Using Nuclear Receptor Activity to Stratify Hepatocarcinogens

Imran Shah*, Keith Houck, Richard S. Judson, Robert J. Kavlock, Matthew T. Martin, David M. Reif, John Wambaugh, David J. Dix

National Center for Computational Toxicology, Office of Research and Development, United States Environmental Protection Agency, Research Triangle Park, North Carolina, United States of America

Abstract

Background: Nuclear receptors (NR) are a superfamily of ligand-activated transcription factors that control a range of cellular processes. Persistent stimulation of some NR is a non-genotoxic mechanism of rodent liver cancer with unclear relevance to humans. Here we report on a systematic analysis of new in vitro human NR activity data on 309 environmental chemicals in relationship to their liver cancer-related chronic outcomes in rodents.

Results: The effects of 309 environmental chemicals on human constitutive androstane receptors (CAR/NR1I3), pregnane X receptor (PXR/NR1I2), aryl hydrocarbon receptor (AhR), peroxisome proliferator-activated receptors (PPAR/NR1C), liver X receptors (LXR/NR1H), retinoic X receptors (RXR/NR2B) and steroid receptors (SR/NR3) were determined using in vitro data. Hepatic histopathology, observed in rodents after two years of chronic treatment for 171 of the 309 chemicals, was summarized by a cancer lesion progression grade. Chemicals that caused proliferative liver lesions in both rat and mouse were generally more active for the human receptors, relative to the compounds that only affected one rodent species, and these changes were significant for PPAR (p < 0.001), PXR (p < 0.01) and CAR (p < 0.05). Though most chemicals exhibited receptor promiscuity, multivariate analysis clustered them into relatively few NR activity combinations. The human NR activity pattern of chemicals weakly associated with the severity of rodent liver cancer lesion progression (p < 0.05).

Conclusions: The rodent carcinogens had higher in vitro potency for human NR relative to non-carcinogens. Structurally diverse chemicals with similar NR promiscuity patterns weakly associated with the severity of rodent liver cancer progression. While these results do not prove the role of NR activation in human liver cancer, they do have implications for nuclear receptor chemical biology and provide insights into putative toxicity pathways. More importantly, these findings suggest the utility of in vitro assays for stratifying environmental contaminants based on a combination of human bioactivity and rodent toxicity.

Introduction

Nuclear receptors (NR) are a superfamily of ligand-activated transcription factors that regulate a broad range of biological processes including development, growth and homeostasis. NR ligands include hormones [1] and lipids [2] but also xenobiotics [3]. We are interested in NR because of their involvement in non-genotoxic rodent liver cancer [4], a frequently observed effect in chronic toxicity testing [5] and often a critical effect in risk assessments of chemicals. Inferring the risk of chemical-induced human liver cancer from rodent studies is difficult because the underlying mechanisms are poorly understood. Persistent activation of NR is believed to be a possible mode of action [6,7] operative in various pathways leading to cancer [8]. This raises a public health concern because some environmental chemicals are human NR activators and non-genotoxic rodent hepatocarcinogens including: pesticides [9,10], persistent chemicals [11], and plastics ingredients [6]. In addition, there is very little available biological information for thousands of environmental chemicals so that new tools are needed to characterize their potential for toxicity [12–15].

We are generating human in vitro NR assay data for hundreds of environmental chemicals as a part of the ToxCast project [15]. Most of the Phase I ToxCast chemicals have undergone long-term testing experiments in rodents and their chronic hepatic effects have been curated and made publicly available in the Toxicology Reference Database (ToxRefDB) [5]. Although small sets of chemicals have been evaluated using selected NR in the past, ToxCast is the largest public data set on chemicals, encompassing concentration-dependent NR activity and chronic outcomes including liver cancer. Hence, these data provide a unique opportunity to investigate relationships between in vitro NR activation and rodent hepatic effects.

Our objective is to stratify chemicals based on their putative mode of action for human toxicity using data ranging from in vitro molecular assays to in vivo rodent outcomes from ToxCast [16] and other available resources. We have previously evaluated supervised machine learning approaches [17] and used them to classify
chemicals by chronic toxicity outcomes using in vitro data. In this analysis we used an unsupervised multivariate analysis of NR activities and rodent liver lesions to investigate a potential mode of action for non-genotoxic hepatocarcinogenesis.

Results

Nuclear Receptor Activity

Human NR activity for 309 environmental chemicals was obtained from in vitro high-throughput screening (HTS) experiments. Duplicates and triplicates for eight chemicals were included for quality control purposes. HTS data were collected for 10 out of the 48 human NR, selected based on availability of assays and potential relevance to toxicology, including: members of the NR1, NR2, NR3 and NR4 subfamilies. The aryl hydrocarbon receptor (AhR) data was also included because of its potential role in xenobiotic metabolism and non-genotoxic liver cancer [18]. A total of 54 HTS assays were used to interrogate different facets of receptor activation including: ligand binding in a cell-free system (cell-free HTS); reporter gene activation in HEK293 human cells [19] (cell-based HTS); multiplexed cis-activation and trans-activation assays for transcription factors in human HepG2 cells [20] (Multiplexed Transcription Reporter); and, multiplexed gene expression assays of xenobiotic metabolizing enzymes regulated by specific NR in primary human hepatocytes (Multiplexed Gene Expression). Data for chemical-assay pairs were collected in concentration-response format and either the AC50 concentration or the Lowest Effective concentration (LEC) were reported (additional details are provided in supplementary methods, Text S1).

Aggregate Nuclear Receptor Activity

To summarize the activity of chemicals across the NR superfamily we aggregated the ToxCast assays for genes and NR groups as follows: retinoic X receptor-like (RXR; RXRβ/β; NR2B); peroxisome proliferator-activated receptor-like (PPAR; PPARα/α; NR1C); constitutive androstane receptor (CAR; CAR1/2; NR113/4); pregnane X receptor (PXR; NR112); liver X receptor-like (LXR; LXRα/β; FXR; NR1H1); and steroid receptor-like (SR; ERα/β, ERRα/β, AR). These are shown visually in Figure 1(a). As there were differences in the number and types of assays for each group, aggregate activity was calculated as the average potency across the assays measured by the AC50 or LEC (described in Methods). This approach aggregated NR binding, activation, agonism or antagonism results into a single assessment of activity.

The aggregate activity of each of 309 chemicals was calculated across all assayed NR with the results visualized as the heatmap in Figure 1(b). In this visualization, the rows represent the NR: RXR, LXR, AhR, SR, PPAR, and PXR. Columns correspond to chemicals. The value of each cell is the aggregate scaled activity of a chemical-NR pair, and the column intensities signify the aggregate NR activity profile for each chemical (see Methods). The intensity of the colors signifies the degree of activity, where gray is inactive, yellow is the least active and red the most active. The dendrogram to the left of the NR shows their functional similarity across all 309 chemicals as two main groups. The first group contains CAR and PXR, which are most similar in their response across the chemicals, followed by AhR. The second group includes PPAR, LXR, SR and RXR. The descending order of similarity between: CAR, PXR, PPAR and SR is consistent with receptor homology, CAR and PXR are members of NR1 (thyroid hormone receptor-like), PPAR includes members of NR1C (peroxisome proliferator-activator receptor), SR represents subfamily NR3 (steroid receptor-like; estrogen and androgen). On the other hand, the activities of RXR are not similar to other NR1 members and AhR belongs to the basic Helix-Loop-Helix/Per/Arnt/Sim (bHLH-PAS) superfamily, which is distinct from NR.

Combinatorial Nuclear Receptor Activity

The chemicals were clustered by similarity of aggregate NR activity into 7 putative groups (A-G) (described in Methods). The average activity profile of the NR groups (NRG) are shown in the columns of Figure 1(c). The rows signify the NR and their order from top to bottom shows decreasing promiscuity and potency. The letters and numbers in parentheses below each column are the cluster designation and the number of chemicals in each cluster, respectively. The colors signify the activity of a NR across the NRG: red shows consistent activity and yellow inconsistent activity. For example, the first column from the left of the heatmap shows NRG A, which contains 41 chemicals that tend to activate AhR, PXR, CAR, PPAR and in some cases also SR or LXR. These results concisely describe how the 309 chemicals and 54 molecular assays can be summarized by different groups of combinatorial NR activity. The NRG correctly grouped 6 out of the 8 replicate chemicals (Table 1). For the remaining two chemicals, the duplicate Dibutyl phthalate samples had low NR activity and grouped closely in NRG F and NRG G (these samples were separately sourced substances from two different vendors). The triplicate Prosulfuron samples did not group correctly and further analysis revealed this to likely be due to degradation of the parent chemical prior to conducting the assays.

Comparing NR Activity with Cancer Lesion Progression

In vivo rat and mouse long-term histopathology outcomes for chemicals were gathered from ToxRefDB [5] and organized by severity of lesions progressing to cancer. Of the 309 ToxCast chemicals, 232 were tested in 2-year chronic feeding studies in both rat and mouse, and were characterized by liver histopathology as follows: 61 caused no observable effects and 171 chemicals caused a range of lesions of varying severity.

The 61 chemicals negative for any liver injury include: Ethylluralin, Fenamiphos, Isazofos, and Propetamphos (NRG A); Cyafoxamid and Fenhexamid (NRG B); Fenpyroximate, Rotenone, Tebupirimfos (NRG C); and (51/61) in NRG D, E, F and G (see Dataset S4). Since the absence of rat or mouse liver toxicity is unusual after sustained treatment with a chemicals for two years, it can indicate an insufficient treatment dose (among other factors). When we reviewed the treatment protocols for these 61 chemicals we found that 7/10 chemicals in NRG A, B and C may have been administered at insufficient doses to produce hepatic effects. For example, Rotenone is a potent mitochondrial inhibitor and commonly used as a pesticide. It can cause rodent gastrointestinal injury at roughly 150 parts per million (ppm), however, it was only tested at a maximum dose of 3.75 ppm in the chronic study. Hence, we could not be certain about the absence of liver toxicity for these 61 chemicals despite a lack of nuclear receptor activity in a majority of 51 cases.

Lesion Progression and Nuclear Receptor Activities

We assumed that dose selection was not an issue for the 171 chemicals that produced at least some liver toxicity in chronic rodent testing. Out of these 171 chemicals, 66 were mild hepatotoxins, 43 produced different grades of proliferative lesions in rat and mouse, and 13 chemicals caused neoplastic lesions in both species. The severity and concordance of hepatic lesions across these 171 chemicals were clustered by similarity into eight lesion progression groups shown in Figure 2(c) (see Methods). The aggregate NR activities were systematically compared across all lesion progression groups (LPG) and visualized in Figure 3. The rows in Figure 3 correspond to the eight lesion progression groups...
(LPG I, II, III, IV, V, VI, VII, VIII) shown in Figure 2(c), and the columns are the NR: AhR, CAR, PXR, PPAR, LXR, SR, RXR. Each cell in the heatmap shows the ratio of the mean NR activities of chemicals in a LPG compared to all other LPG. The statistical significance of differences in mean NR activity was evaluated by permutation and corrected for multiple testing (see Methods). AhR, PPAR, SR and RXR showed 9% to 250% higher average activity for chemicals in LPG I as compared to the other chemicals but only PPAR showed a statistically significant ($p < 0.001$) increase of 150%. For LPG II chemicals, all NR showed some increased activity.
Table 1. Chemicals grouped by nuclear receptor activity and lesion progression.

| A | B              | C              | D              | E              | F                                 | G                                           |
|---|----------------|----------------|----------------|----------------|-----------------------------------|---------------------------------------------|
| I | Fludioxonil    | Diclofop-methyl| Diethylhexyl   | Carbaryl       | Isoxaflutole                      | 2,5-Pyridinedicarboxylic-acid, dipropyl ester |
|   | Lactofen       | Diclofop-methyl| phthalate      |                | Pyometrozine                      |                                             |
|   | Oxadiazon      | Diclofop-methyl|                |                | Tepraloxydim                      |                                             |
|   | Imazalil       |                |                |                |                                   |                                             |
|   | Malathion      |                |                |                |                                   |                                             |
|   | Vinclozolin    |                |                |                |                                   |                                             |
| II| Bensulide      | Fentin         | Buprofezin     | Fenamidone     | Butafenacil                       | Clodinafop-propargyl                        |
|   | Bensulide      | Fluazinam      | Fenarimol      |                | Diphenylamine                     | Pyriothiobac-sodium                         |
|   | Bensulide      | Spirodiclofen  | Fluthiacet-methyl|              | Fenoxycarb                        |                                             |
|   | Dithiopyr      |                |                |                |                                   |                                             |
|   | MGK            |                |                |                |                                   |                                             |
|   | Triflumizole   |                |                |                |                                   |                                             |
| III| Indoxacarb     | Bromoxynil     | Lindane        | Clofentezine   | Dicofol                           | 3-lodo-2-propynylbutyl-carbamate             |
|    | Iprodione      |                |                |                |                                   |                                             |
|    | Linuron        |                |                | Prochloraz     | Nitrpyrin                         | Dazomet                                      |
|    | Propiconazole  |                |                | Propyzamide    |                                   | Fenoxaprop-ethyl                            |
|    | Thiazopyr      |                |                |                |                                   |                                              |
| IV | Isoxaben       | Cinmethylin    | Hesythiazox    | Benfluralin    | Maneb                             | 2-Phenyl/phenol                             |
|    | Methiathion    |                |                | Benomyl        |                                   | Acephate                                     |
|    | Triadimefon    |                |                | Bifenazate     | Propoxur                          | Amitraz                                     |
|    | Triadimenol    |                |                | Bromacil       | Terbacil                          | Bentazone                                   |
|    | Tribufos       |                |                | Fenitrothion   |                                   | Cloprop                                      |
|    |                |                |                | Norflurazon    |                                   | Daminozide                                   |
|    |                |                |                | Thiophanate-methyl| Dimethoate                     |                                              |
|    |                |                |                | Triflusuron-methyl| Triclofibar                    |                                              |
| V  | Cyclanilide    | Tebufenpyrad   | Ametryn        | Acetochlor     | Dichlobenil                       | Mevinphos                                   |
|    |                |                |                | Dimethenalid   | Simazine                          |                                              |
|    |                |                |                |                |                                   |                                              |
| VI | Tetraconazole  |                |                |                |                                   |                                              |
|    |                |                |                |                |                                   |                                              |
| VII| Flutanil       | Bisphenol A    | Carboxin       |                | 2,4-DB                            | Acetamiprid                                 |
|    | Oxyfluorfen    |                | Carfentrazone-ethyl|            | Butylate                          | Asulam                                      |
|    | Triticonazole  | d-cis,trans-Allthrin| Diuron    | Chlorypyros-methyl | Azamethiphos                     |                                              |
|    |                |                |                |                | Fipronil                          | Clorophene                                  |
|    |                |                |                |                | Metaxyl                          | Cymoxanil                                    |
|    |                |                |                |                | Prallethin                        | Hexazinone                                   |
|    |                |                |                |                | Flufenpyr-ethyl                   | Mesoosulfuron-methyl                          |
|    |                |                |                |                | Prosulfuron                       | Fluimiclor-pentyl                            |
|    |                |                |                |                | Sulfentrazene                     | Proxaluron                                   |
|    |                |                |                |                | Trifloxystrobin                   | Myclobutanil                                 |
|    |                |                |                |                | Prometon                          | Thidiazuron                                   |
|    |                |                |                |                | Prosulfuron                       |                                              |
|    |                |                |                |                | Tefluthrin                        |                                              |
except LXR, but only PPAR and PXR had statistically significant (p < 0.05) increases in activity of 80% and 50%, respectively. There were no statistically significant differences in NR activities for chemicals that produced only mouse proliferative lesions, however, the subset of mouse carcinogens showed a 30% increase in AhR activity but a 30% decrease in PPAR activity. Chemicals that produced only rat hepatic neoplasms had a 75% increase in PPAR activity, 23% increase in CAR activity and 30% increase LXR activity but none were statistically significant.

**Lesion Progression and Nuclear Receptor Activity Groups**

The comparison between the LPG and NRG between 171 chemicals is visualized in Figure 4(c). The rows in Figure 4(c) are the eight lesion progression groups (LPG I, II, III, IV, V, VI, VII, VIII) shown in Figure 4(b) and the columns are the seven NR activity groups (NRG A, B, C, D, E, F, and G) shown in Figure 4(c). Each circle represents chemicals that have similar human NR activity and degree of rodent lesion progression. The size of each circle visualizes the proportion of chemicals across the LPG (rows) and NRG (columns), while the color signifies confidence in assignment of chemicals to each group (see Methods).

We designate each joint group by concatenating the identifier as: LPG-NRG, and interpret the first row of Figure 4(c), which corresponds to LPG I. The first circle from the left represents group I-A, which is formed by the intersection of 13 chemicals in LPG I and 29 chemicals in NRG A. The three chemicals in I-A (shown in the first row and first column of Table 1) were most consistently active with PPAR, but also showed some activity with PXR, CAR, and LXR. Oxadiazon is a herbicide with known human PXR activity [21]. Lactofen is a poly-phenyl herbicide with PPAR activity in mice [22]. Similarly, I-B, the second circle from the left, is the intersection of 13 chemicals in LPG I and 10 chemicals in NRG B. Chemicals in I-B were the most consistently active with PPAR, but also showed some activity with PXR, CAR, AhR and SR. Imazalil is a triazole fungicide that perturbs human genes regulated by AhR [23] and is also a PXR activator [21]; malathion is an organophosphorus pesticide with known SR activity [24]; vinclozolin, a dicarboximide fungicide is also a known SR activator [25]; and diclofop-methyl has been shown to be PPAR active in rats. Group I-D only contains diethylhexylphthalate (DEHP), which is a key plastics monomer, and has been shown to activate PPAR [27], PXR [28], and CAR [29]. In group I-E we have carbaryl, which is a carbamate insecticide with AhR [30] and SR [31] activity. Lastly, chemicals in I-F and I-G had negligible NR activity, which could suggest that they act through other pathways.

Chemicals in LPG II produced only putative pre-neoplastic liver lesions in rat and mouse but there is limited prior knowledge about their NR activities.

LPG III only contains mouse hepatocarcinogens predominantly active with AhR, PXR and CAR, but some propensity for PPAR, LXR and SR. In III-A, the dicarboximide fungicide, iprodione, has been shown to activate AhR in human HepG2 cells [32]; linuron activates CAR in mice [33]; CAR in rat [34], and the triazole fungicide, propiconazole, activates CAR, PXR and PPAR in mice [35]. The four chemicals in III-D namely, permethrin, lindane, prochloraz and propyzamide, are most consistently active for CAR, followed by AhR and PXR. In hepatocytes, permethrin [36] and lindane [37] induce expression of the CAR, AhR and PXR target xenobiotic metabolizing enzymes (XME), CYP2B6, CYP1A1/2, and CYP3A4, respectively. Prochloraz has only been observed to activate CAR and AhR [37]. Chemicals in III-E have lesser overall NR activity but are generally more active with AhR and to a lesser degree with CAR, PXR and PPAR. One of the chemicals in III-D, cyproconazole, has been shown to induce expression of a cytochrome P-450 in the 2B subfamily (CYP2B10), an XME regulated by CAR across different mouse strains [38], however, the expression of CYP1A1/2 was not measured in this study.

The relationship between NR activation and cancer lesion progression is visualized by the location and size of circles: when the NR activity is greatest (NRG A), many of the chemicals are rodent hepatocarcinogens (LPG I) or just mouse carcinogens (LPG I-V); and when NR activity is the least (NRG G), most of the chemicals produce mild or no lesions (LPG VII, VIII). For intermediate grades of NR activity (NRG B-F), the relationships are more complex: PPAR, PXR and SR activators (NRG B) produced stage (iii) lesions (neoplastic) in both species (LPG I, V); most CAR and PXR (NRG D) activators produced stage (ii) lesions but some were also hepatocarcinogens; AhR, CAR and PXR activators (NRG E) were mostly mouse hepatocarcinogens.

More importantly, the association of LPG I through VIII with NRG A through G, shown in Figure 4(c) is statistically significant with a p-value of 0.013 using Fisher's exact test. There is greater than 95% confidence that the observations on human nuclear receptor activity and rodent cancer lesion progression are not by chance alone.

---

**Table 1. Cont.**

| A            | B            | C            | D            | E            | F            | G            |
|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| VIII         | Cyprodinil   | Dimethomorph | Alachlor     | Acibenzolar-S-Methyl | Emamectin benzoate | 2,4-Dichlorophenoxyacetic acid (2,4-D) |
| Et oxazole   | S-Bialaphthrin | Picloram     | Icaridin     | Chlorsulfuron |
| Flutemarlat  | Ethofumesate  | Thiabendazole| Paclotinuzole | Chlorsulfuron |
| Hexaconazole | Flusilazole   | Penoxoalum   | Cyhalofop-butyl |
| Methoxyfenozide | Fosthiazate | Triclosan     | Dichlorprop   |
| Phosalone    |              |              |              |              |              |              |
| Pyraclostrobin |              |              |              |              |              |              |
| Tebufenozide |              |              |              |              |              |              |

Chemicals assigned to nuclear receptor groups (columns) and lesion progression groups (rows).

*doi:10.1371/journal.pone.0014584.t001*
Discussion

Chemical-induced activation of NR has been evaluated previously using HTS [3,39,40] but ToxCast is the largest publicly available data set in terms of chemicals (309), number and diversity of NR activities (7), NR assays (54), and associated rodent in vivo toxicity data in ToxRefDB [5]. By analyzing the data, we show that these chemicals concurrently activate multiple members of the NR superfamily (NRG) in combinations that have not been possible to systematically describe before. Since the 309 chemicals may not be a representative sample of all environmental pollutants and because we did not measure all NR, it is difficult to say...
whether these nuclear receptor groups (NRG) are universal. Yet our findings were generally consistent with what is known about the NR activities for some chemicals.

Histopathologic observations in the liver have been also been organized by severity for acute [4] and chronic injury in the past. In our analysis, we integrated diverse phenotypic observations of disease symptoms progressing from adaptive changes to neoplastic lesions. In addition, we also summarized cancer progression data across rat and mouse to contrast subtle differences in the severity of adverse chronic outcomes. While this simplified the computational analysis of phenotypic data, it also represents three possible limitations. First, all stages of lesion progression may not have been observed at the terminus of a chronic bioassay. Second, we did not consider the impact of gender and developmental stages, which can be quite important in chemical carcinogenesis. Third, we did not use information about the concentration at which lesions were observed. This may be especially problematic for chemicals that are dose limited (e.g. acetylcholinesterase inhibitors, many of which are in the current data set), so that doses that might lead to liver toxicity are never reached.

Finding robust relationships in real datasets is difficult because measurements can be noisy or irrelevant, and observations can be uninformative. While our analysis is not immune from these issues we tried to mitigate their influence in two main ways. First, we combined data on disparate molecular assays into an aggregate measure of NR activity. The accuracy of this aggregate activity can be demonstrated by the correct categorization of most replicate chemicals into the same NRG (see Table 1.), despite differences in NR assay profiles. Second, we grouped sparse observations on histopathologic effects into three stages of lesion severity in hepatocarcinogenesis. By independently organizing the observations at these disparate biological scales, we found coherent bioactivity profiles in relation to pathologic states.

Our findings have three main implications for toxicity testing. First, it may be important to screen chemicals for multiple NR activities for assessing the hazard of non-genotoxic liver cancer. Second, the visualization in Figure 4 suggests a possible approach for interpreting disparate NR assays in the context of rodent liver cancer severity, and also shows the uncertainties in using these data for chemical prioritization. Third, NR activation by environmental chemicals may be more conserved between rodents and humans than previously believed [42]. This is corroborated partly by comparison with the literature and also by similarities between the aggregate activities of nuclear receptors across chemicals, which appear to recapitulate their evolutionary relationships (Figure 3(b)). Such a gradual functional divergence in the NR superfamily is consistent with protein evolution [43] but it may also lead to conservation of NR activities between rodents and humans. Relating these responses to divergent phenotypic outcomes, however, requires a deeper understanding of non-genotoxic pathways to cancer.

Chronic animal testing is infeasible for the many thousands of chemicals in commerce, but it is currently the gold-standard for estimating human cancer risk. The EPA ToxCast program is systematically assessing the value of high-throughput technologies for...
screening environmental chemicals’ ability to impact toxicity pathways leading to human diseases such as cancer. Our objective was to develop a tool for efficiently stratifying thousands of environmental chemicals based on their perturbation of events leading to adverse outcomes. Here we focused on liver cancer because it is frequently observed across the 309 ToxCast chemicals, and on NR activity since it is a putative key event in rodent carcinogenesis. Through a unique analysis of these data we found that human NR activity profiles for the chemicals stratified their liver cancer lesion progression in rodents. This relationship between the in vitro molecular assays to in vivo rodent outcomes identifies putative mode of action, advances our understanding of nuclear receptor interactions with environmental chemicals, and suggests approaches for efficient tiered testing for environmental carcinogens.

**Methods**

**Multiplexed Gene Expression in Human Primary Hepatocytes**

This is a collection of multiplexed gene expression assays focused on Phase I and II xenobiotic metabolizing enzymes and transporters. Human primary cell cultures were treated with
chemicals at 5 concentrations (0.004–40 μM) for 6, 24 and 48 hr. Concentration- and time-response profiles of chemicals were measured by the expression of key nuclear receptor target genes, activities of CYP1A enzymes (EROD), and cell morphology. Fourteen gene targets were monitored by quantitative nucleic acid protection assay including six representative cytochrome P-450 genes, four hepatic transporters, three Phase II conjugating enzymes, and one endogenous metabolism gene involved in cholesterol synthesis. The target genes associated with nuclear receptor pathways are as follows: CYP1A1 and CYP1A2 with AhR, ABCB1, ABCC2, CYP2B6, CYP2C9, CYP2C19 and UGT1A1 with CAR, CYP3A4, GSTA2, SLC01B1 and SULT2A1 with PXR, HMGC2S2 with PPAR, and ABCB1 with FXR. Assays were run in primary human hepatocyte cultures by CellDirect Invitrogen Inc. (Durham, NC), in collaboration with EPA.

Multiplexed Transcription Reporter Assays
A multiple reporter transcription unit (MRTU) library consisting of 48 transcription factor binding sites was transfected into the HepG2 human liver hepatoma cell line [20]. In addition to the cis-acting reporter genes (CIS), a modification of the approach was used to generate a trans-system (TRANS) with a mammalian one-hybrid assay consisting of an additional 25 MRTU library reporting the activity of nuclear receptor super-family members. Based on an initial cytotoxicity screen, the maximum tolerated concentration (MTC) was derived as one-third the calculated IC50 or, if no IC50 was determined, the MTC was set to 100 μM. Chemicals were then tested in the CIS and TRANS assays at seven concentrations starting at the MTC and followed by three-fold serial dilutions. These assays were performed by Attagene Inc. (Morrisville NC) under contract to EPA.

Cell-free HTS Assays
These are a collection of biochemical assays measuring binding constants and enzyme inhibition values. Chemicals were initially screened at a single concentration in duplicate wells at a concentration of 10 μM for cytochrome P-450 assays and 25 μM for all others. Chemicals that showed significant activity were then run in concentration response format, from which an AC50 value was extracted. For concentration response, 8 concentrations were tested in the ranges 0.00914–20 μM for cytochrome P-450 assays and 0.0229–50 μM for other assays. These assays were run by Caliper Life Sciences (Hanover, MD) under contract to EPA. Short assay descriptions are available at: http://www.caliperls.com/products/contract-research/in-vitro/.

Cell-based HTS Assays
These assays measure binding constants and enzyme inhibition values for nuclear receptors. The targets include AR, ER, FXR, LXR, PPAR, PPARβ, PPARγ, RXRα, RXRβ and PXR. Each of the nuclear receptor targets was measured in agonist mode. Assays were run at the NIH Chemical Genomics Center (Rockville, MD).

In vitro data
All data used in this analysis are publicly available from the ToxCast website (www.epa.gov/toxcast). The analysis was conducted using the R statistical language (www.r-project.org). For each chemical, a representative combination was derived either the AC50 (50% maximal activity concentration) or LEC (Lowest Effective Concentration) in μM denoted as, cij. All chemicals are provided in Dataset S3 and all assays are given in Dataset S1. The procedure for evaluating the quality of cij are described in the supplementary methods (Text S1) and all assay results are provided in Dataset S2. In order to facilitate comparison across the assays the cij were transformed by the formula, cij′ = log10 (max(s))−log10(cij) where cij′ represents a the potency on an ascending scale, and max(s) is the maximum concentration (lowest potency) across the data set.

Aggregating assay results
The aggregate NR activity φ, where φ ∈ [CAR, PXR, PPAR, AhR, SR, RXR], was calculated using assays aij. The aggregate scaled NR activity score for each chemical xij was calculated as the average concentration value across the assays, scaled by the maximum value across all chemicals using Equation 1.

\[ x_{ij} = \frac{1}{n_p^a} \sum_{a=1}^{n_p} c_{ij} \]

\[ \max_{a} \left( \frac{1}{n_p} \sum_{a=1}^{n_p} c_{ij} \right) \]

The number of assays in aij is given by nφ. Hence, the complete NR activity of each chemical was defined as a vector using Equation 2.

\[ x_i = \{ x_{i\text{CAR}}, x_{i\text{PXR}}, x_{i\text{PPAR}}, x_{i\text{AhR}}, x_{i\text{SR}}, x_{i\text{RXR}, x_{i\text{LXR}}} \} \]

The NR activities of all chemicals were expressed as a matrix X, where the rows are chemicals (φ) and columns are the NR activities (ϕ).

Lesion Progression in Hepatocarcinogenesis
In vivo rat and mouse long-term histopathology outcomes were extracted for 171 chemicals from ToxRefDB and organized by severity into three stages [5] including: (i) non-proliferative, (ii) putative pre-neoplastic and (iii) neoplastic lesions. For each chemical, the incidence of hepatic tissue lesions was summarized across rat and mouse species as a 6-dimensional cancer lesion progression vector (LPV), which is depicted in Figure 2(a). The resulting 37 unique LPV are shown in Figure 2(b). These LPV were clustered by similarity (described below) into eight lesion progression groups (LPG). The LPG are visualized as a heatmap in Figure 2(c) and denoted by uppercase Roman numerals. The colors of each cell in this heatmap are proportional to the number of chemicals in the LPG that induce a specific lesion type.

Statistical tests
In addition to the two-sided t-test, 10^6 permutations were carried out to empirically estimate the significance (type-I error) of reported univariate statistics. The p-values calculated in this way were adjusted for multiple comparisons using the false discovery rate [44] (FDR) correction. The statistical significance of associations between NRG and LPG were evaluated using Fisher’s exact test.

Clustering
Hierarchical clustering was carried out using the Euclidean distance metric and Ward’s minimum variance method for agglomeration. The LPV and the NR data sets were partitioned using k-means [45] clustering for K = 2 to K = 50, the resulting partitions were analyzed using the silhouette method [46], and
the value of k was selected by examining the average cluster width. This procedure was used to partition chemicals into groups of NR activity, $NRG = \{A, B, C, D, E, F, G\}$, and groups of cancer lesion progression, $LPG = \{I, II, III, IV, V, VI, VII, VIII\}$.

Cluster stability

The assignment of chemicals to NRG and LPG was evaluated by a cluster stability score, which was calculated using a subspace sampling approach [47]. A subspace, $q$, of the data was defined by randomly selecting $\phi^q$ assays and $\psi^q$ chemicals, where $\phi^q \subset \phi$ and $\psi^q \subset \psi$. For the chemicals $\phi^q$, the aggregate scaled activities across $\phi^q$ were calculated using Equation 1 to create the subspace data matrix $X^q$. The matrix $X^q$ was analyzed by k-means clustering (see above), and chemicals $\phi^q$ were then assigned to subspace clusters $\rho^q$. The partitions in $\rho^q$ were matched with the partitions $\rho$ from the complete data set based on the maximum number of common members. That is, $\rho^q \equiv \rho_k$ when $\max(|\psi^q \cap \psi_k|) > 0$, and $f^q_k$ is the fraction of chemicals from $\rho$ in $\rho^q$. The subspace sampling was conducted $N_q = 1000$ times and the quality $\omega_k$ of the partition $\rho_k$ was calculated as $\omega_k = \frac{1}{N_q} \sum f^q_k$.

Disclaimer

The United States Environmental Protection Agency through its Office of Research and Development reviewed and approved this publication. Reference to specific commercial products or services does not constitute endorsement.

Supporting Information

Text S1 Supplementary Methods. Found at: doi:10.1371/journal.pone.0014584.s001 (0.03 MB TXT)

Dataset S1 Description of assays. Found at: doi:10.1371/journal.pone.0014584.s002 (0.01 MB TXT)

Dataset S2 Assay results for each chemical. Found at: doi:10.1371/journal.pone.0014584.s003 (0.06 MB TXT)

Dataset S3 Description of all chemicals used in the analysis. Found at: doi:10.1371/journal.pone.0014584.s004 (0.05 MB TXT)

Dataset S4 Nuclear Receptor Groups for chemicals negative for any liver toxicity. Found at: doi:10.1371/journal.pone.0014584.s005 (0.00 MB TXT)

Acknowledgments

We would like to thank Drs. David Malarek (NIHEIS) and Doug Wolf (EPA) for helpful discussions on histopathology and the progression to cancer. We also thank our Tox21 partners at the NIH Chemical Genomics Center, especially Drs. Menghong Xia and Ruili Huang, for generation of much of the chemical-nuclear receptor interaction data.

Author Contributions

Conceived and designed the experiments: IS JW DJD. Performed the experiments: IS KH JW. Analyzed the data: IS KH RSJ MTM DMR JW DJD. Contributed reagents/materials/analysis tools: KH RSJ MTM JW. Wrote the paper: IS KH RSJ RJK MTM DMR JW DJD. Contributed to the study design: KH RJK MTM DJD. Processed ToxCast data for analysis: RSJ. Processed the ToxReDB data for analysis: MTM DMR. Contributed to the statistical analysis: DJD. Edited the manuscript: KH DJD.

References

1. Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schutz G, et al. (1995) The nuclear receptor superfamily: the second decade. Cell 83: 835–839.
2. Whitfield GK, Jaruzak PW, Hauser CA, Hauser MR (1999) Steroid hormone receptors: evolution, ligands, and molecular basis of biologic function. Journal of Cellular Biochemistry 75.
3. Timsit YE, Negishi M (2007) CAR and PXR: the Xenobiotic-Sensing receptors. Steroids 72: 231246.
4. Oliver JD, Roberts RA (2002) Receptor-mediated hepaticcarcinogenesis: role of hepatoocyte proliferation and apoptosis. Pharmacology & Toxicology 91: 1–7.
5. Martin MT, Judson RS, Reif DM, Kavlock RJ, Dix DJ (2009) Profiling chemicals based on chronic toxicity results from the U.S. EPA ToxRef database. Environmental Health Perspectives 117: 392–399.
6. Klaunig JE, Balbich MA, Baetcke KP, Cook JC, Corton JC, et al. (2003) PPARalpha agonist-induced rodent tumors: modes of action and human relevance. Critical Reviews in Toxicology 33: 655–780.
7. Elcombe CR, Outham J, Foster JR, Stone S, Hannall S, et al. (2002) Prediction of rodent nongenotoxic carcinogenicity: evaluation of biochemical and tissue changes in rodents following exposure to nine nongenotoxic NTP carcinogens. Environmental Health Perspectives 110: 363–75.
8. Hanahan D, Weinberg RA (2000) The hallmarks of cancer. Cell 100: 57–70.
9. Kretschmer XG, Baldwin WS (2005) CAR and PXR: xenosensors of endocrine disruptors? Chemico-biological interactions 155: 111–128.
10. Goetz AK, Dix DJ (2009) Mode of action for reproductive and hepatic toxicity inferred from a genomic study of triazole antifungals. Toxicological Sciences: An Official Journal of the Society of Toxicology 95: 37–67.
11. Abdelfattah AG, Prat V, Turner HS, Roberfroid M (1991) The modulation of rat liver carcinogenesis by peroxyacetic acid, a peroxione proliferator. Toxicology and Applied Pharmacology 111: 530–537.
12. Combes R, Barratt M, Bulb M (2006) An overall strategy for the testing of chemicals for human hazard and risk assessment under the EU REACH system. Alternatives to Laboratory Animals: ATLA 34 Suppl 1: 15–27.
13. Doe JE, Boobis AR, Blacker A, Delbaro V, Doerrer NG, et al. (2006) A tiered approach to systemic toxicity testing for agricultural chemical safety assessment. Critical reviews in toxicology 36: 37–68.
14. Judson R, Richard A, Dix DJ, Houck K, Martin M, et al. (2009) The toxicity data landscape for environmental chemicals. Environmental Health Perspectives 117: 685–695.
15. Dix DJ, Houck KA, Martin MT, Richard AM, Setzer RW, et al. (2007) The ToxCast program for prioritizing toxicity testing of environmental chemicals. Toxicological sciences: an official journal of the Society of Toxicology 95: 5–12.
16. Judson RS, Houck KA, Kavlock RJ, Knudsen TB, Martin MT, et al. (2010) In vitro screening of environmental chemicals for targeted testing prioritization: the ToxCast project. Environmental health perspectives 118: 405.
17. Judson R, Elloson F, Setzer RW, Li Z, Shah I (2008) A comparison of machine learning algorithms for chemical toxicity classification using a simulated multi-scale data model. BMC Bioinformatics 9: 241.
18. Nebert DW, Dalton TP, Okey AB, Gonzalez EF (2004) Role of aryl hydrocarbon receptor-mediated induction of the CYP1 enzymes in environmental toxicity and cancer. J Biol Chem 279: 23847–23850.
19. Ingelse J, Johnson RL, Simeonov A, Xia M, Zheng W, et al. (2007) High-throughput screening assays for the identification of chemical probes. Nat Chem Biol 3: 466–470.
20. Romanov S, Medvedev A, Gambharim M, Poltoratskaya N, Moeser M, et al. (2008) Homogenous reporter system enables quantitative functional assessment of multiple transcription factors. Nat Meth 5: 253–260.
21. Lemarie G, Maar W, Pascussi J, Bollon A, Rabenoldina F, et al. (2006) Identification of new human pregnancy receptor ligands among pesticides using a stable reporter cell system. Toxicological Sciences: An Official Journal of the Society of Toxicology 91: 501–509.
22. Butler EG, Tanaka T, Ichida T, Maruyama H, Leber AP, et al. (1988) Induction of hepatopni carcinoma proliferation in mice by lactofen, a diphenyl ether herbicide. Toxicology and Applied Pharmacology 93: 72–80.
23. Sergent T, Dupont I, Jassogne C, Ribonett L, van der Heiden E, et al. (2009) CYP1A1 induction and CYP3A4 inhibition by the fungicide imazalil in the human intestinal caco-2 cell-competition with other conazole pesticides. Toxicology Letters 184: 159–168.
24. Chen H, Xiao J, Hu G, Zhou J, Xiao H, et al. (2002) Estrogenicity of organophosphorus and pyrethroid pesticides. Journal of Toxicology and Environmental Health Part A 65: 1419–1455.
25. Klotz WR, Monosson E, Gomesak MP, Laws SC, Gray LE (1994) Identification of nuclear receptors in cancer.
26. Takeuchi S, Matsuda T, Kobayashi S, Takahashi T, Kojima H (2006) In vitro screening of 200 pesticides for agonistic activity via mouse peroxisome proliferator-activated receptor (PPAR)alpha and PPARgamma and quantitative analysis of in vivo induction pathway. Toxicology and Applied Pharmacology 217: 233–244.

27. Martinasso G, Maggiora M, Trombetta A, Canuto RA, Muzio G (2006) Effects of di(2-ethylhexyl) phthalate, a widely used peroxisome proliferator and plasticizer, on cell growth in the human keratinocyte cell line NCTC 2544. Journal of Toxicology and Environmental Health Part A 69: 353–365.

28. Hurst CH, Wexman DJ (2004) Environmental phthalate monoesters activate pregnane x receptor-mediated transcription. Toxicology and Applied Pharmacology 199: 266–274.

29. Evrillard A, Moletti-Lakhal I, Mogha A, Lasserre F, Polizzi A, et al. (2009) di(2-ethylhexyl) phthalate (DEHP) activates the constitutive androstane receptor (CAR): a novel signalling pathway sensitive to phthalates. Biochemical Pharmacology 77: 1735–1746.

30. Boronat S, Casado S, Navas JM, Pia B (2007) Modulation of aryl hydrocarbon receptor transactivation by carbaryl, a nonconventional ligand. The FEBS Journal 274: 3327–3339.

31. Lemaire G, Meul W, Mauvais P, Balaguer P, Rahmani R (2006) Activation of alpha- and beta-estrogen receptors by persistent pesticides in reporter cell lines. Life Sciences 79: 1160–1169.

32. Washington T, Tchounwou PB (2004) Cytotoxicity and transcriptional activation of stress genes in human liver carcinoma (HepG2) cells exposed to iprodione. International Journal of Environmental Research and Public Health 1: 12–20.

33. Takeuchi S, Iida M, Yabushita H, Matsuda T, Kojima H (2008) In vitro screening for aryl hydrocarbon receptor agonistic activity in 200 pesticides using a highly sensitive reporter cell line, DR-EcoScreen cells, and in vivo mouse liver cytochrome P450-1A induction by propanil, diuron and linuron. Chemosphere 74: 155–165.

34. Scollon EJ, Starr JM, Godin SJ, DeVito MJ, Hughes MF (2009) In vitro metabolism of pyrethroid pesticides by rat and human hepatic microsomes and cytochrome p450 isozymes. Drug Metabolism and Disposition: The Biological Fate of Chemicals 37: 221–228.

35. Dubois M, Plaisance H, Thom JP, Kremers P (1996) Hierarchical cluster analysis of environmental pollutants through p450 induction in cultured hepatic cells. Ecotoxicology and Environmental Safety 34: 205–215.

36. Scollon EJ, Starr JM, Godin SJ, DeVito MJ, Hughes MF (2009) In vitro metabolism of pyrethroid pesticides by rat and human hepatic microsomes and cytochrome p450 isozymes. Drug Metabolism and Disposition: The Biological Fate of Chemicals 37: 221–228.

37. Dubois M, Plaisance H, Thom JP, Kremers P (1996) Hierarchical cluster analysis of environmental pollutants through p450 induction in cultured hepatic cells. Ecotoxicology and Environmental Safety 34: 205–215.

38. Peffer RC, Moggs JG, Pastoor T, Currie RA, Wright J, et al. (2007) Mouse liver effects of cyproconazole; a triazole fungicide; role of the constitutive androstane receptor. Toxicological Sciences: An Official Journal of the Society of Toxicology 99: 315–325.

39. Parker GJ, Law TL, Lenoch FJ, Bolger RE (2000) Development of high throughput screening assays using reverse transcription: Nuclear Receptor-Ligand Binding and Kinase/Phosphatase assays. J Biomol Screen 5: 77–86.

40. Zhu Z, Kim S, Chen T, Liu J, Bell A, et al. (2004) Correlation of High-Throughput pregnane x receptor (PXR) transactivation and binding assays. J Biomol Screen 9: 533–540.

41. Lohnhöfer EK, Boerman GA, Phillips KL, Heindlof AN, Malacray DE, et al. (2006) Application of visualization tools to the analysis of histopathological data enhances biological insight and interpretation. Toxicologic pathology 34: 921–8.

42. Maglich JM, Stolz CM, Goodwin B, Hawkins-Brown D, Moore JT, et al. (2002) Nuclear pregnane x receptor and constitutive androstane receptor regulate overlapping but distinct sets of genes involved in xenobiotic detoxification. Mol Pharmacol 62: 638–646.

43. Tokuniku N, Tasfik DS (2009) Protein dynamism and evolvability. Science (New York, NY) 324: 203–207.

44. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. .

45. Hartigan JA, Wong MA (1979) Algorithm AS 136: A k-means clustering algorithm. Applied Statistics. pp 100–108.

46. Rousseau PJ (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics 20: 53–65.

47. Parsons L, Haque E, Liu H (2004) Subspace clustering for high dimensional data: a review. SIGKDD Explor Newsl 6: 90–105.