The Relative Transcription Index: A Gene Expression Based Metric for Prioritization of Drug Candidates

Sujoy Ghosh*, Mike A. Watson and Jon L. Collins

GlaxoSmithKline, 5 Moore Drive, Research Triangle Park, NC 27709, USA

Abstract: Efficient compound selection remains a key challenge in drug discovery today. The goal is to identify developable drug candidates early in the screening process while simultaneously flagging compounds with off-target effects indicative of liabilities or alternate indications. This goal overlaps but is distinct from the goal of toxicogenomics which is focused primarily on identifying toxicity signatures of lead candidates in key tissues. We propose a framework where global changes in gene expression levels in response to compounds can be used as an objective metric for early compound prioritization. We call this metric the Relative Transcription Index (RTI). RTI is a measure of the relative activity of compounds as ascertained by their effects on transcription at a genome-wide level. Compounds with a low RTI affect the expression of only a few genes whereas compounds with a high RTI affect the expression of a large number of genes. This information is useful for differentiating compounds that, based on phenotypic assays alone, may appear to be equally efficacious. Since compounds with high RTI are more likely to display off-target effects, the RTI metric, if implemented early in the screening process, can become a valuable tool for compound selection. The utility of the RTI metric is demonstrated by its application to two different gene expression datasets - one involving modulators of the liver X receptor (LXR) and the other concerning antibacterial compounds belonging to diverse mechanistic classes.

Keywords: Index, global gene expression, compound prioritization, screening, drug discovery.

INTRODUCTION

Efficient compound selection is a key priority and challenge facing drug discovery today [1]. The goal is to identify developable drug candidates early in the screening process while simultaneously flagging and possibly eliminating compounds with potential off-target effects and other liabilities. Ever-escalating costs for the discovery and development of pharmaceuticals make it imperative that evidence-based compound selection and prioritization take place as early as is reasonably possible. Although some attempts have been made for the high-throughput physicochemical profiling of compounds [9], the majority of existing methods for determining the efficacy and safety of newly synthesized compounds depend on user-defined assays, typically *in vitro* or in cell culture systems. They usually measure single or a handful of endpoints and provide valuable but limited information about compound function. This is due to the fact that the assays only inform on the measured endpoints but provide no information on the general effects of compounds. Such general effects often manifest later as off-target effects leading to either compound liability or alternate indications for the compound. A well-known example of an off-target drug effect is that of induction of cytochrome P450 genes by the anti-tuberculosis drug rifampicin *via* the activation of the pregnane X receptor [19]. The intended target of rifampicin is RNA polymerase [20]. However, induction of cytochrome P450 genes often leads to excessive metabolism of the drug leading to suboptimal or failed treatment. Open-ended measurements that can measure several hundreds or thousands of endpoints simultaneously are needed to gain an understanding of general compound effects. One way of acquiring such information is by the incorporation of 'omics' based technologies early in the drug discovery process [4, 8]. These technologies embrace methods adapted to the measurement of the full or near-full complement of biomolecules in cells including RNA (transcriptomics), proteins (proteomics), lipids (lipidomics) and other metabolites (metabonomics). The data generated from these technologies can be characterized as system response profiles [5-7] and provide information on the general state of a biological system. Several laboratories have applied whole-genome transcriptomics to elucidate pharmacological or toxicological cellular responses to drugs [24-26]. In this article, we report on the application of transcriptomics for generating a gene expression based metric for the relative ranking of compounds. The principle of this approach is exemplified below through a class of compounds that act as modulators of the liver X receptor (LXR, 15) and on a class of antibacterial compounds.

LXR is a ligand-activated transcription factor that influences diverse endocrine functions [16, 17]. Activation of LXR by endogenous or exogenous ligands leads to well-characterized transcriptional responses dependent on the cellular context [10, 11]. In human hepatocytes, specific genes related to the synthesis and transport of lipids are induced (ABCG1, SREBP, FAS etc.) whereas in macrophages, pro-inflammatory genes such as interleukin 6 (IL-6) are down regulated [12, 13]. LXR modulator compounds differentially regulate the activity of LXR [14]. The efficacy of LXR modulators are typically measured *via* phenotypic assays such as measurements of triglyceride biosynthesis, cholesterol efflux or inhibition of interleukin 6 protein releases. In the current study, we profiled genome-wide transcriptional response in cultured human hepatocytes (HUH7) and macrophages (THP1) treated with 20 LXR modulators. Based on the global gene expression responses elicited by each compound, we created an index of transcription (Rela-

*Address correspondence to this author at GlaxoSmithKline, 5 Moore Drive, Research Triangle Park, NC 27709, USA; Tel: 919-483-0803; Fax: 919-315-3749; E-mail: sg45653@gsk.com

1386-2073/07 $50.00+.00

tive Transcription Index, RTI), relative to a benchmark LXR modulator, GW3965A, and used this index to score each compound based on their effects on gene expression. In other words, we determined if the compounds were modulating only the expected subset of LXR responsive genes or if they were inducing or repressing a larger set of genes probably by activating other targets beside LXR. Our results demonstrate that compounds that appear to be equally effective as LXR modulators (measured by phenotypic assays) often have dramatic differences in their effects on gene expression. Thus valuable information about compound characteristics that is not captured by the phenotypic assays is obtained from the RTI and provides a data-driven procedure for finer discrimination among compounds undergoing screening.

We extended the RTI algorithm to investigate a second, publicly available dataset [23] that measured the transcriptional response of over 4000 genes from the bacteria *Bacillus subtilis* to 37 different antibacterial compounds representing 6 major mechanisms of action classes (cell wall biosynthesis inhibitors, DNA topology inhibitors, fatty acid biosynthesis inhibitors, folate biosynthesis inhibitors, membrane-active ionophores and protein biosynthesis inhibitors). Our results point to a relation between the mechanistic class of a compound and it's RTI. Specifically, the protein biosynthesis inhibitor antibiotics elicit the greatest number of changes in gene expression compared to compounds belonging to other mechanistic classes (higher RTI).

METHODS

Compound Selection

A diverse set of 15 full and partial LXRα/β agonist chemotypes were selected based on data from a cell-free ligand-sensing assay (LiSA) using LXRα and LXRβ ligand-binding domains [18]. Activities and efficacies ranged from 5 μM-10 nM and 20-100 percent of GW3965, respectively. An additional exemplar from 5 of these chemotypes was selected to provide structure-activity information, giving rise to a total of 20 LXR ligands being profiled in the described experiments.

Cell-Based (Phenotypic) Assays

THP1 cells were differentiated with 40 ng/mL of 1α25-dihydroxy vitamin D3 (EMD Biosciences, San Diego, CA, USA) for 72 h in T225 flasks. The cells were then plated at $1x10^5$ cells per well of a 96-well plate, and compound was added in decreasing doses from 5 μM down to 2.3 nM final. After 6 h, the cells were stimulated with a final concentration of 100 ng/mL of LPS (Sigma-Aldrich, St. Louis, MO, USA) and incubated for an additional 18 h. The media were removed from the cells and assayed using a human IL-6 ELISA (R&D Systems, Minneapolis, MN, USA).

Human HUH7 cells were plated at $2.5x10^5$ cells per well in 96 well plates, and compound was added in decreasing doses from 5 μM down to 2.3 nM. Cells were incubated for 72 h. After the third day, media was removed and the cells lysed in Infinity triglyceride reagent (Thermo Electron, Melbourne, AUS) containing 0.01% digitonin (Sigma-Aldrich). Plates are incubated at 37 ^oC for 30-120 min and read at 30 min intervals on a plate reader at 540 nM.

Sample Preparation and Hybridization to Affymetrix Genechips

Total RNA samples from cell cultures were isolated by Trizol reagent (Invitrogen, Carlsbad, CA, USA) and purified on a RNeasy column (Qiagen, Valencia, CA, USA), prior to labeling for hybridization to Affymetrix U133A human arrays (Affymetrix, Santa Clara, CA, USA). Subsequent sample processing, labeling and hybridization of labeled nucleic acids to the Affymetrix chips were done as specified by the manufacturer [3]. Each compound treatment, including vehicle, was done in triplicate for each of the 2 cell-lines.

DATA PREPARATION

Gene expression signals from the hybridized Affymetrix microarray were generated *via* the MAS 5.0 software from Affymetrix. All expression data was normalized by global scaling to a trimmed average intensity of 150 per chip. We considered a minimum signal of 50 units in at least one of the treatments to be the inclusion criterion for probesets. Based on these criteria, we retained 8238 probesets from HUH7 samples and 7575 probesets from THP1 samples for further analysis. Typically one or more probesets correspond to known genes or expressed sequence tags.

A principal components analysis was performed on all samples (per cell type) using gene signals as the input variables and treated samples as observations. None of the samples behaved as outliers (data not shown), and consequently, the full set of samples was retained for analysis. Gene signals were logged (to base 2). For each treatment group, the average signal for each gene was computed from the individual logged signals.

STATISTICAL ANALYSIS

The statistical analysis described below details the procedure used for computing the Relative Transcription Index of compounds. For each gene we first computed a median log signal from the compound-wise average log signals for that gene. Then, for each gene, we determined the absolute deviation of the compound-wise average log signals from the gene's median log signal. Next we computed the median of the absolute deviation of the log signals, also known as MAD. From these values, we computed a compound-specific value, known as a robust z-score, for each gene according to the following relations:

$$z_{ij} = \frac{(x_{ij} - Median(x_i))}{1.4826 * MAD(x_i)}, \quad (1)$$

$$MAD(x_i) = Median(Abs(x_{ij} - median(x_{ij}))) \quad (2)$$

for $i = 1,2,3...,n$ genes ; $j = 1,2,3...,m$ compounds, where z_{ij} is the robust z-score for gene i and compound j, x_{ij} is the average log signal for gene i and compound j, x_i is the vector of average log signals for gene i. $Median(x_i)$ is the median log signal for gene i across all compounds and $MAD(x_i)$ is the median absolute deviation for gene i. The constant 1.4826 is a correction factor designed to make the estimator scale factor consistent with the usual scale parameter of a normal distribution, in this case the standard devia-

tion [2]. For a given compound, the robust z-score for a gene indicates the level of overexpression or underexpression of the gene compared to the average (median) expression of that gene across all compounds. The robust z-test heuristic states that an absolute z-score greater than 3.5 should be considered an extreme value or an outlier. Thus a compound with an absolute z-score greater than 3.5 for any gene indicates that the expression for that gene was significantly up or downregulated by that compound when compared to the expression of that gene for the other compounds. The number of genes with an absolute z-score greater than the threshold is used as a metric to label the overall ability of a compound to cause significant changes in gene expression. These numbers are then scaled in reference to the numbers observed for a common reference compound (or vehicle). We define these scaled values as the Relative Transcription Index (RTI) of the compounds. RTI for a compound is expressed as $RTI_j = \log_2\left(\frac{\#g_j \text{ with } |z|_{>3.5}}{\#g_{cref} \text{ with } |z|_{>3.5}}\right)$, where $\#g_j$ with $|z|_{>3.5}$ is the number of genes with absolute modified z-scores greater than 3.5 for compound j and $\#g_{cref}$ with $|z|_{>3.5}$ is the number of genes with absolute modified z-scores greater than 3.5 for the reference compound. The logarithm of the ratio is taken in order to equalize the scale of RTI values that are larger or smaller than that of the reference. The RTI value of the reference equals to zero after log transformation.

RESULTS

RTI assessment is independent of phenotypic efficacy assessments: Table **1** shows the Relative Transcription Index of the LXR modulators in the two cell lines assayed, along with phenotypic assay results. This allows one to compare relative agreement of a RTI score of a compound with its phenotypic score (obtained from a phenotypic assay). For the HUH7 system, the phenotypic score of a LXR modulator is a measure of its relative ability to induce triglyceride synthesis at a concentration of 500 nM compared to the reference compound (GW683965A). Fig. **1** is a plot comparing the

Table 1. Relative Transcription Index (RTI) and Phenotypic Assay Scores for LXR Modulator Compounds

Compound	THP1 Number of Outlier Genes ($\|z_i\|>3.5$)	THP1 RTI Score	THP1 Inhibition of IL-6[a] Release (% GW3965[b] Activity at 500 nM)	HUH7 Number of Outlier Genes ($\|z_i\|>3.5$)	HUH7 RTI Score	HUH7 TG[c] Synthesis (% GW3965 Activity at 500 nM)
Veh	48	0.13	**0**	10	0.00	0
C1	264	2.58	**ND**	1212	6.92	ND
C2	14	-1.65	**50**	490	5.61	ND
C3	342	2.96	**ND**	1059	6.73	ND
C4	32	-0.46	**25**	381	5.25	50
C5	15	-1.55	**90**	660	6.04	150
C6	9	-2.29	**90**	137	3.78	90
C7	19	-1.21	**20**	16	0.68	20
C8	95	1.11	**ND**	27	1.43	ND
C9	45	0.03	**55**	981	6.62	11
C10	114	1.37	**96**	5	-1.00	15
C11	33	-0.42	**90**	61	2.61	ND
C12[b]	44	0.00	**100**	10	0.00	100
C13	45	0.03	**100**	10	0.00	100
C14	50	0.18	**90**	7	-0.51	50
C15	167	1.92	**90**	78	2.96	10
C16	34	-0.37	**70**	8	-0.32	30
C17	35	-0.33	**40**	12	0.26	10
C18	87	0.98	**20**	22	1.14	90
C19	16	-1.46	**ND**	66	2.72	ND
C20	68	0.63	**90**	18	0.85	100

Results from each of the two cell lines (HUH7 and THP1) are presented. For each compound, the number of outlier genes (genes with absolute, robust z-score ≥ 3.5), the associated RTI score and the phenotypic assay score are presented.

[a]IL 6 = interleukin 6.

[b]Compound 12, GW683965A, was used as the reference compound.

[c]TG = triglyceride.

RTI scores and triglyceride synthesis scores for 15 compounds (the remaining 5 compounds do not have triglyceride synthesis scores and were therefore not plotted). Compounds C13, C18 and C20 are phenotypically similar to the reference compound and also have RTI scores close to that of the reference. However, this does not hold true in all cases. For example, compounds C5 and C6, which also appear phenotypically equivalent to the reference, have very high RTI scores indicating that they change expression of a much larger number of genes compared to the reference compound. Thus, based on the RTI scores, one can rank phenotypically similar compounds in the following order: C12,13,18,20 > C6 > C5 in order of increasing non-specific transcriptional effects. This ranking can then be used in compound prioritization and selection.

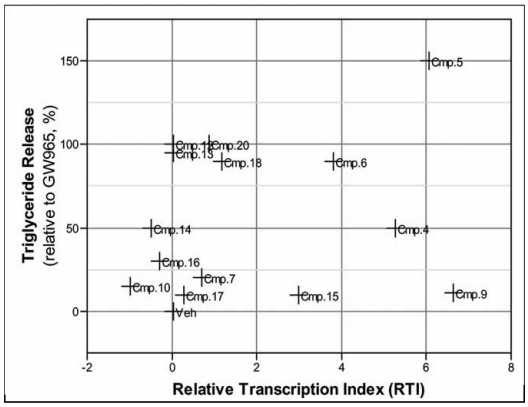

Fig. (1). Ranking of compounds based on RTI and phenotype assay scores from HUH7 cells. The phenotypic assay (triglyceride release relative to reference compound, GW683965) results for each compound is plotted on the Y-axis and the corresponding RTI scores are plotted on the X-axis. Compounds for which both phenotypic and RTI data exist are plotted.

RTI is context dependent: From Table **1**, a comparison among the compounds for the number of outlier genes ($|z| > 3.5$) and the corresponding RTI scores reveals context (cell line) specific differences. In the HUH7 system, compounds C2, C3, C4, C5, C6, and C10 upregulate a large number of genes compared to the reference compound (C12, GW683965A) and consequently have large positive RTI scores. Since the reference compound induces only 10 genes beyond the outlier range, there is not enough room to obtain large negative RTI scores for the other compounds relative to the reference. For the THP1 system, the positive and negative RTI scores are more evenly distributed (maximum 2.96 for C3 and minimum -2.29 for C6). This is due, in part, to the reference compound affecting the expression of a larger number of genes beyond the outlier range (89 genes).

Variability in the degree of overlap among outlier genes: In Fig. **2**, we performed a compound-by-compound comparison of the percent overlap between the outlier genes ($|z|>3.5$) generated by each compound. For each compound, the self-comparison value for percent overlap was set at 100%. Overall, the degree of overlap among any two compounds was relatively small (average overlap of 4.6% for THP1 system and 5.9% for HUH7 system). Only one exception was noted for a cluster of compounds in the HUH7 system. This cluster consisted of compounds C2, C4, C5 and C9 for which 60-73% of outlier genes were in common with each other.

RTI may be related to mechanistic class: In order to test the extent of applicability of the RTI, we performed an analysis on a different dataset containing transcription data from the bacterium *Bacillus subtilis* treated with 37 antibacterial compounds representing diverse mechanistic classes. In this dataset, we observed a class-dependence in the RTI scores. The transcription dataset was generated by Hutter *et al.* [23], and is publicly available over the World Wide Web. We computed the RTI for each compound relative to cefoxitin (an arbitrarily chosen reference) based on gene expression profiles observed at 40 min of treatment with the compounds. Fig. **3** shows the RTI for each compound, grouped according to its mechanism of action. According to Fig. **3**, compounds displaying high RTI are enriched in the protein biosynthesis inhibitor class, compared to any other classes. This finding demonstrates that in certain circumstances, RTI is also useful for ranking mechanism-based compound classes in addition to the ranking of individual compounds.

DISCUSSION

Current methods for compound screening and selection rely on directed phenotypic assays investigating a compound's efficacy and safety. These assays convey information about the selected targets and endpoints but fail to inform on off-target effects. These off-target effects are often the basis for adverse reactions or opportunities for alternate therapeutic applications. The earlier these off-target effects are identified and acted upon, the better it will impact the drug discovery process by picking the winners and losers early and reducing expensive attrition in the late, resource-intensive phases of drug development. Although the need for utilizing gene expression for compound selection and prioritization has been articulated [21, 22], an appropriate metric for translating that need into practice has not been reported. Moreover, most of the attention has been focused on predicting drug toxicity from gene expression (toxicogenomics) and not so much on the earlier phases of compound selection preceding lead optimization, when several molecules from multiple chemotypes are contenders and only a few can be progressed for structure-activity relationship studies.

To address the need for more insightful compound characterization early in drug discovery, we propose the Relative Transcription Index or RTI as a metric for compound characterization and prioritization. The RTI measures the impact of a compound on global gene expression and thereby provides a measure of general compound behavior. One application of the RTI is towards the characterization of a set of compounds directed against the same molecular target. Our results with LXR modulators show that the RTI scores for the compounds are context dependent and importantly, do not necessarily correlate with phenotypic efficacy scores. Also, compounds with similar RTI scores can do so by affecting

(**a**)

HUH	Veh	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20
Veh	100.0	30.0	20.0	50.0	30.0	20.0	0.0	0.0	0.0	30.0	0.0	0.0	0.0	0.0	0.0	0.0	10.0	20.0	0.0	0.0	0.0
C1		100.0	9.4	22.4	8.3	13.7	5.0	0.6	0.4	15.8	0.2	1.6	0.4	0.1	0.2	1.6	0.2	0.5	0.4	0.7	0.3
C2			100.0	15.9	54.5	61.2	3.5	0.0	0.8	73.9	0.2	0.4	0.2	0.0	0.0	0.2	0.2	0.6	0.2	0.8	0.0
C3				100.0	5.4	9.6	2.8	0.4	1.3	14.9	0.1	2.2	0.5	0.2	0.4	3.0	0.3	0.5	0.8	2.0	0.8
C4					100.0	60.4	5.8	0.0	0.8	74.3	0.3	1.3	0.3	0.0	0.0	1.0	0.5	0.5	0.3	1.0	0.3
C5						100.0	2.7	0.2	0.3	66.4	0.2	0.9	0.2	0.0	0.0	1.4	0.5	0.3	0.3	0.9	0.6
C6							100.0	0.7	1.5	24.8	0.7	4.4	0.7	0.0	0.0	2.2	0.0	0.0	2.2	2.9	0.0
C7								100.0	18.8	12.5	6.3	12.5	0.0	12.5	0.0	0.0	0.0	12.5	6.3	12.5	0.0
C8									100.0	14.8	0.0	11.1	3.7	3.7	3.7	14.8	0.0	0.0	0.0	18.5	3.7
C9										100.0	0.1	0.6	0.3	0.1	0.0	1.1	0.3	0.2	0.3	0.7	0.2
C10											100.0	0.0	0.0	20.0	20.0	0.0	0.0	20.0	20.0	0.0	0.0
C11												100.0	3.3	1.6	0.0	8.2	0.0	1.6	1.6	14.8	3.3
C12													100.0	10.0	0.0	10.0	0.0	20.0	0.0	0.0	10.0
C13														100.0	0.0	0.0	0.0	10.0	10.0	0.0	30.0
C14															100.0	14.3	0.0	0.0	0.0	0.0	0.0
C15																100.0	0.0	1.3	1.3	5.1	1.3
C16																	100.0	0.0	0.0	0.0	0.0
C17																		100.0	0.0	0.0	0.0
C18																			100.0	4.5	4.5
C19																				100.0	0.0
C20																					100.0

(**b**)

THP	Veh	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20
Veh	100.0	6.3	0.0	18.8	4.2	0.0	4.2	8.3	8.3	0.0	2.1	4.2	0.0	2.1	2.1	8.3	8.3	6.3	2.1	2.1	0.0
C1		100.0	0.4	53.4	1.5	0.4	0.0	0.0	2.7	1.1	1.5	3.0	0.4	1.1	1.1	5.3	0.0	2.3	1.9	1.1	1.1
C2			100.0	14.3	7.1	7.1	0.0	7.1	14.3	7.1	0.0	0.0	0.0	7.1	0.0	14.3	0.0	0.0	7.1	7.1	0.0
C3				100.0	0.6	0.6	0.0	0.9	2.6	1.2	2.0	0.6	0.0	0.9	0.9	4.4	0.6	1.5	1.5	1.5	0.9
C4					100.0	6.3	3.1	3.1	12.5	18.8	0.0	0.0	0.0	0.0	0.0	12.5	3.1	0.0	0.0	0.0	0.0
C5						100.0	0.0	0.0	0.0	13.3	6.7	0.0	0.0	0.0	0.0	6.7	0.0	6.7	0.0	0.0	0.0
C6							100.0	11.1	22.2	0.0	0.0	0.0	11.1	0.0	0.0	22.2	11.1	11.1	0.0	0.0	11.1
C7								100.0	10.5	0.0	0.0	0.0	0.0	5.3	0.0	21.1	26.3	5.3	10.5	5.3	5.3
C8									100.0	4.2	5.3	0.0	2.1	1.1	0.0	3.2	4.2	0.0	1.1	1.1	1.1
C9										100.0	2.2	2.2	0.0	0.0	4.4	8.9	4.4	0.0	2.2	0.0	0.0
C10											100.0	0.0	9.6	11.4	10.5	3.5	1.8	1.8	3.5	0.9	7.9
C11												100.0	0.0	0.0	0.0	66.7	3.0	12.1	3.0	3.0	0.0
C12													100.0	27.3	22.7	4.5	0.0	2.3	4.5	0.0	15.9
C13														100.0	20.0	8.9	0.0	2.2	6.7	0.0	20.0
C14															100.0	2.0	2.0	0.0	10.0	2.0	20.0
C15																100.0	2.4	3.6	3.6	0.6	2.4
C16																	100.0	5.9	5.9	5.9	0.0
C17																		100.0	5.7	0.0	0.0
C18																			100.0	1.1	8.0
C19																				100.0	0.0
C20																					100.0

Fig. (2). Overlap among gene sets. (**a**) Percent overlap among the outlier genes ($|z| > 3.5$) on a compound-by-compound basis in the HUH7 study. The overlap among genes for each compound with itself is always 100%. The value at a row-column intersection reflects the percent of outlier genes that are in common between the two compounds occupying that row and that column. (**b**) Same analysis as (a) but for the THP1 study. Compounds are identified as per Table **1**.

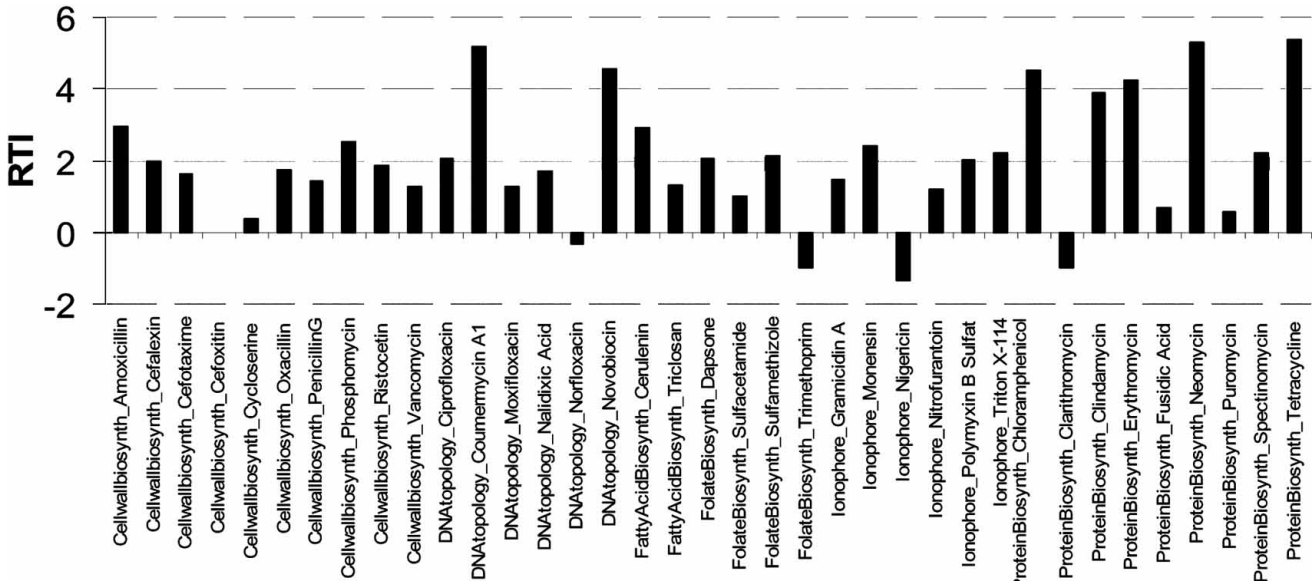

Fig. (3). Relative Transcription Index of Antibacterial Compounds. Bacterial gene expression data from 37 antibacterial compounds belonging to 6 major mechanisms of action classes were compared. The RTI for each compound was calculated relative to cefoxitin (arbitrarily used as reference).

largely non-overlapping sets of genes. These findings point to the fact that different compounds, designed against the same molecular target, can confer unique effects on global gene expression that cannot be known from directed phenotypic assays. In this context RTI provides value as an independent measure of compound function and a metric for compound selection. Additionally, as shown with bacterial gene expression data, RTI may also inform about general properties of mechanism-based compound classes in addition to individual compound rankings.

The scope of the RTI is best appreciated during the early phase of compound screening. In its current format, the RTI is not explicitly designed to be a toxicity screen and compounds with obvious toxic effects do not require the RTI. However, the possibility of adverse effects is expected to be greater for compounds with large RTIs. The RTI is most useful for compounds that appear to be equally efficacious in directed cell-based *in vitro* assays without overt signs of toxicity. It is useful when established methods for compound prioritization cannot provide the information needed for ranking.

In this study, we have confined ourselves to the presentation of the RTI as a practical, quantitative measure of compound behavior. Additional statistical and biological analysis of the sets of genes induced by compounds with large RTI can lead to insights into compound-mediated toxicity or compound-mediated alternative therapeutic benefits. We have also confined ourselves to investigating the effects of compounds on transcription alone. This is a practical consideration reflecting the maturity and adaptability of the gene expression profiling platform technologies today. In the future, the concept of the RTI can be extended to global proteomic and metabonomic profiling of compounds. The concept of RTI is thus generalizable. This will be particularly useful for characterizing compounds that do not have significant effects on transcription.

ACKNOWLEDGEMENTS

We thank Dr. Jack Liu for helpful discussions on the statistical analysis and Patrick McMillen for chip hybridizations and data preparation. The authors report no conflicts of interest.

REFERENCES

[1] Muegge, I. *Med. Res. Rev.*, **2003**, *23*,302.
[2] Rousseeuw, P. *Statistical Methods for Engineers and Scientists*, McGraw-Hill; New York, **1990**.
[3] *Affymetrix Genechip Expression Analysis Technical Manual*, **2003** (http://www.affymetrix.com/support/technical/manual/expression_manual.affx)
[4] Nicholson, J.K.; Wilson, I.D. *Nat. Rev. Drug Discov.*, **2003**, *2*,668.
[5] Hood, L.; Perlmutter, R. *Nat. Biotechnol.*, **2004**, *22*, 1215.
[6] Van der Greef, J.; McBurney, R.N. *Nat. Rev. Drug Discov.*, **2005**, *4*, 961.
[7] Butcher, E.C.; Berg, E.L.; Kunkel, E.J. *Nat. Biotechnol.*, **2004**, *22*, 1253.
[8] Pinhasov, A.; Mei, J.; Amaratunga, D.; Amato, F.A.; Lu, H.; Kauffman, J.; Xin, H.; Brenneman, D.E.; Johnson, D.L.; Andrade-Gordon, P; Ilyin, S.E. *Comb. Chem. High Throughput Screen.*, **2004**, *7*, 133.
[9] Kerns, E.H. *J. Pharm. Sci.*, **2001**, *90*, 1838.
[10] Hummasti, S.; Laffitte, B.A.; Watson, M.A.; Galardi, C.; Chao, L.C.; Ramamurthy, L.; Moore, J.T.; Tontonoz, P. *J. Lipid Res.*, **2004**, *45*, 616.
[11] Walczak, R.; Joseph, S.B.; Laffitte, B.A.; Castrillo, A.; Pei, L.; Tontonoz, P. *J. Biol. Chem.*, **2004**, *279*, 9905.
[12] Schultz, J.R.; Tu, H.; Luk, A.; Repa, J.J.; Medina, J.C.; Li, L.; Schwendner, S; Wang, S.; Thoolen, M.; Mangelsdorf, D.J.; Lustig, K.D.; Shan, B. *Genes Devel.*, **2000**, *14*, 2831.
[13] Joseph, S.B.; Castrillo, A.; Laffitte, B.A.; Mangelsdorf, D.J.; Tontonoz, P. *Nat. Med.*, **2003**, *13*, 213.
[14] Hegele, R.A.; Robinson, J.F. *Curr. Drug Targets Cardiovasc. Haematol. Disord.*, **2005**, *5*, 31.
[15] Williams, S.P.; Bledsoe, R.K.; Collins, J.L.; Boggs, S.; Lambert, M.H.; Miller, A.B.; Moore, J.T.; McKee, D.D.; Moore, L.; Nichols,

J.; Parks, D.; Watson, M.; Wisely, B.; Willson, T.M. *J. Biol. Chem.*, **2003**, *278*, 27138.
[16] Willson, T.M.; Brown, P.J.; Sternbach, D.D.; Henke, B.R. *J. Med. Chem.*, **2000**, *43*, 527.
[17] Willson, T.M.; Moore, J.T. *Mol. Endocrinol.*, **2003**,*16*,1135.
[18] Collins, J.L.; Fivush, A.M.; Watson, M.A.; Galardi, C.M.; Lewis, M.C.; Moore, L.B.; Parks, D.J.; Wilson, J.G.; Tippin, T.K.; Binz, J.G.; Plunket, K.D.; Morgan, D.G.; Beaudet, E.J.; Whitney, K.A.; Kliewer, S.A.; Willson, T.M. *J. Med. Chem.*, **2002**, *45*, 1963.
[19] Chen, Z.; Raymond, K. *Ann. Clin. Microbiol. Antimicrob.*, **2006**, *5*, doi:10.1186/1476-0711-5-3.
[20] Telenti, A. *Thorax*, **1998**, *53*, 793.
[21] Barros, S.A. *Pharmacogenomics*, **2005**, *6*, 547.
[22] Bugelski, P.J. *Curr. Opin. Drug Discov. Devel.*, **2002**, *5*, 79.
[23] Hutter, B.; Schaab, C.; Albrecht, S.; Borgmann, M.; Brunner, N.A.; Freiberg, C.; Ziegelbauer, K.; Rock, C.O.; Ivanov, I.; Loferer, H. *Antimicrob. Agents Chemother.*, **2004**, *48*, 2838.
[24] Scherf, U.; Ross, D.T.; Waltham, M.; Smith, L.H.; Lee, J.K.; Tanabe, L.; Kohn, K.W.; Reinhold, W.C.; Myers, T.G.; Andrews, D.T.; Scudiero, D.A.; Eisen, M.B.; Sausville, E.A.; Pommier, Y.; Botstein, D.; Brown, P.O.; Weinstein, J.N. *Nat. Genet.*, **2000**, *24*, 236.
[25] Hamadeh, H.K.; Bushel, P.R.; Jayadev, S.; DiSorbo, O.; Bennett, L; Li, L.; Tennant, R.; Stoll, R.; Barrett, J.C.; Paules, R.S.; Blanchard, K.; Afshari, C.A. *Toxicol Sci.*, **2002**, *67*, 232.
[26] Way, J.M.; Harrington, W.W.; Brown, K.K.; Gottschalk, W.K.; Sundseth, S.S.; Mansfield, T.A.; Ramachandran, R.K.; Willson, T.M.; Kliewer, S.A. *Endocrinology*, **2001**, *142*, 1269.

Received: February 2, 2007 **Revised: March 7, 2007** **Accepted: March 14, 2007**