Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Interdiscip Sci. Author manuscript; available in PMC 2019 June 01.

Published in final edited form as:

Interdiscip Sci. 2018 June ; 10(2): 449–454. doi:10.1007/s12539-016-0194-3.

A computational systems biology approach for identifying candidate drugs for repositioning for cardiovascular disease

Alvin Z. Yu[#a] and **Stephen A. Ramsey**[#a,#b,*]

[#a]Oregon State University, Department of Biomedical Sciences, 106 Dryden Hall, Corvallis, OR 97331, United States

[#b]Oregon State University, School of Electrical Engineering and Computer Science, 1148 Kelley Engineering Center, Corvallis, OR, 97331, United States

Abstract

We report an *in silico* method to screen for receptors or pathways that could be targeted to elicit beneficial transcriptional changes in a cellular model of a disease of interest. In our method we integrate: (i) a dataset of transcriptome responses of a cell line to a panel of drugs; (ii) two sets of genes for the disease; and (iii) mappings between drugs and the receptors or pathways that they target. We carried out a Gene Set Enrichment Analysis (GSEA) test for each of the two gene sets against a list of genes ordered by fold-change in response to a drug in a relevant cell line (HL60), with the overall score for a drug being the difference of the two enrichment scores. Next, we applied GSEA for drug targets based on drugs that have been ranked by their differential enrichment scores. The method ranks drugs by the degree of anticorrelation of their gene-level transcriptional effects on the cell line with the genes in the disease gene sets. We applied the method to data from (i) CMap 2.0; (ii) gene sets from two transcriptome profiling studies of atherosclerosis; and (iii) a combined dataset of drug/target information. Our analysis recapitulated known targets related to CVD (e.g., PPARγ; HMG-CoA reductase, HDACs) and novel targets (e.g., amine oxidase A, δ-opioid receptor). We conclude that combining disease-associated gene sets, drug-transcriptome-responses datasets and drug-target annotations can potentially be useful as a screening tool for diseases that lack an accepted cellular model for *in vitro* screening.

Keywords

atherosclerosis; gene expression analysis; drug repositioning; bioinformatics

Introduction

Due to functional pleiotropy of drug targets and poly-pharmacological drug-target mechanisms, many approved drugs likely have undiscovered therapeutic applications. In light of the significant cost-efficiencies of bringing an already approved drug to market under a new indication (i.e., drug repositioning; [1]), the recent availability of

* stephen.ramsey@oregonstate.edu; Tel. +1-541-737-5609; Fax +1-541-737-2730 .

Software availability

The software for this project is available under a free-software, open-source license at github.com/ramseylab/drugscan.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

comprehensive transcriptome profiling data for the effects of drugs on cell lines [2] has spurred interest in computationally screening for new indications for existing drugs, i.e., *in silico* drug repositioning [3].

Cardiovascular diseases (CVD) and their main underlying pathology, the chronic inflammatory disease atherosclerosis, are together the leading cause of death. To the extent that current lipid-lowering drugs for CVD prevention are estimated to reduce CVD mortality by only 20% [4], there is a need for new therapeutic approaches. Atherosclerosis is particularly attractive for a computational approach because there is not a natural *in vitro* cellular assay for drug-to-*phenotype* screening for this disease. A key cellular constituent in atherosclerotic plaque, the macrophage (an innate immune cell of the myeloid lineage), is both an enticing therapeutic target and has an analogous human myeloid cell line, HL60, that is a workhorse cell line in pharmacology related to hematopoiesis [5].

In this work, we investigated whether candidate atheroprotective or cardioprotective drugs can be identified by applying a rank-based statistical test (Gene Set Enrichment Analysis, or GSEA; [6]) to measurements of drug-induced differential expression of atherosclerosis-related genes ("gene sets") in a physiologically relevant human cell line. For the transcriptome profiling data on cell line drug responses, we used measurements from the Connectivity Map 2.0 (CMap2) database [2] for differential expression of 12,135 genes in HL60 cells that were treated with vehicle or one of 1,229 drugs. We used CVD-related gene sets from two transcriptome studies of human tissues: a study comparing three types of atherosclerotic arteries (aorta, coronary, and carotid) with normal arteries [7] (Cagnin *et al.,* 161 genes), and a study comparing unstable vs. stable carotid plaques as determined by specific molecular markers [8] (Puig *et al.,* 1,271 genes). Thus, our analysis included four sets of genes; two for genes that are up- or down-regulated in atherogenesis, and two for genes that are up- or down-regulated as plaque becomes unstable. Using the gene sets and the CMap2 data, we screened for drugs that reduced HL60 expression of genes in the "upregulated" gene sets and increased expression of genes in the "downregulated" gene sets, as measured by enrichment scores. We used a novel permutation-based approach to assess the significance of each enrichment score. We based our choice of weight factor for the analysis on accuracy results that we obtained by applying GSEA and a weighted Kolmogorov-Smirnov test with various weights to simulated data (the first such analysis of which we are aware).

Materials and Methods

Connectivity Map analysis

We obtained probe intensity files (CEL files; Affymetrix HG-U133A and HT_HG-U133A GeneChips) for HL60 experiments spanning 1,229 drugs (1,406 files in all) from the Connectivity Map 2.0 website (broadinstitute.org/cmap). We mapped probe intensities to 12,135 probesets using the Entrez Gene-based probesets from the University of Michigan Custom CDF project (brainarray.mbni.med.umich.edu) release 18.0.0, and we obtained probeset-level intensities using the justRMA function in the Bioconductor software package "affy". For each comparison of a drug to vehicle, we selected only probesets for which the within-sample-group-average $\log_2$ intensity is at least six (the background hybridization

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

signal level) in either the drug or vehicle sample group. For each probeset, we computed the average $\log_2$ ratio of the intensity in drug-treated to vehicle-treated HL60 cells. Then, for each drug, we ranked all above-background probesets (genes) based on their $\log_2$ ratios. All subsequent analyses were carried out in the R statistical computing environment.

Synthetic dataset analysis to determine optimal weighting

For the synthetic dataset analysis, we generated 1,000 "positive" gene set ranks and 1,000 "control" gene set ranks (each of size 80 genes). For both types of gene sets, rank assignments were randomly sampled from {1, … , 1,229} without replacement. Rank assignments were selected with uniform probability for the case of a "control" gene set, and rank assignments for "positive" gene sets were selected with a bias probability for rank j defined by $p_j = j^{-q}/\Sigma_{j'}(j')^{-q}$, where the values of q tested were 0.1, 0.2, 0.3, and 0.4. Expression ratios r_j for the set of 8,000 genes were obtained from the average $\log_2$ ratios at each rank, across all drugs. The enrichment score E was computed for each of the 1,000 gene sets, and gene sets were ordered by the E scores and the area under the sensitivity vs. false positive error rate (i.e., area under the receiver operating characteristic) curve was computed using the ROCR software package.

Selection of disease-associated gene sets

We obtained genes from Table S2 of the Cagnin *et al.* article [7] that were classified as "up" (64 genes) or "down" (97 genes), for the gene sets for coronary artery disease vs. normal arteries. From Table S3 of the Puig *et al.* study [8], we selected genes that were classified as "inflamed" (900 genes) for the "up" gene set, and genes that were classified as "stable" (371 genes) for the "down" gene set. In both cases, genes were selected based on the differential expression analysis in the original study (FDR < 0.05 for the Cagnin *et al.* study and FDR < 0.12 for the Puig *et al.* study). The HGNC identifiers from the Cagnin *et al.* and the Puig *et al.* studies were converted to Entrez Gene identifiers using the Web-based tool DAVID (david.ncifcrf.gov).

Permutation method for computing *P* values

For each drug and gene set resulting in a E score, we computed E(random) scores for 1,000 randomly generated gene sets. For each random gene set, the genes' ranks were sampled (uniform probability, no replacement). We computed the empirical P value of E as the cCDF of E in the distribution of E (random), using kernel density estimation (kCDF function in the sROC R package). For each pair of gene sets, a P value cutoff was determined at which the estimated false discovery rate would be 0.05 [9]. The resulting P value cutoffs for the gene set pairs based on the Cagnin *et al.* [7] and the Puig *et al.* [8] gene sets were $P = 0.017$ and $P = 0.03$, respectively.

TTD and DrugBank databases

Drug-target data matrices were downloaded from two databases, DrugBank and Therapeutic Target Database (TTD), and matched to the drugs from the Connectivity Map. Drugs were matched using two different strategies, by its given name and by its CAS number. CAS number for each drug from the Connectivity Map were collected from DrugBank, Sigma-

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Aldrich and ChemSpider websites. Overall, 525 of 1,087 drugs matched to DrugBank's database and 568 of 1,087 matched to TTD's database. When combined, 604 of 1,087 drugs were mapped to at least one target. Each drug target that was associated with at least three drugs was tested against the ranked list of drugs using GSEA.

Results and Discussion

Our method (Fig. 1) entails generating an overall disease association score ΔE for a given drug, based on two disease-related gene sets S^{up} and S^{down} and based on the expression $\log_2$-ratios r_i of N genes ($i \in \{1,\ldots, N\}$) in drug-treated vs. vehicle-treated cells, where $g_j \in \{1,\ldots, N\}$ is the gene whose expression $\log_2$ ratio r_{g_j} has rank j (rank 1 means highest positive $\log_2$ ratio), as described below. For each gene set S (where S is either the set S^{up} for genes that are upregulated in disease vs. normal tissue, or the gene set S^{down} for genes that are downregulated in disease), we compute a weighted enrichment statistic $E(S)$ by computing, for all $i \in \{1,\ldots, N = |S|\}$,

$$P_i^{\text{miss}}(S) = \sum_{\substack{g_j \notin S \\ j \le i}} \frac{|r_j|^{zw}}{N_M}; \qquad P_i^{\text{hit}}(S) = \sum_{\substack{g_j \in S \\ j \le i}} \frac{|r_j|^{w}}{N_R},$$

where $w \ge 0$ is the weight factor for the expression ratio, the case $z = 0$ corresponds to the GSEA method, the case $z = 1$ corresponds to a weighted Kolmogorov-Smirnov test, and where the normalization factors N_M and N_R are defined by

$$N_M = \Sigma_{g_j \notin S} |r_j|^{zw}; \qquad N_R = \Sigma_{g_j \in S} |r_j|^{w}.$$

We compute the enrichment score $E(S)$ by the maximum deviation from zero,

$$E(S) = \text{extreme}_i\left(P_i^{\text{hit}}(S) - P_i^{\text{miss}}(S)\right).$$

.

The overall drug-to-disease association score is the difference in the enrichment score between the two disease-associated gene sets,

$$\Delta E = E\left(S^{\text{up}}\right) - E\left(S^{\text{down}}\right).$$

Based on uncertainty in the literature regarding the optimal selection of the per-gene weight factor w for the statistical test (where $w = 0$ corresponds to the unweighted test) and on whether the unweighted test ($z = 0$) is more accurate than a weighted Kolmogorov-Smirnov (WKS) or weighted GSEA test ($z = 1$) [6], we simulated drug-response data and gene-set rankings under the null hypothesis (i.e., that the gene set is unrelated to the drug response) and under the alternate hypothesis (see Materials & Methods). We found that the *unweighted*

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

GSEA (i.e., $z = 0$, $w = 0$) test was the most accurate ($P < 10^{-4}$ for all pairwise comparisons with unweighted GSEA; Table 1), and thus we chose $w = 0$ and $z = 0$ as the parameter values.

To rank drugs by their potential therapeutic benefit in the context of a specific pathophysiological process (SPP) (atherogenesis for the Cagnin *et al.* gene sets; and plaque destabilization for the Puig *et al.* gene sets) [7, 8], we used a drug's ΔE score *and* P value. Specifically, for each SPP, we ranked drugs by their ΔE values, such that a top-ranked drug would have effects on HL60 cells that are the most anti-correlated with the "directions" of the gene sets for the SPP. Additionally, for each drug and gene set, we tested the null hypothesis that there is no association between the gene set and the drug response by computing E for the gene set for randomly permuted gene ranks, from which we empirically determined a P value. We eliminated any drug whose P values for both of the atherogenesis [7] gene sets did not satisfy $P < 0.017$ or whose P values for both of the destabilization [8] sets did not satisfy $P < 0.03$ (reflecting the latter SPP's larger gene set size). For each remaining drug, we averaged its ranks for the two SPPs (atherogenesis and plaque destabilization) to obtain an overall drug ranking; the top ten drugs are shown in Table 2 (the complete set of quantitative results for all drugs is provided as Online Resource 1).

Based on the observation that the two top-scoring drugs share a common mechanism (histone deacetylase inhibition), we investigated whether there are other drug target receptors or pathways that are enriched among the top-scoring drugs. For each drug, we obtained its list of annotated target receptors and pathways, based on information from two databases, DrugBank [10] and TTD [11]. We then applied the unweighted GSEA method to the ranks of drugs that are associated with a specific target, vs. the ranks of the complete set of drugs, to obtain an E score for the drug target (i.e., the receptor or pathway). The top-scoring drug targets (Table 3) include PPARγ (whose pharmacological activation has been shown to reduce heart attacks [12]) and HMG CoA-reductase (the rate-limiting enzyme in cholesterol biosynthesis, and the primary target of the statin class of drugs for CVD prevention and treatment; this finding is also consistent with reports that statins exert anti-inflammatory effects in macrophages [13]). Other specific molecular targets, like amine oxidase [14], cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A [15], and phospholipase A2 [16], are also reported to have roles in atherosclerosis.

Conclusions

Our finding that two of the top-scoring drugs (ranked by their transcriptional effects in HL60 cells) are HDAC inhibitors is intriguing in light of the broad range of pathophysiological processes that involve HDACs in CVD [17,18]. Of relevance to atherosclerosis, in macrophages, regulation of cholesterol metabolic enzymes in response to statins is mediated by HDAC repression that can be recapitulated by trichostatin A treatment [19]. More generally, HDACs have been reported to have anti-inflammatory effects on macrophages [20]. The other HDAC inhibitor, valproic acid, has recently been demonstrated to attenuate atherosclerosis in a mouse model [21]. While in this example application of our approach we analyzed the effects of drugs on expression levels of protein-coding genes, our approach could in principle be readily applied to screen for drugs that modulate expression levels of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

disease-associated microRNAs, (which are of significant interest in drug discovery for cancer and other diseases [22–24] using data from a drug-microRNA expression database such as Pharmaco-miR [25].

Our analysis shows that our computational approach, including the novel permutation test that we have introduced, is both practical and effective for identifying novel potential drugs and drug targets for two pathophysiological processes related to CVD (atherosclerosis and plaque destabilization). Our work also provides evidence in support of using unweighted GSEA for *in silico* drug screening, and our overall computational approach is readily applicable to other disease contexts.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by the US National Institutes of Health (award HL098807 to S.A.R.), the Medical Research Foundation of Oregon (New Investigator Grant award to S.A.R.), Oregon State University (Division of Health Sciences Interdisciplinary Research Grant award to S.A.R. and University Honors College DeLoach Work Scholarship to A.Y.), the National Science Foundation (award numbers 1557605-DMS and 1553728-DBI to S.A.R.), and the Oregon State University Center for Genome Research and Biocomputing.

References

1. Yarchoan M, Arnold S. Repurposing Diabetes Drugs for Brain Insulin Resistance in Alzheimer Disease. Diabetes. 2014; 63:2253–2261. [PubMed: 24931035]
2. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. Science. 2006; 313:1929–1935. [PubMed: 17008526]
3. Li J, Zheng S, Chen B, Butte AJ, Swamidass SJ, Lu Z. A survey of current trends in computational drug repurposing. Brief Bioinform. 2015 ePub ahead of print.
4. Kostis WJ, Cheng JQ, Dobrzynski JM, Cabrera J, Kostis JB. J Am Coll Cardiol. 2012; 59(6):572–582. [PubMed: 22300691]
5. Sun H, Wang Y. Apoptosis of human leukemic HL-60 cells induced to differentiate by treatment with RA or DMSO. Cell Res. 1995; 5:181–186.
6. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005; 102:15545–50. [PubMed: 16199517]
7. Cagnin S, Biscuola M, Patuzzo C, Trabetti E, Pasquali A, et al. Reconstruction and functional analysis of altered molecular pathways in human atherosclerotic arteries. BMC Genomics. 2009; 10:13. G. [PubMed: 19134193]
8. Puig O, Yuan J, Stepaniants S, Zieba R, Zycband E, et al. A gene expression signature that classifies human atherosclerotic plaque by relative inflammation status. Circ Cardiovasc Genet. 2011; 4:595–604. [PubMed: 22010137]
9. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J Roy Stat Soc B. Dec 1; 1994 57(1):289–300.
10. Law V, Knox C, Djoumbou Y, Jewison T, Guo AC, et al. DrugBank 4.0: shedding new light on drug metabolism. Nucl Acids Res. 2014; 42:1091–1097.
11. Zhu F, Shi Z, Qin C, Tao L, Liu X, et al. Therapeutic Target Database update 2012 : a resource for facilitating target-oriented drug discovery. Nucl Acids Res. 2012; 40:1128–1136.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

12. Dormandy JA, Charbonnel B, Eckland DJA, Erdmann E, Massi-Benedetti M, Moules IK, et al. Secondary prevention of macrovascular events in patients with type 2 diabetes in the PROactive Study (PROspective pioglitAzone Clinical Trial In macroVascular Events): a randomised controlled trial. Lancet. 2005; 366(9493):1279–89. [PubMed: 16214598]
13. Tuomisto TT, Lumivuori H, Kansanen E, Hakkinen SK, Turunen MP, van Thienen JV, et al. Simvastatin has an anti-inflammatory effect on macrophages via upregulation of an atheroprotective transcription factor, Kruppel-like factor 2. Cardiovasc Res. 2008; 78(1):175–84. [PubMed: 18192240]
14. Meszaros Z, Karadi I, Csanyi A, Szombathy T, Romics L, Magyar K. Determination of human serum semicarbazide-sensitive amine oxidase acitivity: a possible clinical marker of atherosclerosis. Eur J Drug Metab Pharmacokinet. 1999; 24:299–302. [PubMed: 10892891]
15. Aizawa T, Wei H, Miano JM, Abe J, Berk BC, Yan C. Role of Phosphodiesterase 3 in NO/cGMP-Mediated Antiinflammatory Effects in Vascular Smooth Muscle Cells. Circ Res. 2003; 93:406–413. [PubMed: 12919948]
16. Zalewski A, Macphee C. Novel Approaches to the Treatment of Dyslipidemia Role of Lipoprotein-Associated Phospholipase A2 in Atherosclerosis Biology, Epidemiology , and Possible Therapeutic Target. Arterioscler Thromb Vasc Biol. 2005; 25(5):923–31. [PubMed: 15731492]
17. Eom GH, Kook H. Posttranslational modifications of histone deacetylases: implications for cardiovascular diseases. Pharmacol Therap. 2014; 143(2):168–80. [PubMed: 24594235]
18. McKinsey T. Therapeutic potential for HDAC inhibitors in the heart. Annu Rev Pharmacol Toxicol. 2012; 52:303–319. [PubMed: 21942627]
19. Feig JE, Shang Y, Rotllan N, Vengrenyuk Y, Wu C, Shamir R, et al. Statins Promote the Regression of Atherosclerosis via Activation of the CCR7-Dependent Emigration Pathway in Macrophages. PLoS ONE. 2011; 6(12):e28534. [PubMed: 22163030]
20. Wu C, Li A, Leng Y, Li Y, Kang J. Histone Deacetylase Inhibition by Sodium Valproate Regulates Polarization of Macrophage Subsets. DNA Cell Biol. 2012; 31:592–599. [PubMed: 22054065]
21. Bowes AJ, Khan MI, Shi Y, Robertson L, Werstuck GH. Valproate attenuates accelerated atherosclerosis in hyperglycemic apoE-deficient mice: evidence in support of a role for endoplasmic reticulum stress and glycogen synthase kinase-3 in lesion development and hepatic steatosis. Am J Pathol. 2009; 174:330–342. [PubMed: 19095952]
22. Zeng X, Zhang X, Zou Q. Integrative approaches for predicting microRNA function and prioritizing disease-related microRNA using biological interaction networks. Brief Bioinformatics. 2016; 17:193–203. [PubMed: 26059461]
23. Liu Y, Zeng X, He Z, Zou Q. Inferring microRNA-disease associations by random walk on a heterogeneous network with multiple data sources. IEEE/ACM Trans Comput Biol and Bioinf. 2016; ePub ahead of print. doi: 10.1109/TCBB.2016.2550432
24. Wang Q, Wei L, Guan X, et al. Briefing in family characteristics of microRNAs and their applications in cancer research. Biochim Biophys Acta. 2014; 1844:191–197. [PubMed: 23954304]
25. Rukov JL, Wilentzik R, Jaffe I, et al. Pharmaco-miR: linking microRNAs and drug effects. Briefings in Bioinformatics. 2014; 15:648–659. DOI: 10.1093/bib/bbs082 [PubMed: 23376192]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

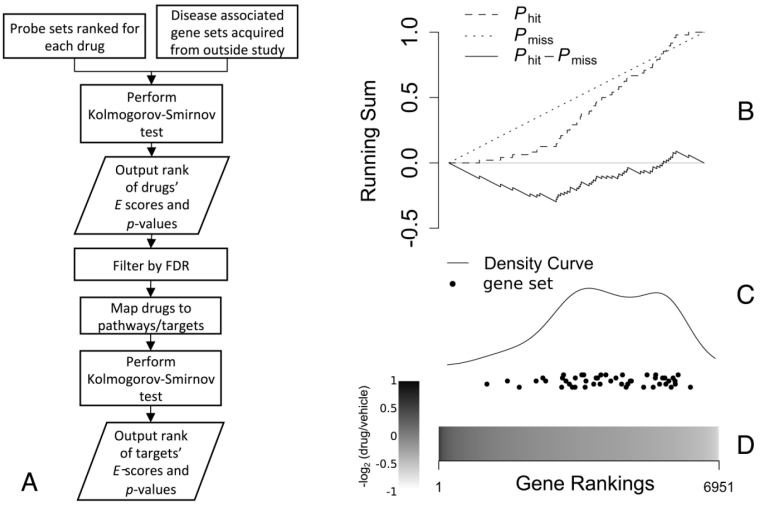

Figure 1. How we score a drug's gene expression response against a disease-associated gene set

(A) Flow-chart of the overall analysis workflow. (B) Example erichment score analysis test of gene expression responses of HL60 cells to the drug oxolinic acid when tested with a a coronary artery disease gene set. (C) Rank positions of coronary artery disease genes among all genes, ranked by differential expression in response to the drug oxolinic acid in HL60 cells. Line plot, kernel density of the gene ranks. (D) Grayscale intensity scale bar for the $-\log_2$ gene expression ratios for oxolinic acid-treated vs. vehicle-treated HL60 cells.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Accuracy results for detecting drug-disease associations within a synthetic dataset, for weighted Kolmogorov-Smirnov (K-S) and GSEA tests

Column headers are as follows: "Weighted KS", weighted K-S test; AUC, mean area under the curve for sensitivity vs. false positive error rate; AUC s.d., standard deviation of the AUC. The difference between the mean AUCs for the unweighted and GSEA ($w = 0.25$) test is significant at $P < 10^{-4}$ (paired Student's t-test). The comparison of the unweighted and weighted $w = 0.25$ are also significant ($P < 10^{-4}$). Shown here are the results for the rank bias parameter value $q = 0.2$ (see Materials & Methods), but the relative accuracy results using all other values of q were consistent with the above results.

Test Type	*z*	*w*	AUC	AUC s.d.
Weighted KS	1	1	0.9615	0.0058
Weighted KS	1	0.5	0.9855	0.0025
Weighted KS	1	0.25	0.9906	0.0016
GSEA	0	1	0.9646	0.0054
GSEA	0	0.5	0.9864	0.0024
GSEA	0	0.25	0.9908	0.0017
Unweighted	0	0	**0.9922**	0.0014

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 2

The ten top-ranked drugs based on our drug repositioning screen using measured drug responses in HL60 cells

Drugs' Δ*E* score ranks were separately computed for the two SPPs and then and averaged. HDAC inhibitors (denoted by *), scored well in the rankings. The lowest rank corresponds to the highest significance.

Drug Name	CAS Identifier	Cagnin Δ*E* score	Cagnin Up *P* value	Cagnin Down *P* value	Puig up *P* value	Puig down *P* value	Puig Δ*E* score	Cagnin Rank	Puig Rank	Average Rank
trichostatin A(*) ≠	58880-19-6	−0.583	5.66×10^{-17}	4.58×10^{-35}	5.82×10^{-22}	7.08×10^{-66}	−0.270	8	13	10.5
valproic acid (*)	99-66-1	−0.523	1.34×10^{-8}	9.66×10^{-5}	5.08×10^{-61}	1.60×10^{-45}	−0.290	19	6	12.5
sulfaguanidine	57-67-0	−0.537	1.14×10^{-3}	1.29×10^{-11}	6.15×10^{-37}	1.08×10^{-42}	−0.274	15	11	13.0
mafenide	7761-27-5	−0.634	5.81×10^{-10}	2.62×10^{-79}	5.37×10^{-76}	2.41×10^{-3}	−0.223	3	46	24.5
dehydrocholic acid	81-23-2	−0.513	6.97×10^{-10}	1.77×10^{-4}	1.54×10^{-41}	7.91×10^{-13}	−0.243	23	29	26.0
trichostatin A (*) †	58880-19-6	−0.494	8.73×10^{-4}	1.74×10^{-8}	1.60×10^{-3}	2.13×10^{-50}	−0.244	27	28	27.5
hydrastinine	6592-85-4	−0.535	1.58×10^{-2}	2.43×10^{-52}	9.06×10^{-17}	3.22×10^{-5}	−0.228	16	41	28.5
Trolox C	53188-07-1	−0.465	3.05×10^{-3}	3.65×10^{-5}	1.91×10^{-15}	1.40×10^{-33}	−0.245	38	27	32.5
harman	486-84-0	−0.455	1.93×10^{-9}	3.79×10^{-3}	7.76×10^{-131}	8.01×10^{-14}	−0.259	46	21	33.5
brinzolamide	138890-62-7	−0.637	6.58×10^{-4}	1.31×10^{-80}	1.61×10^{-69}	1.16×10^{-2}	−0.203	2	83	42.5

≠For trichostatin A, = drug from Sigma;

†= drug from CalBiochem.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

High-scoring drug targets, based on an unweighted GSEA test of a drug target against the complete list of drugs ranked by their *E* scores for CVD (for $E > 0.5$).

Target	*E* score	# of drugs
Amine oxidase [flavin-containing] A	0.736	3
Peroxisome proliferator activated receptor gamma	0.701	3
3-hydroxy-3-methylglutaryl-coenzyme A reductase	0.647	3
cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A	0.594	3
Phospholipase A2	0.593	3
Cholinesterase	0.584	3
δ-opioid receptor	0.553	3