of many of these targets makes it difficult to predict in advance how these pathways will be "wired" in TF-1 cells. Nevertheless, the enrichment analysis provided several leads for further investigation. The value of a PS is to identify chemical matter operating via novel modes or discover new mechanisms for existing compounds. Therefore, understanding the "wiring" of the screening cell is important, if only to mechanistically contextualize the hit compounds based on the novelty and desirability of the MOA. For the TRAIL PS in TF-1 cells, the target analysis and the pathway analysis suggest mechanistic assays appropriate for triage of the chemical matter including: GSK3, tubulin binding, cell cycle analysis and intrinsic apoptosis initiation such as p53 activation or mitochondrial membrane permeabilization.

Enrichment analysis was also performed on the DNA fragmentation PS campaign. One particularly interesting series of tetrahydro-γ-carboline derivatives[70] (Figure 5) was identified as a potent inducer of apoptosis. This series contained more than 20 apoptosis-inducing compounds efficacious in HL-60 cells and other tumor cell lines such as DLCL2. Mechanistic studies demonstrated a cell cycle block prior to mitosis, whereas in vitro assays established that these compounds did not directly interfere with tubulin.

In a search for alternative targets associated with this cell cycle phenotype, mitotic kinesin Eg5 (also known as kinesin spindle protein (KSP)) was evaluated as a potential MOA. This protein is required for bipolar spindle formation in the early phase of mitosis.[71] Eg5 inhibition is known to cause mitotic cell cycle arrest and eventually apoptosis in cancer cells.[72] When these compounds were evaluated in the Eg5 biochemical assay, a strong correlation was discovered between Eg5 inhibition IC_{50} and apoptosis induction (Figure 5).[70] This correlation was particularly convincing considering the inherent discrepancies between biochemical and whole cell assays.

Our analysis algorithm retrospectively generated a large number of statistically relevant targets for the DNA fragmentation PS output. Eg5 was indeed identified by our enrichment method with a highly significant *p*-value. Interestingly, filtering the known function[73] of these targets by keywords associated with the mitosis-blocking phenotype highlighted only a few candidate targets including Eg5 and tubulin (Table 2). Other mitotic targets identified included CDK5 and GSK3 due to their phosphorylation of microtubule-binding protein tau[74] and the mitotic kinase Aurora2.[75]

Therefore, filtering the output of our algorithm by function generated a set of relevant putative targets for mechanistic interrogation. Although applied to this chemical series with hindsight, Eg5 was correctly identified as a strong candidate target using our approach.

Table 2. Selected Output from Enrichment Analysis of Confirmed DNA Fragmentation PS Actives[a]

target	number of actives	number of inactives	% of actives	*p*-value of over-representation
CDK5	169	1728	9	0[b]
GSK3b	81	2081	4	0[b]
Tubulin	170	46	79	0[b]
Aurora2	73	1575	4	0[b]
Eg5	22	2969	1	0[b]

[a]"Number of actives" column represents term A of Figure 2, while "number of inactives" column represents term C. Output is filtered by keywords associated with mitosis or microtubules based on protein function available from UniProt.[73] [b]*p*-value of 0 is too small to detect.

DISCUSSION

Our enrichment method offers several advantages. First and foremost, it connects multiple compounds to multiple genes simultaneously and thus bridges the gap between biological and chemical space in a systematic data-driven manner to deconvolute a PS.

PS analysis methods available in the literature usually focus on a particular active compound rather than a target to discover biological associations.[1] In contrast, our method uses all screened compounds as evidence to form the hypothesis that compound interaction at a particular target results in the desired phenotype.

Other PS analysis methods fail to use data from historical screens. Recognizing that an institution may have many historical screens and that most compounds screened in the current PS are already evaluated in these screens, we use this treasure trove to establish or refute a hypothesis about a target. Moreover, due to the massive lookup of all compounds' data in all screens, the method can simultaneously form hypotheses about all targets for which data is available.

Another advantage of the enrichment method is that it reduces sampling bias while quantifying association between a target and a PS. If many actives in the PS are also active for a particular target, one may be tempted to form the hypothesis that compound interaction at that target results in the desired phenotype. But this could happen by chance because an extraordinarily large number of compounds were tested in both the target's screen and the PS. This is especially true for targets in advanced stages of the drug discovery and development pipeline of an institution. Many more compounds have been screened against these targets than the targets whose programs were terminated early, thus creating a sampling bias. Spurious association with these targets will lead to wasteful follow-up experiments. The enrichment analysis prevents them by correctly adjusting for sampling bias.

The enrichment method consists of a major step in deconvoluting a PS, and its results can be used in multiple ways. A PS is used when a particular aspect of biology has several unknowns. Our analysis starts uncovering them based on two types of tools newly discovered to be associated with the desired phenotype: targets and compounds.

A target closely associated with a PS can serve as a novel target for the particular disease area. Targets can be used to identify pathways that play a major role in the desired phenotype and thus identify even more targets in those pathways that were not screened previously. This approach also opens the doors for polypharmacology. Moreover, the

uncovered biology suggests appropriate assays to further elucidate the mechanisms of the desired phenotype.

Our method has value in prioritizing candidate targets for mechanistic deconvolution studies or chemical triage. Depending on the novelty and desirability of each target, this may lead to specific targets with existing chemical matter requiring validation, such as GSK3 for the TRAIL PS, or suggest antitarget or target-based assays in the triage of PS chemical matter, such as evaluating tubulin binding or p53 activation. Over-represented targets are good candidate hypotheses for attractive chemical matter with unknown mechanisms. This can be particularly powerful when combined with other mechanistic information such as the mitotic block observed for the tetrahydro-γ-carboline series for the DNA fragmentation PS. Although retrospectively applied, this technique correctly highlighted Eg5 in a short list of relevant candidate targets.

In addition, our method enables repurposing of an institution's own compounds. An institution may have a great number of proprietary compounds that were screened against the targets now hypothesized to be associated with the PS. These compounds can now be progressed rapidly if they confer the desired phenotype. The scale of our method provides such access to millions of compounds in an institution's repository. This proved to be the case for GSK3, where many additional inhibitors were available to the project team beyond those evaluated in the TRAIL PS.

While the enrichment method offers several advantages, it has a few limitations related to data, methodology, and implementation, which are discussed below. Certain limitations may be overcome by possible future extensions to the method.

Since the method heavily uses the institution's database, accuracy of annotation in the database is essential for its success. Data from historical screens is often entered by different laboratories at different locations within the institution, often from many legacy institutions, over several years. In addition, multiple assays with different protocols may exist for the same target or targets of interest. Special care needs to be taken to ensure that data is entered consistently and that results from different screens can be compared.

Even if data in the database is accurate, erroneous associations may appear if the actives common to the target and the PS are not selective for that target. In the case of TRAIL potentiation, the majority of correlations are to kinases and must be interpreted with care due to similarity of ATP binding pockets. This potential promiscuity can lead to a bystander effect due to cross-reactivity of chemical matter, which must be discounted to fully validate a target. For example, cross-reactivity of GSK3 and CDKs was particularly problematic for the TRAIL PS. One simple way to overcome this issue is to apply a selectivity filter and remove nonselective compounds before performing the enrichment analysis. However, apparent selectivity is restricted by available profiling; limited existing data may erroneously suggest high selectivity.

Interpretation is particularly difficult when new targets arise from a chemical series with known activities associated with the desired phenotypic outcome. With GSK3 inhibition dominating the TRAIL PS, whether fully validated as a MOA or not, other putative targets associated with chemical matter having some degree of GSK3 activity required significant care in analysis and in many cases were deprioritized for follow-up. Nonetheless, inhibiting multiple targets can certainly influence overall cellular efficacy as known in the literature for dual inhibition of JNK and GSK3. Potential dual activities are often overlooked in the triage process, whereas our enrichment analysis method provides an opportunity to mine in-house data for potential correlations of synergistic or ablative additional "off-target" activities.

Moreover, one needs to be cognizant about the bias in both the chemistry and biology aspects of the data. On the chemistry side, compound collection is usually biased, and some compound families are typically represented more than others are, due to restrictions on the availability of synthetic approaches as well as patent considerations. One may perform enrichment analysis based on compound families rather than compounds to reduce this bias, keeping in mind that arbitrary decisions need to be then made about how to define a compound family and how to treat compound families with both active and inactive compounds in a screen. Bias can be eliminated only by screening the PS and the target with all possible compounds, which is practically impossible.

On the biology side, enrichment analysis is applicable only to the targets already screened within the institution. This can be overcome to some extent by using external databases containing a variety of targets, such as PubChem MLSCN.[76] However, one should review externally performed screens carefully and normalize their data before including them in the enrichment analysis. This may be particularly useful for commercial screening collections or evaluation of known clinical compounds.

Screening data may unwittingly harbor biases because assays do not necessarily distinguish between allosteric or orthosteric interactions, nor functional efficacies such as agonism and antagonism. Additionally, they often only capture a subset of a target's pharmacology that is not necessarily valid across different substrates or biological contexts.

In terms of methodology, our algorithm requires parameters to be set for defining actives. It currently uses only one cutoff value for all screens. However, compound activities in biochemical and cellular assays often have different scales, and hence one may want to apply different cutoff values for different types of screens. In either case, defined cutoffs are arbitrary. Instead of applying cutoffs, one can enhance the method by treating activities as continuous variables and using regression type approaches.

One can argue that χ-squared test may give similar results to the Fisher exact method employed by our enrichment analysis. While this may be true when all four numbers in the contingency table are large, the number of actives in both PS and a target-based screen is often less than 5, which violates the assumptions for a χ-squared test. Fisher exact test, however, is based on hypergeometric distribution and is the method of choice for such situations. Alternatively, one can use the Barnard exact test, a powerful test that estimates the nuisance parameter to maximize the p-value and uses more support points especially when sample size is small.[77] Hypergeometric probability under the null hypothesis in the Fisher exact test does not depend on the nuisance parameter. Results of both tests are expected to be similar due to a large overall sample size (number of inactives in both PS and a target-based screen is usually very large).

On the implementation side, due to the large amount of data retrieved from the database and processed on-the-fly, the database interaction part was coded using a multithreaded approach where more than 100 simultaneous data requests were made to the database. Upon return of the results, the data was processed and accumulated in computer memory using

dynamic memory allocation capabilities available in .NET.[21] As a result, a significant amount of information was being kept in the random access memory of the computer and at least 10 GB of RAM is recommended to run an application for a large number (500 000) of tested compounds.

While we have focused on one particular application of the method, it can be extended to work with many other applications. In the current application, we compared PS data with historical target-based screening data. Alternatively, if we compare a PS with historical PS's, we can determine biological association between the phenotype in question and other phenotypes. Additionally, one can also compare data from a target-based screen with data from historical target-based screens. Thus, our method serves as a powerful tool for biological interpretation of any HTS.

Moreover, this approach can be used to discover closely related targets in nonphenotypic whole cell assays as well as to investigate polypharmacology of compounds when the desired pharmacological effect is obtained by nonspecific inhibition of more than one target. For utilizing the same algorithm, such whole cell (or any other in vivo or ex vivo) assay would be treated as a phenotypic assay and the analysis described in the manuscript would take place. If any other assay or target would be very closely associated with the given target, one might conclude that that target is also relevant for the desired cell, or ex vivo, or in vivo effect. Such treatment of a non-PS assay as a PS assay can be theoretically justified by the similarity of the binding sites of two related targets, which will bind the same or similar compounds. Further research is needed to justify such applications of the method.

Another useful extension of our method is to perform pathway enrichment analysis in a statistically rigorous manner. Currently, one can identify enriched targets based on our method and then identify pathways enriched in these targets using methods noted earlier. This approach is not entirely accurate because (i) targets are not weighed by the chemistry evidence and (ii) many targets are not pathway-specific. First, compounds active in a PS will have different association *p*-values with the targets in a pathway, and association may be quite poor with some targets in the pathway. Second, if a pathway is broadly defined, it contains many genes shared across many other pathways (e.g., AKT) and thus actives for such genes should not be considered evidence that a PS is associated with that particular pathway.[78−80] One needs to integrate this information in a rigorous fashion to connect compounds directly to a pathway.

CONCLUSION

Our method takes advantage of the rich information available from historical screens to identify targets associated with a PS in a systematic manner. We believe that it should be an essential tool in an institution's arsenal for analysis of PSs.

ASSOCIATED CONTENT

Supporting Information

HTS protocols for TRAIL and DNA fragmentation,[81] DLCL2 viability assay,[70] and EG5 pharmacological assay,[70] together with the underlying source code for processing the data and the compiled binaries. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: adrawid@gmail.com.

Present Addresses

For Valery Polyakov, now Senior Investigator I, Novartis Institute for Biomedical Research, 4560 Horton Street, Emeryville, CA 94608-2916, use valery.polyakov@gmail.com.

For Neil Moorcroft, now Research Scientist, Chromocell Corporation, 685 U.S. Highway One, North Brunswick, NJ 08902, use neilmoo@centurylink.net.

For Amar Drawid, now Business Consultant, ZS Associates, 212 Carnegie Center, Princeton NJ 08540, use adrawid@gmail.com.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank the TRAIL and DNA fragmentation project teams and respective screening groups for data utilized both as input to the algorithm and in discussion of subsequent target-based hypotheses.

REFERENCES

(1) Hart, C. P. Finding the target after screening the phenotype. *Drug Discovery Today* **2005**, *10*, 513−519.

(2) Wale, N.; Karypis, G. Target fishing for chemical compounds using target-ligand activity data and ranking based methods. *J. Chem. Inf. Model.* **2009**, *49* (10), 2190−2201.

(3) Haggarty, S. J. The principle of complementarity: Chemical versus biological space. *Curr. Opin. Chem. Biol.* **2005**, *9* (3), 296−303.

(4) Swinney, D. C.; Anthony, J. How were new medicines discovered? *Nat. Rev. Drug Discovery* **2011**, *10* (7), 507−19.

(5) Hosack, D. A.; Dennis, G., Jr.; Sherman, B. T.; Lane, H. C.; Lempicki, R. A. Identifying biological themes within lists of genes with EASE. *Genome Biol.* **2003**, *4* (10), R70.

(6) Huang, D. W.; Sherman, B. T.; Tan, Q.; Kir, J.; Liu, D.; Bryant, D.; Guo, Y.; Stephens, R.; Stephens, R.; Baseler, M. W.; Lane, H. C.; Lempicki, R. A. DAVID Bioinformatics Resources: Expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res.* **2007**, *35* (Suppl2), W169−175.

(7) Dennis, G., Jr.; Sherman, B. T.; Hosack, D. A.; Yang, J.; Gao, W.; Lane, H. C.; Lempicki, R. A. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* **2003**, *4* (5), P3.

(8) Ashkenazi, A.; Dixit, V. Death receptors: Signaling and modulation. *Science* **1998**, *281* (28), 1305−1308.

(9) Abdulghani, J.; El-Deiry, W. S. TRAIL receptor signaling and therapeutics. *Expert Opin. Ther. Targets* **2010**, 1091−1108.

(10) Koschny, R.; Walczak, H.; Ganten, T. The promise of TRAIL—potential and risks of a novel anticancer therapy. *J. Mol. Med. (Heidelberg, Ger.)* **2007**, *85*, 923−935.

(11) Li, Y.; Wang, H.; Wang, Z.; Makhija, S.; Buchsbaum, D.; LoBuglio, A.; Kimberly, R.; Zhou, T. Inducible resistance of tumor cells to tumor necrosis factor-related apoptosis-inducing ligand receptor 2-mediated apoptosis by generation of a blockade at the death domain function. *Cancer Res.* **2006**, *66* (17), 8520−8528.

(12) Booth, N. L.; Sayers, T. J.; Brooks, A. D.; Thomas, C. L.; Jacobsen, K.; Goncharova, E. I.; McMahon, J. B.; Henrich, C. J. A cell-based high-throughput screen to identify synergistic TRAIL sensitizers. *Cancer Immunol. Immunother.* **2009**, *58* (8), 1229−1244.

(13) Schimmer, A. D.; Thomas, M. P.; Hurren, R.; Gronda, M.; Pellecchia, M.; Pond, G. R.; Konopleva, M.; Gurfinkel, D.; Mawji, I. A.; Brown, E.; Reed, J. C. Identification of small molecules that sensitize resistant tumor cells to tumor necrosis factor-family death receptors. *Cancer Res.* **2006**, *66* (4), 2367−2375.

(14) Elmore, S. Apoptosis: A review of programmed cell death. *Toxicol. Pathol.* **2007**, *35* (4), 495−516.

(15) Guchelaar, H. J.; Vermes, A.; Vermes, I.; Haanen, C. Apoptosis: Molecular mechanisms and implications for cancer chemotherapy. *Pharm. World Sci.* **1997**, *19* (3), 119−125.
(16) Fleiss, J. L. *Statistical Methods for Rates and Proportions*, 2nd ed.; John Wiley & Sons: New York, 1981; p 236.
(17) Lenovo. http://shop.lenovo.com/ISS_Static/WW/EMAT/products/workstation/thinkstation/en/pdf/C20-C20x_DS_EN.pdf (accessed on September 14, 2011).
(18) Microsoft. http://windows.microsoft.com/en-US/windows7/help (accessed on September 14, 2011).
(19) Intel. http://ark.intel.com/products/47922/Intel-Xeon-Processor-X5650-(12M-Cache-2_66-GHz-6_40-GTs-Intel-QPI) (accessed on September 14, 2011).
(20) Microsoft. http://msdn.microsoft.com/library/zw4w595w.aspx (accessed on September 14, 2011).
(21) Microsoft. http://msdn.microsoft.com/en-us/library/ms228593.aspx (accessed on September 14, 2011).
(22) Ekins, S.; Nikolsky, Y.; Bugrim, A.; Kirillov, E.; Nikolskaya, T. Pathway mapping tools for analysis of high content data. *Methods Mol. Biol.* **2007**, *356*, 319−350.
(23) Rayasam, G. V.; Tulasi, V. K.; Sodhi, R.; Davis, J. A.; Ray, A. Glycogen synthase kinase 3: more than a namesake. *Br. J. Pharmacol.* **2009**, *156*, 885−898.
(24) Lesuisse, D.; Tiraboschi, G.; Krick, A.; Abecassis, P.-Y.; Dutruc-Rosset, G.; Babin, D.; Halley, F.; Châtreau, F.; Lachaud, S.; Chevalier, A.; Quarteronet, D.; Burgevin, M.-C.; Amara, C.; Bertrand, P.; Rooney, T. Design of potent and selective GSK3β inhibitors with acceptable safety profile and pharmacokinetics. *Bioorg. Med. Chem. Lett.* **2010**, *20* (7), 2344−2349.
(25) Lesuisse, D.; Dutruc-Rosset, G.; Tiraboschi, G.; Dreyer, M. K.; Maignan, S.; Chevalier, A.; Halley, F.; Bertrand, P.; Burgevin, M.-C.; Quarteronet, D.; Rooney, T. Rational design of potent GSK3beta inhibitors with selectivity for Cdk1 and Cdk2. *Bioorg. Med. Chem. Lett.* **2010**, *20* (6), 1985−1989.
(26) Liao, X.; Zhang, L.; J. Thrasher, B.; Du, J.; Li, B. Glycogen synthase kinase-3beta suppression eliminates tumor necrosis factor-related apoptosis-inducing ligand resistance in prostate cancer. *Mol. Cancer Ther.* **2003**, *2*, 1215−1222.
(27) Beurel, E.; Blivet-Van Eggelpoël, M.-J.; Kornprobst, M.; Moritz, S.; Delelo, R.; Paye, F.; Housset, C.; Desbois-Mouthon, C. Glycogen synthase kinase-3 inhibitors augment TRAIL-induced apoptotic death in human hepatoma cells. *Biochem. Pharmacol.* **2009**, *77* (1), 54−65.
(28) Cantarella, G.; Di Benedetto, G.; Pezzino, S.; Risuglia, N.; Bernardini, R. TRAIL-related neurotoxicity implies interaction with the Wnt pathway in human neuronal cells in vitro. *J. Neurochem.* **2008**, *105*, 1915−1923.
(29) Beurel, E.; Jope, R. S. The Paradoxical Pro- and Anti-apoptotic Actions of GSK3 in the Intrinsic and Extrinsic Apoptosis Signaling Pathways. *Prog. Neurobiol.* **2006**, *79* (4), 173−189.
(30) Scaffidi, C.; Fulda, S.; Srinivasan, A.; Friesen, C.; Li, F.; Tomaselli, K. J.; Debatin, K.-M.; Krammer, P. H.; Peter, M. E. Two CD95 (APO-1/Fas) signaling pathways. *EMBO J.* **1998**, *17* (6), 1675−1687.
(31) Linseman, D. A.; Butts, B. D.; Precht, T. A.; Phelps, R. A.; Le, S. S.; Laessig, T. A.; Bouchard, R. J.; Florez-McClure, M. L.; Heidenreich, K. A. Glycogen synthase kinase-3beta phosphorylates Bax and promotes its mitochondrial localization during neuronal apoptosis. *J. Neurosci.* **2004**, *24*, 9993−10002.
(32) Maurer, U.; Charvet, C.; Wagman, A. S.; Dejardin, E.; Green, D. R. Glycogen synthase kinase-3 regulates mitochondrial outer membrane permeabilization and apoptosis by destabilization of MCL-1. *Mol. Cell* **2006**, *21* (6), 749−760.
(33) Bilim, V.; Ougolkov, A.; Yuuki, K.; Naito, S.; Kawazoe, H.; Muto, A.; Oya, M.; Billadeau, D.; Motoyama, T.; Tomita, Y. Glycogen synthase kinase-3: A new therapeutic target in renal cell carcinoma. *Br. J. Cancer* **2009**, *101*, 2005−2014.
(34) Yang, L.; Lin, C.; Liu, Z.-R. P68 RNA helicase mediates PDGF-induced epithelial mesenchymal transition by displacing Axin from beta-catenin. *Cell* **2006**, *127* (1), 139−155.
(35) Luo, J.; Chen, J.; Deng, Z.-L.; Luo, X.; Song, W.-X.; Sharff, K. A.; Tang, N.; Haydon, R. C.; Luu, H. H.; He, T. C. Wnt signaling and human diseases: What are the therapeutic implications? *Lab. Invest.* **2007**, *87* (2), 97−103.
(36) Senthivinayagam, S.; Mishra, P.; Paramasivam, S. K.; Yallapragada, S.; Chatterjee, M.; Wong, L.; Rana, A.; Rana, B. Caspase-mediated cleavage of β-catenin precedes drug-induced apoptosis in resistant cancer cells. *J. Biol. Chem.* **2009**, *284* (20), 13577−13588.
(37) Yost, C.; Torres, M.; Miller, J. R.; Huang, E.; Kimelman, D.; Moon, R. T. The axis-inducing activity, stability, and subcellular distribution of beta-catenin is regulated in Xenopus embryos by glycogen synthase kinase 3. *Genes Dev.* **1996**, *10* (12), 1443−1454.
(38) Wu, D.; Pan, W. GSK3: A multifaceted kinase in Wnt signaling. *Trends Biochem. Sci.* **2010**, *35* (3), 161−168.
(39) Polakis, P. The oncogenic activation of beta-catenin. *Curr. Opin. Genet. Dev.* **1999**, *9* (1), 15−21.
(40) Rodriguez, R.; Muller, S.; Yeoman, J. A.; Trentesaux, C.; Riou, J.-F.; Balasubramanian, S. A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres. *J. Am. Chem. Soc.* **2008**, *130* (47), 15758−15759.
(41) Ou, T.-M.; Lu, Y.-J.; Zhang, C.; Huang, Z.-S.; Wang, X.-D.; Tan, J.-H.; Chen, Y.; Ma, D.-L.; Wong, K.-Y.; Tang, J. C.-O.; Chan, A. S.-C.; Gu, L.-Q. Stabilization of G-Quadruplex DNA and Down-Regulation of Oncogene c-myc by Quindoline Derivatives. *J. Med. Chem.* **2007**, *50* (7), 1465−1474.
(42) Kuribayashi, K.; Krigsfeld, G.; Wang, W.; Xu, J.; Mayes, P. A.; Dicker, D. T.; Wu, G. S.; El-Deiry, W. S. TNFSF10 (TRAIL), a p53 target gene that mediates p53-dependent cell death. *Cancer Biol. Ther.* **2008**, *7* (12), 2034−2038.
(43) Ding, L.; Yuan, C.; Wei, F.; Wang, G.; Zhang, J.; Bellail, A. C.; Zhang, Z.; Olson, J. J.; Hao, C. Cisplatin Restores TRAIL apoptotic pathway in Glioblastoma-Derived Stem Cells through Up-regulation of DR5 and Down-regulation of c-FLIP. *Cancer Invest.* **2011**, *29* (8), 511−520.
(44) Vazquez, A.; Bond, E. E.; Levine, A. J.; Bond, G. L. The genetics of the p53 pathway, apoptosis and cancer therapy. *Nat. Rev. Drug Discovery* **2008**, *7*, 979−987.
(45) Urashima, M.; Teoh, G.; Chauhan, D.; Ogata, A.; Shirahama, S.; Kaihara, C.; Matsuzaki, M.; Matsushima, H.; Akiyama, M.; Yuza, Y.; Maekawa, K.; Anderson, K. C. MDM2 Protein Overexpression Inhibits Apoptosis of TF-1 Granulocyte-Macrophage Colony-Stimulating Factor-Dependent Acute Myeloblastic Leukemia Cells. *Blood* **1998**, *92* (3), 959−967.
(46) Rosas, M.; Birkenkamp, K. U.; Lammers, J.-W. J.; Koenderman, L.; Coffer, P. J. Cytokine mediated suppression of TF-1 apoptosis requires PI3K activation and inhibition of Bim expression. *FEBS Lett.* **2005**, *579* (1), 191−198.
(47) Zhu, Y. M.; Bradbury, D. A.; Russell, N. H. Wild-type p53 is required for apoptosis induced by growth factor deprivation in factor-dependent leukaemic cells. *Br. J. Cancer* **1994**, *69* (3), 468−472.
(48) Chen, K.-F.; Tai, W.-T.; Liu, T.-H.; Huang, H.-P.; Lin, Y.-C.; Shiau, C.-W.; Li, P.-K.; Chen, P.-J.; Cheng, A.-L. Sorafenib overcomes TRAIL resistance of hepatocellular carcinoma cells through the inhibition of STAT3. *Clin. Cancer Res.* **2010**, *16* (21), 5189−5199.
(49) Kim, Y.-S.; Jin, H.-O.; Seo, S.-K.; Woo, S. H.; Choe, T.-B.; An, S.; Hong, S.-I.; Lee, S.-J.; Lee, K.-H.; Park, I.-C. Sorafenib induces apoptotic cell death in human non-small cell lung cancer cells by down-regulating mammalian target of rapamycin (mTOR)-dependent survivin expression. *Biochem. Pharmacol.* **2011**, *82* (3), 216−226.
(50) Meng, X. W.; Lee, S.-H.; Dai, H.; Loegering, D.; Yu, C.; Flatten, K.; Schneider, P.; Dai, N. T.; Kumar, S. K.; Smith, B. D.; Karp, J. E.; Adjei, A. A.; Kaufmann, S. H. Mcl-1 as a buffer for proapoptotic Bcl-2 family members during TRAIL-induced apoptosis: A mechanistic basis for sorafenib (Bay 43-9006)-induced TRAIL sensitization. *J. Biol. Chem.* **2007**, *282* (41), 29831−29846.
(51) Rosato, R. R.; Almenara, J. A.; Coe, S.; Grant, S. The multikinase inhibitor sorafenib potentiates TRAIL lethality in human

leukemia cells in association with Mcl-1 and cFLIPL down-regulation. *Cancer Res.* **2007**, *67* (19), 9490−9500.
(52) Ding, W.; Cai, T.; Zhu, H.; Wu, R.; Tu, C.; Yang, L.; Lu, W.; He, Q.; Yang, B. Synergistic antitumor effect of TRAIL in combination with sunitinib in vitro and in vivo. *Cancer Lett.* **2010**, *293* (2), 158−166.
(53) Karasic, T. B.; Hei, T. K.; Ivanov, V. N. Disruption of IGF-1R signaling increases TRAIL-induced apoptosis: a new potential therapy for the treatment of melanoma. *Exp. Cell Res.* **2010**, *316* (12), 1994−2007.
(54) Pennarun, B.; Kleibeuker, J. H.; Oenema, T.; Stegehuis, J. H.; de Vries, E. G. E.; de Jong, S. Inhibition of IGF-1R-dependent PI3K activation sensitizes colon cancer cells specifically to DR5-mediated apoptosis but not to rhTRAIL. *Cell. Oncol.* **2011**, *34* (3), 245−259.
(55) Kim, M.; Liao, J.; Dowling, M. L.; Voong, K. R.; Parker, S. E.; Wang, S.; El-Deiry, W. S.; Kao, G. D. TRAIL inactivates the mitotic checkpoint and potentiates death induced by microtubule-targeting agents in human cancer cells. *Cancer Res.* **2008**, *68* (9), 3440−3449.
(56) Gasparian, M. E.; Domnina, L. V.; Ivanova, O. Y.; Izyumov, D. S.; Lomakin, A. Y.; Popova, E. N.; Yagolovich, A. V.; Pletjushkina, O. Y.; Dolgikh, D. A.; Chernyak, B. V. Cytoskeleton inhibitors combined with TRAIL induce apoptosis in HeLa carcinoma cells overexpressing antiapoptotic protein bcl-2. *Biochemistry (Moscow)* **2008**, *73* (3), 358−362.
(57) Dhanasekaran, D. N.; Reddy, E. P. JNK signaling in apoptosis. *Oncogene* **2008**, *27* (48), 6245−6251.
(58) Guan, B.; Yue, P.; Lotan, R.; Sun, S.-Y. Evidence that the human death receptor 4 is regulated by activator protein 1. *Oncogene* **2002**, *21* (20), 3121−3129.
(59) Mahalingam, D.; Keane, M.; Pirianov, G.; Mehmet, H.; Samali, A.; Szegezdi, E. Differential activation of JNK1 isoforms by TRAIL receptors modulate apoptosis of colon cancer cell lines. *Br. J. Cancer* **2009**, *100* (9), 1415−1424.
(60) Mucha, S. R.; Rizzani, A.; Gerbes, A. L.; Camaj, P.; Thasler, W. E.; Bruns, C. J.; Eichhorst, S. T.; Gallmeier, E.; Kolligs, F. T.; Göke, B.; De Toni, E. N. JNK inhibition sensitises hepatocellular carcinoma cells but not normal hepatocytes to the TNF-related apoptosis-inducing ligand. *Gut* **2009**, *58* (5), 688−698.
(61) Beurel, E.; Blivet-Van Eggelpoël, M.-J.; Kornprobst, M.; Moritz, S.; Delelo, R.; Paye, F.; Housset, C.; Desbois-Mouthon, C. GSK-3β inhibition by lithium confers resistance to chemotherapy-induced apoptosis through the repression of CD95 (Fas/APO-1) expression. *Biochem. Pharmacol.* **2009**, *77* (1), 54−64.
(62) Sánchez-Pérez, T.; Ortiz-Ferrón, G.; López-Rivas, A. Mitotic arrest and JNK-induced proteasomal degradation of FLIP and Mcl-1 are key events in the sensitization of breast tumor cells to TRAIL by antimicrotubule agents. *Cell Death Differ.* **2010**, *17* (5), 883−894.
(63) Vermeulen, K.; Berneman, Z. N.; Van Bockstaele, D. R. Cell cycle and apoptosis. *Cell Proliferation* **2003**, *36* (3), 165−175.
(64) Kim, E. H.; Kim, S. U.; Shin, D. Y.; Choi, K. S. Roscovitine sensitizes glioma cells to TRAIL-mediated apoptosis by down-regulation of survivin and XIAP. *Oncogene* **2004**, *23* (2), 446−456.
(65) Festa, M.; Petrella, A.; Alfano, S.; Parente, L. R-roscovitine sensitizes anaplastic thyroid carcinoma cells to TRAIL-induced apoptosis via regulation of IKK/NF-kappaB pathway. *Int. J. Cancer* **2009**, *124* (11), 2728−2736.
(66) Ortiz-Ferrón, G.; Yerbes, R.; Eramo, A.; López-Pérez, A. I.; De Maria, R.; López-Rivas, A. Roscovitine sensitizes breast cancer cells to TRAIL-induced apoptosis through a pleiotropic mechanism. *Cell Res.* **2008**, *18* (6), 664−676.
(67) Retzer-Lidl, M.; Schmid, R. M.; Schneider, G. Inhibition of CDK4 impairs proliferation of pancreatic cancer cells and sensitizes towards TRAIL-induced apoptosis via downregulation of survivin. *Int. J. Cancer* **2007**, *121* (1), 66−75.
(68) Ham, Y.-M.; Choi, K.-J.; Song, S.-Y.; Jin, Y.-H.; Chun, M.-W.; Lee, S.-K. Xylocydine, a Novel Inhibitor of Cyclin-Dependent Kinases, Prevents the Tumor Necrosis Factor-Related Apoptosis-Inducing Ligand-Induced Apoptotic Cell Death of SK-HEP-1 Cells. *J. Pharmacol. Exp. Ther.* **2004**, *308* (3), 814−819.
(69) Dey, A.; Tergaonkar, V.; Lane, D. P. Double-edged swords as cancer therapeutics: Simultaneously targeting p53 and NF-kappaB pathways. *Nat. Rev. Drug Discovery* **2008**, *7* (12), 1031−1040.
(70) Gross, A.; Li, R.; Majid, T.; Moorcroft, N.; Yu, K.; Zilberstein, A. Preparation of tetrahydrocarboline derivatives as Eg5 inhibitors for treating cell proliferation disorders. PCT Int. Appl. WO 2011084439, Dec 15, 2010; *Chem. Abstr.* **2011**, *155*, 182041.
(71) Endow, S. A. Microtubule motors in spindle and chromosome motility. *Eur. J. Biochem.* **1999**, *262* (1), 12−18.
(72) Chin, G. M.; Herbst, R. Induction of apoptosis by monastrol, an inhibitor of the mitotic kinesin Eg5, is independent of the spindle checkpoint. *Mol. Cancer Ther.* **2006**, *5* (10), 2580−2591.
(73) *UniProt Knowledgebase (UniProtKB)*. UniProt release 2013_11. http://www.uniprot.org (accessed December 2, 2013).
(74) Flaherty, D. B.; Soria, J. P.; Tomasiewicz, H. G.; Wood, J. G. Phosphorylation of human Tau protein by microtubule-associated kinases: GSK3β and cdk5 are key participants. *J. Neurosci. Res.* **2000**, *62* (3), 463−472.
(75) Hannak, E.; Kirkham, M.; Hyman, A. A.; Oegema, K. Aurora-A kinase is required for centrosome maturation in Caenorhabditis elegans. *J. Cell Biol.* **2001**, *155* (7), 1109−1116.
(76) Wang, Y.; Bolton, E.; Dracheva, S.; Karapetyan, K.; Shoemaker, B. A.; Suzek, T. O.; Wang, J.; Xiao, J.; Zhang, J.; Bryant, S. H. An overview of the PubChem BioAssay resource. *Nucleic Acids Res.* **2010**, *38* (Jan), D 255−266.
(77) Mehta, C. R.; Senchaudhuri, P. Conditional versus Unconditional Exact Tests for Comparing Two Binomials. http://www.cytel.com/Papers/twobinomials.pdf (accessed September 22, 2013).
(78) Schaefer, C. F.; Anthony, K.; Krupa, S.; Buchoff, J.; Day, M.; Hannay, T.; Buetow, K. H. PID: The Pathway Interaction Database. *Nucleic Acids Res.* **2009**, *37*, D 674−679.
(79) Kanehisa, M.; Goto, S.; Sato, Y.; Furumichi, M.; Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* **2012**, *40* (Jan), D109−114.
(80) Kanehisa, M.; Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **2000**, *28* (1), 27−30.
(81) Garyantes, T.; Li, Z.; Yan, Y.; Anthony, J.; Yu, J. ; Zilberstein, A. DNA fragmentation assay using DNA intercalators. PCT Int. Appl. WO 2007149896, June 20, 2007; *Chem. Abstr.* **2007**, *148*, 96018.