Chemical genomics of cancer chemopreventive dithiolethiones

Quynh T.Tran[1,2,3], Lijing Xu[1,3], Vinhthuy Phan[3,4], Shirlean B.Goodwin[2,3], Mostafizur Rahman[2,3], Victor X.Jin[2], Carrie H.Sutter[2,3], Bill D.Roebuck[5], Thomas W.Kensler[6], E.Olusegun George[1,3,4] and Thomas R.Sutter[2,3,*]

[1]Department of Mathematical Sciences and [2]Department of Biology, University of Memphis, Memphis, TN 38152, USA, [3]W. Harry Feinstone Center for Genomic Research and [4]Department of Computer Science, University of Memphis, Memphis, TN 38152, USA, [5]Department of Pharmacology and Toxicology, Dartmouth Medical School, Hanover, NH 03755, USA and [6]Department of Environmental Health Sciences, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD 21205, USA

[*]To whom correspondence should be addressed. Tel: +1 901 678 8391;
Fax: +1 901 678 2458;
Email: tsutter@memphis.edu

3H-1,2-dithiole-3-thione (D3T) and its analogues 4-methyl-5-pyrazinyl-3*H*-1,2-dithiole-3-thione (OLT) and 5-tert-butyl-3*H*-1,2-dithiole-3-thione (TBD) are chemopreventive agents that block or diminish early stages of carcinogenesis by inducing activities of detoxication enzymes. While OLT has been used in clinical trials, TBD has been shown to be more efficacious and possibly less toxic than OLT in animals. Here, we utilize a robust and high-resolution chemical genomics procedure to examine the pharmacological structure–activity relationships of these compounds in livers of male rats by microarray analyses. We identified 226 differentially expressed genes that were common to all treatments. Functional analysis identified the relation of these genes to glutathione metabolism and the nuclear factor, erythroid derived 2-related factor 2 pathway (Nrf2) that is known to regulate many of the protective actions of dithiolethiones. OLT and TBD were shown to have similar efficacies and both were weaker than D3T. In addition, we identified 40 genes whose responses were common to OLT and TBD, yet distinct from D3T. As inhibition of cytochrome P450 (CYP) has been associated with the effects of OLT on CYP expression, we determined the half maximal inhibitory concentration (IC_{50}) values for inhibition of CYP1A2. The rank order of inhibitor potency was OLT $\gg$ TBD $\gg$ D3T, with IC_{50} values estimated as 0.2, 12.8 and >100 μM, respectively. Functional analysis revealed that OLT and TBD, in addition to their effects on CYP, modulate liver lipid metabolism, especially fatty acids. Together, these findings provide new insight into the actions of clinically relevant and lead dithiolethione analogues.

Introduction

The chemicals 3*H*-1,2-dithiole-3-thione (D3T), 4-methyl-5-pyrazinyl-3*H*-1,2-dithiole-3-thione (OLT) and 5-tert-butyl-3*H*-1,2-dithiole-3-thione (TBD) are dithiolethione class chemopreventive agents that

Abbreviations: AFB_1, aflatoxin B_1; ANOVA, analysis of variance; C/EBP, CCAAT/enhancer-binding protein; CORE_TF, conserved and overrepresented transcription factor; CYP, cytochrome P450; D3T, 3*H*-1,2-dithiole-3-thione; FDR, false discovery rate; GO, gene ontology; GST, glutathione *S*-transferase; KW, Kruskal–Wallis; Nrf2, nuclear factor, erythroid derived 2-related factor 2; OLT, 4-methyl-5-pyrazinyl-3*H*-1,2-dithiole-3-thione; PCA, principal component analysis; PPAR, peroxisome proliferator-activated receptor; RT–PCR, reverse transcriptase–polymerase chain reaction; TBD, 5-tert-butyl-3H-1,2-dithiole-thione; TFBS, transcription factor-binding site.

have been shown to inhibit chemical toxicity and carcinogenicity in many target organs (1). They continue to be under preclinical and clinical investigations for use as cancer chemopreventive agents in humans (2). The three compounds are known to activate the Kelch-like erythroid cell-derived protein with Cap 'n' collar homology (ECH)-associated protein 1–nuclear factor, erythroid derived 2-related factor 2 (Nrf2)-signaling pathway (3,4). Kelch-like ECH-associated protein 1 sequesters Nrf2 through binding to the N-terminal regulatory domain of Nrf2 in the cytoplasm (5). Administration of sulfhydryl reactive compounds such as diethylmaleate, sulforaphane or D3T abolishes Kelch-like ECH-associated protein 1 repression of Nrf2 activity, facilitating the translocation and accumulation of Nrf2 in the nucleus (4,6,7). As a transcription factor, Nrf2 binds to the antioxidant response element, forms heterodimers with small Maf (v-maf, avian musculoaponeurotic fibrosarcoma oncogene homolog) proteins and regulates the expression of antioxidant response element-containing genes (6,7). The induction of these cytoprotective proteins enhances resistance against electrophiles, oxidative stress and chemical toxicants (8).

Several studies of OLT have evaluated the pharmacodynamics of this compound in cancer chemoprevention (9). Additional clinical trials were conducted in Qidong, Jiangsu Province, People's Republic of China (10–12). These trials were randomized, placebo-controlled and double-masked studies. In a Phase IIa chemoprevention trial, high dose OLT (500 mg/week) inhibited bioactivation of aflatoxin B_1 (AFB_1), whereas sustained low dose OLT (125 mg/day) increased conjugation of AFB_1 epoxide, leading to increased rates of excretion of aflatoxin–mercapturic acid in urine (12).

TBD is a recent lead dithiolethione compound for cancer chemoprevention. In 2003, Roebuck *et al.* (2) compared the cancer chemopreventive potency of several dithiolethione analogues in male F344 rats. In this study, D3T, OLT and TBD were shown to attenuate the hepatic burden of AFB_1-induced preneoplastic lesions. Relative potency estimates were obtained by comparing the doses of dithiolethiones required to reduce the focal volume percent to 0.1% of the liver volume. D3T proved to be the most potent agent, whereas TBD was equipotent with OLT. Furthermore, it was suggested that TBD may be a more useful chemopreventive agent than OLT because of features that may make it less toxic, specifically the difference in its wavelength absorption and the lack of a pyrazine ring, as well as relative ease and yield of synthesis (2) (see Figure 1A for chemical structures).

Here, we apply a robust chemical genomics approach to compare and contrast the gene expression activities of D3T, OLT and TBD. We show that this procedure is simultaneously able to identify treatment-responsive genes and the known pharmacological relationships that are common among, and distinct between, the treatment groups. Furthermore, we show that the distinct gene clusters, indicative of unique structure–activity relationships, capture the activities of additional pharmacophores present in the chemical analogues in this chemical series.

Materials and methods

Animal treatments

Male Fischer F344 rats (80–100 g) were obtained from Harlan Laboratories (Indianapolis, IN) at 6 weeks of age. The chemicals and procedures for animal care were described previously in detail (13). Animals were randomly assigned into four groups of four animals for each treatment by gavage with 100 μl of vehicle (saturated sucrose), D3T (0.3 mmol/kg body wt), OLT (0.3 mmol/kg body wt) or TBD (0.3 mmol/kg body wt), once every other day over 5 days. Animals were killed 24 h after the third dose and the liver tissues were snap
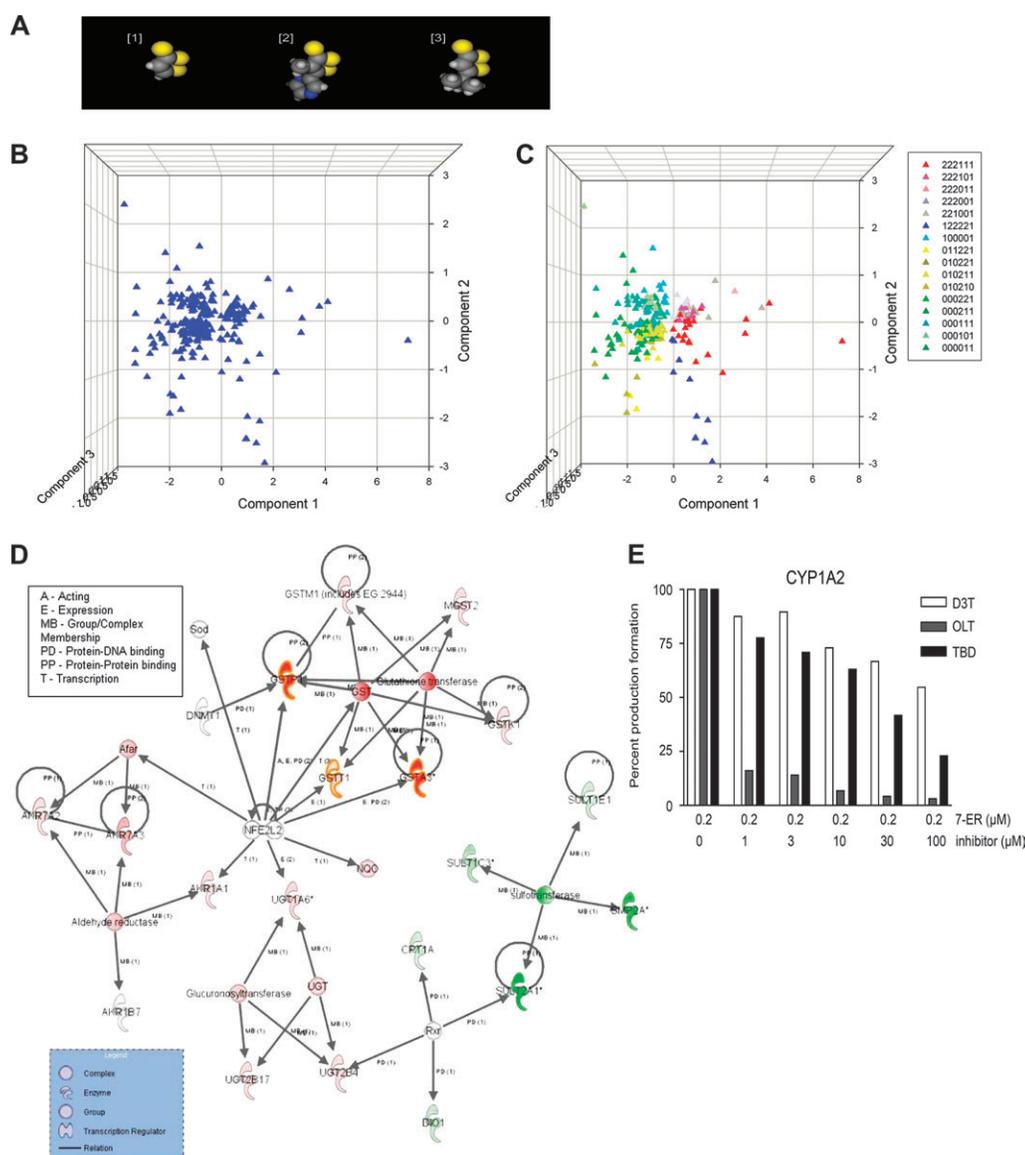
Fig. 1. (**A**) Space filling models of the three dithiolethiones: [1] D3T, [2] OLT and [3] TBD. The atoms are identified by colors: sulfur, yellow; carbon, dark gray; hydrogen, light gray; nitrogen, blue. (**B**) PCA of gene expression using the treatments as variables. The Eigen values of the principal component 1, 2, 3 and 4 are 17.513, 9.863, 5.196 and 0, respectively. The first three principal components contain over 98% of the total variance in the data. (**C**) PCA of gene expression color coded for the patterns identified in Table I, containing greater than six genes (right box). (**D**) Biological network common to D3T, OLT and TBD treatments. The intensity of the node color indicates the degree of upregulation (red) or downregulation (green). The lower legend box (blue) indicates the node shapes in order of complex, enzyme, group, transcription regulator and relation. The upper legend box indicates the relationships between two nodes. (**E**) Inhibition of *CYP1A2* catalyzed 7-ethoxyresorufin-*O*-deethylase by D3T, OLT and TBD. The concentrations of the substrate (7-ER, 7-ethoxyresorufin) and the inhibitors are shown at the bottom.

frozen. The protocol for this study was approved by the Animal Care and Use Committee of the Johns Hopkins University, where the animals were treated.

RNA procedures

Total RNA was isolated from the frozen liver tissue using STAT-60 (Tel-Test, Friendswood, TX). Affymetrix RG-U34A chips (containing 8799 probe sets) were used to measure the messenger RNA levels as described previously (13). The Affymetrix dataset was submitted to the Gene Expression Omnibus database, GSE8882. Real-time reverse transcriptase–polymerase chain reaction (RT–PCR) was performed as described previously (13). The Ct values for all genes were normalized to that of β-actin, and the relative value for the control samples was set as one arbitrary unit.

Cytochrome P450 1A2 inhibition assays

Cytochrome P450 (CYP) 1A2 inhibition assay for 7-ethoxyresorufin *O*-deethylase activity, 1 ml each, contained 5 pmol of human microsomal *CYP1A2* (BD Biosciences, San Jose, CA), 1.4 mM reduced nicotinamide

adenine dinucleotide phosphate and 0.2 μM of the substrate, 7-ethoxyresorufin in 0.1 M KPO_4 buffer, pH 7.4. D3T, OLT and TBD were dissolved in dimethyl sulfoxide and added at concentrations of 0, 1, 3, 10, 30 and 100 μM. The reactions were incubated for 15 min at 37°C, stopped by the addition of 2 ml of ice-cold methanol, vortexed and then centrifuged for 5 min at 4302*g*. The formation of product was determined fluorometrically against a standard of resorufin, using a Varian Cary Eclipse fluorescence spectrophotometer (EL 0206-5717) with an excitation and emission wavelength of 530 and 585 nm, respectively.

Data analysis

Microarray data were preprocessed by GC-Robust Multi-array Average (14). Affymetrix Microarray Suite 5.0 was used to detect the presence and absence of gene expression (15). Genes that were present in 75% of any treatment group were selected using a Perl program. The Kruskal–Wallis (KW) procedure, permutation tests and the recursive procedure were implemented using JAVA. In the

current Affymetrix technology, one gene may be represented by more than one probe set. For these redundant probe sets, only the probe sets with the most numbers of smallest pairwise comparison *P*-values were kept for further analysis. Principal component analysis (PCA) was performed according to Raychaudhuri *et al.* (16). Eigen values for PCA were calculated using Cluster 3.0 (17). SigmaPlot 10.0 (Systat Software, Inc., Chicago, IL, 2002) was used to produce three-dimensional PCA plots with color-coded clusters. Analysis of variance (ANOVA) was performed using a National Institute on Aging Array Analysis Tool (18). Gene ontology (GO) enrichment analysis was performed using Gene Ontology Enrichment Analysis Software Toolkit (19). The significant of each GO class was determined by its *P*-value using default criteria. Biological network and pathway analysis were performed using ingenuity pathway analysis software (Ingenuity System, Redwood City, CA). Common transcription factor-binding sites (TFBS) in promoters of coregulated genes were identified using conserved and overrepresented transcription factor binding sites (CORE_TF) (20). Ensembl gene IDs were used as input into CORE_TF. A 5′ flanking region of 2 kb upstream to 1 kb downstream of the transcription start site was designated as an experimental promoter set since most of TFBSs are located within these regions. A random set of 3000 promoters from rat species was selected to be compared with the experimental set in searching for overrepresented TFBSs (20).

Statistical methods

To ensure robustness, we used the KW rank procedure, which does not require assumptions of data distribution. To account for the small number of replications, exact *P*-values of the KW tests were computed using a permutation resampling method (21) to perform 1.0E6 permutations. For each gene with a significant KW test, post-hoc analyses were performed to identify the pairs of treatments with significantly different expression values using the Wilcoxon rank-sum tests. Since the number of replication per treatment group is small, exact *P*-values of the Wilcoxon tests were computed by a recursive procedure (22). False discovery in the multiple testing was controlled with the false discovery rate (FDR) procedure (23). Based on the results of the post-hoc pairwise comparisons, we clustered the gene responses in relation to treatment effects by applying a hierarchical procedure developed by Sutter *et al.* (24). Specifically, for testing H_0: $\mu_{ag} = \mu_{bg}$ versus H_1: $\mu_{ag} < \mu_{bg}$, where μ_{tg} denotes the expected value of the expression of gene g under treatment t, we use the recurrence relation (22) to compute the exact *P*-value $P_{n,m}(r_g)$, where r_g is the rank sum of the n expression values of gene g under treatment a and m is the number of replications under treatment b. If $P_{n,m}(r) \leq \alpha$, then $\mu_a < \mu_b$ and a 2 was assigned to the result of the comparison (upregulated genes). Another possibility was H_1: $\mu_a > \mu_b$. The *P*-value for this test was computed by $1 - P_{n,m}(r - 1)$. If this *P*-value was not $>\alpha$, then $\mu_a > \mu_b$ and a 0 was assigned (downregulated genes). When both $P_{n,m}(r)$ and $1 - P_{n,m}(r - 1)$ were $>\alpha$, the null hypothesis was accepted and a 1 was assigned to the result (no change).

Results and discussion

Clustering of microarray data

Normal probability plots and Kolmogoroff–Smirnov tests (25,26) demonstrated that the raw gene expression values, as well as the log-transformed expression values, were not normally distributed (data not shown). Therefore, a non-parametric procedure was used to test the differences among the expression levels of each gene under four treatment conditions. After controlling for FDR, 431 significant probe sets were identified. A recursive procedure was then applied to detect differences occurring between any two treatments. An illustration of this analysis is shown for one gene (supplementary Table I is available at *Carcinogenesis* Online), *Gstp1*. After resolving the redundant probe sets, the number of significant genes was reduced from 431 to 367 genes. The genes were then clustered according to their responses to the treatments. For four treatments, there were six pairwise comparisons and 219 possible clusters (24). A 1 was assigned to all insignificant comparisons and 2563 probe sets had the 111111 patterns of pairwise comparisons. The remaining 367 genes were binned into 64 additional clusters (Table I).

Comparison with ANOVA

By comparison, the non-parametric and ANOVA methods appeared to detect similar differences in gene expression (supplementary Table II

is available at *Carcinogenesis* Online), with 93% of the probe sets identified by both methods. However, the ANOVA method identified many more significant probe sets (761 versus 431). As the assumption of normality was not supported, we explored this issue in more detail by RT–PCR. From the results of each procedure, we selected the five genes whose FDR *P*-values were closest to the cutoff threshold (supplementary Table III is available at *Carcinogenesis* Online). RT–PCR was performed on each sample, and Pearson correlation analysis was performed between the microarray and RT–PCR results. We observed a strong correlation between the expression values for the genes identified by the non-parametric method ($R = 0.8576$, *P*-value = 0.0001) (supplementary Figure 1A is available at *Carcinogenesis* Online). However, for the genes selected from ANOVA, the correlation of the expression values between microarray and RT–PCR was low ($R = 0.0879$, *P*-value = 0.446) (supplementary Figure 1B is available at *Carcinogenesis* Online). These results support the robustness of the non-parametric method and provide confidence in the identification and clustering of genes by this procedure, thereby reducing biological noise in the subsequent functional analysis of this data.

PCA

The resolution of our chemical genomics approach was revealed by comparison with PCA, a statistical technique for dimension reduction in multivariate data sets, often used to cluster genes with similar expression profiles (18,27). PCA allows one to observe graphically the key variables of the differences in the gene observations. In this study, experimental treatments were the variables. The percentage variance from PCA showed that the first three components captured most of the information about the observed variability in this experiment. Shown in Figure 1B is the plot of the first three principal components for genes in clusters containing more than six genes, in order to show a clearer view of the data. In Figure 1B, PCA identified only two large clusters that appeared close to one another. However, if we color code these genes (Figure 1C) using the clusters identified by our method (Table I), we observe that the two major groups of genes identified by PCA are identified by our method as upregulated (red and gray tones) and downregulated (green and yellow tones) genes, relative to control. For example, clusters 211001 (dark red) and 011221 (yellow) represent genes showing increased or decreased expression in response to only treatment with D3T. Moreover, we observe that the distinct clusters identified by our procedure tend to group in distinct regions of the PCA plot. This observation indicates a general concordance of these two methods, with our method producing a much finer cluster, expressing an explicit pharmacological relationship between responses and treatments (Figure 1C). For example, clusters 122221 (blue) and 100001 (cyan) represent genes that, although D3T is the most efficacious chemical among the three (2), respond only to OLT and TBD. This unique ability to identify distinct clusters in a series of closely related chemical structures clearly demonstrates the strength of our procedure for chemical genomic applications.

Common pharmacodynamic action of D3T, OLT and TBD

The shared structure–activity relationships of the three dithiolethiones (Figure 1A) was observed in the clusters containing the most responsive genes, clusters 1–8 and 54–65 (patterns 222xxx and 000xxx) consisting of 68 upregulated genes and 158 downregulated genes, respectively (Table I). As D3T, the parent compound, is known to activate the Nrf2 pathway, this observation was expected. Included in the upregulated genes were those known to be mechanistically important to the prevention AFB_1 carcinogenesis, including glutathione *S*-transferase (GST) Yc2 subunit and AFB_1 aldehyde reductase, as well as other well-known Nrf2-regulated genes including NAD(P)H quinone oxidoreductase 1 (*Nqo1*), *Gstp1* and uridine diphosphate glucuronosyltransferase 1 family, polypeptide A6 (*UGT1A6*) (2,6,28).

Functional analysis of the 226 genes in the 222xxx and 000xxx clusters by GO enrichment (19) identified metabolic process and

Table I. The clusters of genes whose patterns identify responses to treatments

Cluster	Pattern[a]						Number of genes
	CD	CO	CT	DO	DT	OT	
1	2	2	2	2	2	1	3
2	2	2	2	1	2	1	5
3	2	2	2	1	1	1	27
4	2	2	2	1	0	1	8
5	2	2	2	1	0	0	1
6	2	2	2	0	1	2	1
7	2	2	2	0	1	1	9
8	2	2	2	0	0	1	14
9	2	2	1	2	1	0	1
10	2	2	1	1	0	1	2
11	2	2	1	0	1	1	2
12	2	2	1	0	0	1	8
13	2	1	2	1	1	2	1
14	2	1	2	1	0	1	2
15	2	1	2	0	1	2	2
16	2	1	2	0	1	1	3
17	2	1	2	0	0	1	3
18	2	1	1	0	1	1	1
19	2	1	1	0	0	1	5
20	1	2	2	2	2	1	9
21	1	2	2	2	1	1	3
22	1	2	2	1	2	1	4
23	1	2	2	1	1	1	1
24	1	2	1	2	1	1	1
25	1	2	1	2	1	0	1
26	1	1	2	2	2	2	1
27	1	1	1	1	1	1	2563
28	1	1	0	1	0	1	1
29	1	1	0	1	0	0	2
30	1	0	1	0	1	1	1
31	1	0	0	1	0	1	4
32	1	0	0	1	0	0	2
33	1	0	0	0	1	1	4
34	1	0	0	0	0	1	12
35	1	0	0	0	0	0	1
36	0	2	2	2	2	1	3
37	0	2	1	2	2	1	1
38	0	2	0	2	1	0	1
39	0	1	2	2	2	1	1
40	0	1	1	2	2	1	9
41	0	1	1	2	2	0	1
42	0	1	0	2	2	1	11
43	0	1	0	2	1	1	9
44	0	1	0	2	1	0	11
45	0	1	0	1	2	1	1
46	0	1	0	1	1	1	4
47	0	1	0	1	1	0	1
48	0	0	2	1	2	2	1
49	0	0	1	2	2	1	3
50	0	0	1	1	2	2	2
51	0	0	1	1	2	1	1
52	0	0	1	1	1	1	2
53	0	0	1	0	1	1	2
54	0	0	0	2	2	1	9
55	0	0	0	2	1	1	29
56	0	0	0	2	1	0	1
57	0	0	0	1	2	2	1
58	0	0	0	1	2	1	6
59	0	0	0	1	1	2	1
60	0	0	0	1	1	1	74
61	0	0	0	1	1	0	6
62	0	0	0	1	0	1	12
63	0	0	0	1	0	0	4
64	0	0	0	0	1	1	13
65	0	0	0	0	0	1	2

[a]CD, control versus D3T; CO, control versus OLT; CT, control versus TBD; DO, D3T versus OLT; DT, D3T versus TBD; OT, OLT versus TBD.

glutathione transferase activity in the top three significant GO classes (Table IIA). Similarly, ingenuity pathway analysis identified glutathione metabolism and metabolism of xenobiotics among the top significant canonical pathways (Table IIB). Also among these canonical pathways was Nrf2-mediated oxidative stress response, a well-known regulator of D3T activity (3,4,6,7,29). The biological network corresponding to Nrf2 was also enriched and is shown in Figure 1C as the major network enriched by D3T, OLT and TBD. In this network, genes involved in glutathione metabolism such as *GST*, *Gstp1*, *GSTt1*, *Gsta3*, *GSTm1*, microsomal *GST2* (*MGST2*) and *GSTk1* were increased. Thus, there was good agreement between the GO and ingenuity pathway analysis for the genes common to these three compounds.

Promoter analysis of the 226 common genes also indicated that these genes were regulated by Nrf2. Among these 226 genes, 191 (84.5%) are well annotated and CORE_TF was able to retrieve their promoter sequences for the analysis. There was a significant overrepresentation (P-value < 0.05, after correction for FDR) for the Nrf2 matrix NRF2_Q4 in 189 of 191 (98.9%) experimental promoters (supplementary Table IV is available at *Carcinogenesis* Online).

The efficacies of OLT and TBD are similar and both weaker than D3T

As the test compounds were administered in equal molar doses, a lower efficacy of OLT, relative to D3T and TBD, was observed in clusters 13–17 and 42–47 (patterns 212xxx and 010xxx) containing 48 genes (Table I). These genes did not change in response to OLT, indicating that OLT was less efficacious in the response of these genes, a finding consistent with the study of Roebuck *et al.* (2) for other end points. Of the observed clusters, a large portion (three of five in clusters 13–17 and four of six in clusters 42–47) showed no response to either OLT or TBD. This observation is also consistent with the reported efficacies of these compounds, which indicates that OLT and TBD are weaker than the parent compound, D3T (2). Similarly, clusters showing the greater efficacy of D3T were also observed, i.e. the 16 genes that responded only to D3T (patterns 211xxx and 011xxx). Finally, there were 23 genes that responded to OLT and D3T, but not to TBD (patterns 221xxx and 001xxx). Collectively, these data support the current understanding that D3T is the most efficacious of the three compounds, with the observed efficacies of D3T $\ggg$ TBD $\geq$ OLT.

Unique OLT and TBD gene clusters indicate novel pharmacodynamic action

While the common clusters discussed above contained 89% of the statistically significant genes, the remaining 11% of the responding genes identified two unique sets of clusters, 20–23 and 31–35 (patterns 122xxx and 100xxx), containing 40 genes responding to OLT and TBD, but not to D3T (Table I).

Two of the most highly induced genes in this set were *Cyp2b6* and *Cyp2b19*, elevated 11- and 7-fold, respectively. Two previous studies have shown that OLT induces members of multiple families of CYP, especially the CYP1A and *Cyp2b6* families in livers of rat (24,30). Furthermore, it has been proposed that the induction of these genes may be related to the inhibitory properties of OLT toward hepatic CYP (31). One explanation for this observed activity of OLT for CYP enzyme induction is its unique structure, compared with the unsubstituted parent dithiolethione, D3T (24). As shown in Figure 1A, OLT contains a 5-pyrazinyl substituent on the dithiolethione ring. This nitrogen-containing, electron-withdrawing ring is believed to make OLT an inhibitor of hepatic CYP (24). To explore this property in light of the previously identified OLT cluster associated with inhibition of CYP (24), we determined the relative potencies of D3T, OLT and TBD for inhibition of CYP1A2-catalyzed 7-ethoxyresorufin-*O*-deethylase (Figure 1E). The rank order of inhibitor potency was OLT $\gg$ TBD $\gg$ D3T, with half maximal inhibitory concentration values estimated as 0.2, 12.8 and >100 μM, respectively. Although the relatively high and multiple doses of 0.3 mmol/kg body wt did not

Table II. GO and functional analyses of the 226 genes regulated by D3T, OLT and TBD and the 40 genes regulated by OLT and TBD only

A

GO ID	GO term	q	m	t	K	P-value
GO:0008152	Metabolic process	62	3872	8799	75	1.7×10^{-10}
GO:0006412	Translation	16	216	8799	75	3.8×10^{-10}
GO:0004364	Glutathione transferase activity	8	31	8799	75	4.8×10^{-10}

B

Canonical pathway	P-value	D3T-, OLT- and TBD-regulated genes
Metabolism of xenobiotics by CYP	5.0×10^{-11}	GSTa3, GSTm1, GSTt1, GSTp1, UGT2B4, UGT1A6, CYP2C18, MGST2, CYP2A12, UGT2B17, GSTk1
NRF2-mediated oxidative stress response	1.9×10^{-09}	AKR7A2, GSTa3, AKR7A3, GSTm1, GSTt1, GSTp1, AKR1A1, FTL, MGST2, Nqo1, GSTk1
Xenobiotic metabolism signaling	8.5×10^{-09}	GSTa3, GSTm1, GSTp1, FTL, UGT1A6, Nqo1, GSTt1, MGST2, UGT2B4, UGT2B17, GSTk1, MAOA
Glutathione metabolism	1.7×10^{-08}	GSTa3, GSTm1, GSTt1, GSTp1, MGST2, IDH2, GSTk1

C

Biological functions	P-value	OLT- and TBD-regulated genes
Molecular transport	3.6×10^{-4}	Ifrd1, SLC34A1, Nr0b2, COMT, Slc37a4, Adh1, Acat1, HSP90AA1, Slc25a10, Cyp2b6, Pck1
Carbohydrate metabolism	5.4×10^{-4}	Slc37a4, Slc25a10, Pck1, SDS
Drug metabolism	8.1×10^{-4}	SLC34A1, CYP2C18, COMT, Adh1, Cyp2b6, HSD17B2
Small molecule biochemistry	8.1×10^{-4}	SLC34A1, Adh1, Pck1, Hacl1, Hmgcs2, FAH, Ifrd1, ATPIF1, Nr0b2, CYP2C18, BCKDHA, COMT, Slc37a4, Acat1, Slc25a10, Cyp2b6, SDS, HSD17B2, LTB4DH, Cyp2b19

q, the number of probe sets in the list responding to D3T, OLT and TBD and associated with the GO term; m, number of probe sets on the chip associated with the GO term; t, total number of probe sets on the chip; K, number of probe sets in the input list for GO enrichment analysis.

permit discrimination of the potential differences between OLT and TBD in this regard, our results do show that the dose–response relationships for enzyme inhibition and Nrf2 activation by these chemicals are quite different. If this unique action of OLT and TBD proves to be an important consideration in their use as prophylactic agents, then further studies to explore the dose–response relationships and mechanism of this response could be performed.

In order to further explore this unique activity, functional analysis of these 40 genes was performed (Table IIC). The top four significant biological functions were related to molecular transport, including molecular transport ($P = 3.2 \times 10^{-4}$), carbohydrate metabolism ($P = 5.4 \times 10^{-4}$), drug metabolism ($P = 8.1 \times 10^{-4}$) and small-molecule biochemistry ($P = 8.1 \times 10^{-4}$). Of interest, most of the downregulated genes in these categories were related to lipid metabolism, especially fatty acids. Of these genes, two are known regulators of the accumulation of triglycerides: the transcription factor nuclear receptor subfamily 0, group B, member 2 (*Nr0b2*, decreased 3-fold) and the regulator interferon-related developmental regulator 1 (*Ifrd1*, decreased 3-fold) (32–34). In addition, several genes involved in fatty acid biosynthesis were also decreased including 2-hydroxyacyl-CoA lyase 1 (*Hacl1*, decreased 7.5-fold), acetyl-CoA C-acetyltransferase (*Acat1*, decreased 1.6-fold), alcohol dehydrogenase 1 (*Adh1*, decreased 2.6-fold) (35) and solute carrier family 25 member 10 (*Slc25a10*, decreased 1.3-fold) (36). Two additional genes important in the control of gluconeogenesis and fat deposition were also downregulated, phosphoenolpyruvate carboxykinase 1 (*Pck1*, decreased 2-fold) and glucose-6-phosphate transporter member 4 (*Slc37a4*, decreased 1.4-fold) (37). In addition, adiponectin receptor 2 (*Adipor2*, increased 6-fold) and 3-hydroxy-3-methylglutaryl-coenzyme A synthase 2 (*Hmgcs2*, decreased 1.4-fold) play central roles in the regulation of glucose and lipid metabolism (38,39).

Promoter analysis was performed on the 40 coexpressed genes to look for common TFBSs. Of the 40 genes, 32 (80%) had well-annotated transcription start sites and could be used for further analysis using CORE_TF. The top two overrepresented TFBSs were the peroxisome proliferator-activated receptors (PPARs), PPARγ and PPARα (Table III). There were 30 genes (93.8%) and 26 genes (81.3%) with

TFBSs for PPARγ and PPARα, respectively (supplementary Table V is available at *Carcinogenesis* Online). These genes include *Nr0b2*, *Pck1*, *Slc37a4*, *Acat1*, *Slc25a10*, *Adipor2*, *Adh1*, *Hmgcs2* and *Hacl1* (also known as phytanoyl-CoA 2-hydroxylase 2, Phyh2), which were associated with lipid metabolism. Studies have shown that *Acat1*, *Slc25a10*, *Adipor2* and *Hmgcs2* are PPARα-regulated genes (38,40), whereas *Nr0b2* and *Pck1* are mediated by PPARγ (41,42). The PPAR nuclear receptors play important roles in the regulation of carbohydrate, lipid and protein metabolism, inflammation and adipocyte differentiation. PPARα is highly expressed in liver where it stimulates fatty acid oxidation and ketogenesis (43). It also regulates the peroxisome and mitochondrial fatty acid oxidation pathway. PPARγ is expressed at low levels in the liver, but can be upregulated by PPARγ activators. These same activators can suppress carcinogenesis in animal models (43,44).

In mice, OLT has been shown to increase the expression of *Cyp2b10* by a constitutive androstane receptor-mediated mechanism (45), which could account for the observed induction of *Cyp2b6* messenger RNAs here. However, the constitutive androstane receptor TFBS was not significantly represented in the 40 gene set. In rat hepatocyte H4IIE cells, OLT has been shown to increase the expression of the GSTA2 via enhanced binding of the CCAAT/enhancer-binding protein (C/EBP) β (46). Although the C/EBPβ TFBS was not significantly represented in the 40 gene set, a recent study shows that >1000 genes are coregulated by PPARγ and C/EBPα or C/EBPβ during adipocyte differentiation, indicating a strong co-operation of these factors (47). Another study showed that the PPARγ activator, 15-deoxy-$\Delta^{12,14}$-prostaglandin J$_2$ (15d-PGJ$_2$), increased the expression of GSTA2, which was lost by deletion of either the C/EBPβ- or PPAR-responsive enhancers, indicating that these factors interact in the regulation of this gene (48).

In another study, 15d-PGJ$_2$ was shown to activate Nrf2, as well as PPARγ (49), demonstrating that such chemicals may activate multiple nuclear receptors in the liver, adding to the complexity of their effects. Combined genetic and chemical genomic studies will be necessary to understand the mechanisms of action of such chemicals. The information and procedures described here should aid in the design and analysis of such studies.

Table III. The top three significant overrepresented matrices found in the promoters of the 40 genes regulated by OLT and TBD using CORE_TF

Name of matrix	P-value	Number of experimental promoters hit	Number of random promoters hit	Number of random promoters	Frequency random[a]
PPARG_03	2.1×10^{-08}	30	1521	2944	0.52
PPAR_DR1_Q2	1.3×10^{-07}	26	1135	2944	0.39

[a]Frequency of hits in the random data.

Supplementary materials

Supplementary Tables I–V and Figure 1 can be found at http://carcin.oxfordjournals.org/

References

1. Wattenberg,L.W. *et al*. (1986) Inhibitory effects of 5-(2-pyrazinyl)-4-methyl-1,2-dithiol-3-thione (Oltipraz) on carcinogenesis induced by benzo[a]pyrene, diethylnitrosamine and uracil mustard. *Carcinogenesis*, **7**, 1379–1381.
2. Roebuck,B.D. *et al*. (2003) Evaluation of the cancer chemopreventive potency of dithiolethione analogs of oltipraz. *Carcinogenesis*, **24**, 1919–1928.
3. Itoh,K. *et al*. (1999) Regulatory mechanisms of cellular response to oxidative stress. *Free Radic. Res.*, **31**, 319–324.
4. Kwak,M.K. *et al*. (2001) Role of transcription factor Nrf2 in the induction of hepatic phase 2 and antioxidative enzymes *in vivo* by the cancer chemoprotective agent, 3H-1, 2-dimethiole-3-thione. *Mol. Med.*, **7**, 135–145.
5. Itoh,K. *et al*. (1999) Keap1 represses nuclear activation of antioxidant responsive elements by Nrf2 through binding to the amino-terminal Neh2 domain. *Genes Dev.*, **13**, 76–86.
6. Itoh,K. *et al*. (1997) An Nrf2/small Maf heterodimer mediates the induction of phase II detoxifying enzyme genes through antioxidant response elements. *Biochem. Biophys. Res. Commun.*, **236**, 313–322.
7. Venugopal,R. *et al*. (1996) Nrf1 and Nrf2 positively and c-Fos and Fra1 negatively regulate the human antioxidant response element-mediated expression of NAD(P)H:quinone oxidoreductase1 gene. *Proc. Natl Acad. Sci. USA*, **93**, 14960–14965.
8. Yu,X. *et al*. (2005) Nrf2 as a target for cancer chemoprevention. *Mutat. Res.*, **591**, 93–102.
9. Gupta,E. *et al*. (1995) Pharmacokinetics and pharmacodynamics of oltipraz as a chemopreventive agent. *Clin. Cancer Res.*, **1**, 1133–1138.
10. Kensler,T.W. *et al*. (1998) Oltipraz chemoprevention trial in Qidong, People's Republic of China: modulation of serum aflatoxin albumin adduct biomarkers. *Cancer Epidemiol. Biomarkers Prev.*, **7**, 127–134.
11. Jacobson,L.P. *et al*. (1997) Oltipraz chemoprevention trial in Qidong, People's Republic of China: study design and clinical outcomes. *Cancer Epidemiol. Biomarkers Prev.*, **6**, 257–265.
12. Wang,J.S. *et al*. (1999) Protective alterations in phase 1 and 2 metabolism of aflatoxin B1 by oltipraz in residents of Qidong, People's Republic of China. *J. Natl Cancer Inst.*, **91**, 347–354.
13. Huang,Y. *et al*. (2006) Identification of novel transcriptional networks in response to treatment with the anticarcinogen 3H-1,2-dithiole-3-thione. *Physiol. Genomics*, **24**, 144–153.
14. Wu,Z. *et al*. (2004) A model-based background adjustment for oliogonucleotide expression arrays. *J. Am. Stat. Assoc.*, **99**, 909–917.
15. Liu,W.M. *et al*. (2002) Analysis of high density expression microarrays with signed-rank call algorithms. *Bioinformatics*, **18**, 1593–1599.
16. Raychaudhuri,S. *et al*. (2000) Principal components analysis to summarize microarray experiments: application to sporulation time series. *Pac. Symp. Biocomput.*, **5**, 452–463.
17. Eisen,M.B. *et al*. (1998) Cluster analysis and display of genome-wide expression patterns. *Proc. Natl Acad. Sci. USA*, **95**, 14863–14868.
18. Sharov,A.A. *et al*. (2005) A web-based tool for principal component and significance analysis of microarray data. *Bioinformatics*, **21**, 2548–2549.
19. Zheng,Q. *et al*. (2008) GOEAST: a web-based software toolkit for Gene Ontology enrichment analysis. *Nucleic Acids Res.*, **36**, W358–W363.
20. Hestand,M.S. *et al*. (2008) CORE_TF: a user-friendly interface to identify evolutionary conserved transcription factor binding sites in sets of co-regulated genes. *BMC Bioinformatics*, **9**, 495.
21. Davidson,A.C. *et al*. (1997) *Bootstrap Methods and Their Application*. Cambridge University Press, Cambridge, UK.
22. Ross,M.S. (2002) *Simulation*. Academic Press, San Diego, CA.
23. Benjamini,Y. *et al*. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B*, **57**, 289–300.
24. Sutter,T.R. *et al*. (2002) Multiple comparisons model-based clustering and ternary pattern tree numerical display of gene response to treatment: procedure and application to the preclinical evaluation of chemopreventive agents. *Mol. Cancer Ther.*, **1**, 1283–1292.
25. Smirnov,N. (1939) Sur les ecarts de la courbe de distribution and epirique (Russian, French summary). *Mathematicheskii Sbornik*, **48**, 3–26.
26. Kolmogoroff,A. (1933) Sulla Determinazione empirica di una legge di distribuzione. *Giornale dell'Istituto Italiano degli Attuari*, **4**, 83–91.
27. Misra,J. *et al*. (2002) Interactive exploration of microarray gene expression patterns in a reduced dimensional space. *Genome Res.*, **12**, 1112–1120.
28. Guengerich,F.P. *et al*. (2001) Reduction of aflatoxin B1 dialdehyde by rat and human aldo-keto reductases. *Chem. Res. Toxicol.*, **14**, 727–737.
29. Wakabayashi,N. *et al*. (2004) Protection against electrophile and oxidant stress by induction of the phase 2 response: fate of cysteines of the Keap1 sensor modified by inducers. *Proc. Natl Acad. Sci. USA*, **101**, 2040–2045.
30. Langouet,S. *et al*. (1997) Effects of administration of the chemoprotective agent oltipraz on CYP1A and *Cyp2b6* in rat liver and rat hepatocytes in culture. *Carcinogenesis*, **18**, 1343–1349.
31. Langouet,S. *et al*. (2000) Inhibition of human cytochrome P450 enzymes by 1,2-dithiole-3-thione, oltipraz and its derivatives, and sulforaphane. *Chem. Res. Toxicol.*, **13**, 245–252.
32. Olswang,Y. *et al*. (2002) A mutation in the peroxisome proliferator-activated receptor gamma-binding site in the gene for the cytosolic form of phosphoenolpyruvate carboxykinase reduces adipose tissue size and fat content in mice. *Proc. Natl Acad. Sci. USA*, **99**, 625–630.
33. Wang,L. *et al*. (2003) Resistance of SHP-null mice to bile acid-induced liver damage. *J. Biol. Chem.*, **278**, 44475–44481.
34. Wang,Y. *et al*. (2005) Targeted intestinal overexpression of the immediate early gene tis7 in transgenic mice increases triglyceride absorption and adiposity. *J. Biol. Chem.*, **280**, 34764–34775.
35. Huyghe,S. *et al*. (2001) Prenatal and postnatal development of peroxisomal lipid-metabolizing pathways in the mouse. *Biochem. J.*, **353**, 673–680.
36. Mizuarai,S. *et al*. (2005) Identification of dicarboxylate carrier *Slc25a10* as malate transporter in de novo fatty acid synthesis. *J. Biol. Chem.*, **280**, 32434–32441.
37. Kanehisa,M. *et al*. (2002) The KEGG databases at GenomeNet. *Nucleic Acids Res.*, **30**, 42–46.
38. Rodriguez,J.C. *et al*. (1994) Peroxisome proliferator-activated receptor mediates induction of the mitochondrial 3-hydroxy-3-methylglutaryl-CoA synthase gene by fatty acids. *J. Biol. Chem.*, **269**, 18767–18772.
39. Yamauchi,T. *et al*. (2007) Targeted disruption of AdipoR1 and AdipoR2 causes abrogation of adiponectin binding and metabolic actions. *Nat. Med.*, **13**, 332–339.
40. Rakhshandehroo,M. *et al*. (2007) Comprehensive analysis of PPARalpha-dependent regulation of hepatic lipid metabolism by expression profiling. *PPAR Res.*, **2007**, 26839.
41. Wang,L. *et al*. (2006) Orphan receptor small heterodimer partner is an important mediator of glucose homeostasis. *Mol. Endocrinol.*, **20**, 2671–2681.
42. Tontonoz,P. *et al*. (1995) PPAR gamma 2 regulates adipose expression of the phosphoenolpyruvate carboxykinase gene. *Mol. Cell. Biol.*, **15**, 351–357.
43. Jiang,J.G. *et al*. (2001) Peroxisome proliferator-activated receptor gamma-mediated transcriptional up-regulation of the hepatocyte growth factor gene promoter via a novel composite cis-acting element. *J. Biol. Chem.*, **276**, 25049–25056.

44. Sarraf,P. *et al*. (1998) Differentiation and reversal of malignant changes in colon cancer through PPARgamma. *Nat. Med.*, **4**, 1046–1052.

45. Merrell,M.D. *et al*. (2008) The Nrf2 activator oltipraz also activates the constitutive androstane receptor. *Drug Metab. Dispos.*, **36**, 1716–1721.

46. Kang,K.W. *et al*. (2003) Essential role of phosphatidylinositol 3-kinase-dependent CCAAT/enhancer binding protein beta activation in the induction of glutathione S-transferase by oltipraz. *J. Natl Cancer Inst.*, **95**, 53–66.

47. Lefterova,M.I. *et al*. (2008) PPARgamma and C/EBP factors orchestrate adipocyte biology via adjacent binding on a genome-wide scale. *Genes Dev.*, **22**, 2941–2952.

48. Park,E.Y. *et al*. (2004) Transactivation of the PPAR-responsive enhancer module in chemopreventive glutathione S-transferase gene by the peroxisome proliferator-activated receptor-gamma and retinoid X receptor heterodimer. *Cancer Res.*, **64**, 3701–3713.

49. Itoh,K. *et al*. (2004) Transcription factor Nrf2 regulates inflammation by mediating the effect of 15-deoxy-Delta(12,14)-prostaglandin j(2). *Mol. Cell. Biol.*, **24**, 36–45.