TOXICOLOGICAL SCIENCES, 175(1), 2020, 98–112

doi: 10.1093/toxsci/kfaa026
Advance Access Publication Date: March 2, 2020
Research Article

Quantitative Transcriptional Biomarkers of Xenobiotic Receptor Activation in Rat Liver for the Early Assessment of Drug Safety Liabilities

Alexei A. Podtelezhnikov,*,[1] James J. Monroe,† Amy G. Aslamkhan,† Kara Pearson,† Chunhua Qin,† Alex M. Tamburino,* Andrey P. Loboda,* Warren E. Glaab,† Frank D. Sistare,†,[2] and Keith Q. Tanis*,[3]

*Human Genetics and Pharmacogenomics; and †Safety Assessment and Laboratory Animal Resources, Merck & Co., Inc., West Point, Pennsylvania 19486-0004

[1]To whom correspondence should be addressed at Human Genetics and Pharmacogenomics, Merck PO Box 4, WP53B-120 770 Sumneytown Pike West Point, PA 19486-0004. Fax: (215) 993-1835. E-mail: alexei_podtelezhnikov@merck.com. [2]Safety Assessment and Laboratory Animal Resources, Merck PO Box 4, WP81-200 770 Sumneytown Pike West Point, PA 19486-0004. Fax: (215) 993-7563. E-mail: frank_sistare@merck.com.[3]Human Genetics and Pharmacogenomics, Merck PO Box 4, WP53B-120 770 Sumneytown Pike West Point, PA 19486-0004. Fax: (215) 993-1835. E-mail: keith_tanis@merck.com.

ABSTRACT

The robust transcriptional plasticity of liver mediated through xenobiotic receptors underlies its ability to respond rapidly and effectively to diverse chemical stressors. Thus, drug-induced gene expression changes in liver serve not only as biomarkers of liver injury, but also as mechanistic sentinels of adaptation in metabolism, detoxification, and tissue protection from chemicals. Modern RNA sequencing methods offer an unmatched opportunity to quantitatively monitor these processes in parallel and to contextualize the spectrum of dose-dependent stress, adaptation, protection, and injury responses induced in liver by drug treatments. Using this approach, we profiled the transcriptional changes in rat liver following daily oral administration of 120 different compounds, many of which are known to be associated with clinical risk for drug-induced liver injury by diverse mechanisms. Clustering, correlation, and linear modeling analyses were used to identify and optimize coexpressed gene signatures modulated by drug treatment. Here, we specifically focused on prioritizing 9 key signatures for their pragmatic utility for routine monitoring in initial rat tolerability studies just prior to entering drug development. These signatures are associated with 5 canonical xenobiotic nuclear receptors (AHR, CAR, PXR, PPARα, ER), 3 mediators of reactive metabolite-mediated stress responses (NRF2, NRF1, P53), and 1 liver response following activation of the innate immune response. Comparing paradigm chemical inducers of each receptor to the other compounds surveyed enabled us to identify sets of optimized gene expression panels and associated scoring algorithms proposed as quantitative mechanistic biomarkers with high sensitivity, specificity, and quantitative accuracy. These findings were further qualified using public datasets, Open TG-GATEs and DrugMatrix, and internal development compounds. With broader collaboration and additional qualification, the quantitative toxicogenomic framework described here could inform candidate selection prior to committing to drug development, as well as complement and provide a deeper understanding of the conventional toxicology study endpoints used later in drug development.

Key words: receptor; gene expression/regulation; toxicogenomics; methods; transcription factors; liver; systems; toxicology; biomarkers; safety evaluation.

A large fraction of pharmaceutical candidates fail during development due to safety liabilities uncovered during regulated preclinical testing or during clinical trials (Stevens and Baker, 2009; Waring *et al.*, 2015). Thus, earlier and more accurate

predictors of risks destined to surface later, either in preclinical or even clinical studies, are important for more efficient, successful, and safer pharmaceutical development. A particularly promising and largely untapped opportunity for the development of predictive risk assessment assays lays just prior to the regulated phases of drug discovery, where test animals are first exposed to repeated high doses of study compound to first assess tolerability and appropriateness for selection to enter development. These are typically the first studies where sufficiently high pharmacologic and metabolic burdens are achieved in living animals that could unmask signals of molecular initiating events that could lead to preclinical or clinical toxicities. Although well-recognized toxicological phenotypes (eg, tumors) may take months, or years to develop, there is opportunity to detect evidence for molecular initiating or protective defensive events during short-term repeated dose studies that are predictive of high downstream potential for preclinical or clinical liabilities.

Indeed, the liver undergoes dramatic molecular adaptation to neutralize and respond to xenobiotic stressors, providing potential sensitive signals for human drug-induced liver injury (DILI), and rodent carcinogenicity biomarker development. Much of this adaptation is orchestrated via transcriptional plasticity, making gene expression an attractive platform for biomarker development (Igarashi *et al.*, 2015; Qin *et al.*, 2016; Sutherland *et al.*, 2018) especially as broad transcriptional profiling technologies allow assessment of numerous biological processes in a single assay. Importantly, drug-induced gene expression changes in liver serve not only as biomarkers of liver injury, but also as sentinels of adaptation in absorption, distribution, metabolism, elimination, and detoxification of chemicals (ADMET), which have the potential to inform toxicologic and carcinogenic predisposition and mechanisms. Determining how the combined fingerprint of these diverse biological responses relates to the presence or absence of long-term toxicological outcomes in both the test species and humans would position these biomarkers for more routine and practical risk assessment utility.

To this end, there have been substantial efforts to characterize the breadth of drug-induced transcriptional responses in rodent liver, most notably the DrugMatrix and Open TG-GATEs efforts where genome-wide transcriptional data were collected on hundreds of compounds at multiple doses and durations (Afshari *et al.*, 2011; Ganter *et al.*, 2005; Igarashi *et al.*, 2015; Te *et al.*, 2016). These and other more focused efforts have resulted in multiple reports demonstrating the promise of transcriptional biomarkers for more accurate and earlier prediction of liver liabilities in the test species (Qin *et al.*, 2016). For example, toxicogenomics has been useful for carcinogenicity prediction (Thomas and Waters, 2016). A recent analysis (Sutherland *et al.*, 2018) applied gene coexpression approaches to the Open TG-GATEs dataset proposing gene networks that may change in rat liver following administration of diverse drugs. The utility of this coexpression framework was proposed for the hypothesized linkage of the gene expression changes to a spectrum of toxicities in chronic rat studies, therefore providing mechanistic insights into later observed toxicities.

There are 2 major hurdles for transitioning transcriptional biomarkers from investigational and hypothesis generating tools to decision-making assays that can be relied upon to inform the propensity for nonclinical and clinical risk. First, analytical performance and quantitative thresholds of significant biomarker responses must be established in the test species under dosing paradigms that match those executed during routine preclinical testing. Second, those biomarker responses and assigned thresholds must be capable of discriminating benign from toxicological outcome at specified target doses either in the preclinical species or upon relevant exposures in clinical trials. In this study, we focused on optimizing analytical performance of multiple biomarkers of liver xenobiotic response with nonclinical or clinical implications. The application of these biomarkers for predicting thresholds of toxicologic risk as a function of dose and exposure relationship, both in the test animal species and at targeted therapeutic exposures in humans is addressed elsewhere (Monroe *et al.*, 2020; Qin *et al.*, 2019).

The primary response to xenobiotic stimulus in liver is mediated by a number of receptors including AHR, CAR, PXR, PPARα, ER (Xu *et al.*, 2005). The secondary effects include response to reactive metabolites and proteasomal stress through NRF1, NRF2, and DNA damage through P53, as well as innate immune response (IIR) (Bugno *et al.*, 2015; Fischer, 2017; Xu *et al.*, 2005). Traditionally, the discovery of gene expression biomarkers for each of these mechanisms began from postulating chemical agents that specifically induce only one of these mechanisms but not the others. The most commonly utilized inducers include TCDD and PCB126 for AHR; phenobarbital for CAR; PCN for PXR; fibrates for PPARα; estrogens for ER; bardoxolone for NRF2; and lipopolysaccharide endotoxins for innate immune-mediated systemic inflammatory response. This approach was indeed very productive in identifying key signature genes for each mechanism of action. Nonetheless, the specificity of gene assignment to each mechanism is limited by chemical specificity of ligands to those receptors, and gene induction through alternative mechanisms, both direct and indirect, could not be ruled out.

In this work, rather than focusing on a limited number of compounds, the gene induction was examined in the context of a large and diverse set of compounds. We sought transcripts with exceptionally strong coexpression, and, when available, evidence of direct binding by the respective regulatory transcription factor to distinguish them from all other direct and indirect regulatory mechanisms. In this manner, we addressed the main drawback of prior studies to define transcriptional biomarkers with optimized specificity for each regulatory factor along with statistical characteristics of significant and specific induction. The resulting signatures set the stage for then defining thresholds of biological and toxicological impact in the context of the magnitude and duration of exposure to harness their utility for informing decision making for preclinical and clinical drug development studies (Monroe *et al.*, 2020; Qin *et al.*, 2019).

MATERIALS AND METHODS

Animals and treatments. Male Han-Wistar [Crl:WI(Han)] and Sprague Dawley [Crl:CD(SD)] rats (Charles River Laboratories, Raleigh, North Carolina) were socially housed and fed PMI Certified Rodent Diet No. 5002 with tap water provided *ad libitum*. Animals were 8 weeks old and 200–275 g at the start of study. Rats were dosed once daily with compound or vehicle control as indicated (Supplementary Table 2). The final necropsy was performed 24 h (±40 min) post final dose. Exsanguination of animals via vena cava was performed under isoflurane-induced anesthesia. Liver slices were flash frozen on dry ice and stored at −80°C until used. Specific study dosing regimens are provided in Supplementary Table 2. All studies were conducted under IACUC approved protocols in AAALAC International accredited test facilities.

Administered pharmaceuticals were specified in the text and figures by either generic names or public development codes. Some chemical names were abbreviated: *N*-nitrosodimethylamine (NDMA), 2,3,7,8-Tetrachlorodibenzo-*p*-dioxin (TCDD), pregnenolone 16α-carbonitrile (PCN), 3-methylcholanthrene (3-MC), 1-(3-chlorobiphenyl-2-ylmethyl)-1*H*-imidazole (Cmpd Z). The rationale for the compound selection and their sourcing is given elsewhere (Monroe *et al.*, 2020).

RNA extraction and sequencing. Total RNA was purified from frozen liver samples using the MagMAX Express Magnetic Particle Processor (Applied Biosystems, No. 4400074), and sequenced at the Covance Genomics Laboratory according to Illumina specifications. Briefly, total RNA samples were converted into cDNA libraries using the TruSeq Stranded mRNA HT Sample Preparation Kit (Illumina, No. RS-122-2103). Starting with 150 ng of total RNA, cytoplasmic ribosomal RNA was removed by hybridization to a biotinylated probe selective for ribosomal RNA species, followed by streptavidin bead binding and sample purification. The resulting cytoplasmic rRNA-depleted sample was chemically fragmented and converted into single-stranded cDNA using reverse transcriptase and random hexamer primers, with the addition of Actinomycin D to suppress DNA-dependent synthesis of the second strand. Double-stranded cDNA was created by removing the RNA template and synthesizing the second strand in the presence of dUTP in place of dTTP. A single A base was added to the 3′ end to facilitate ligation of sequencing adapters, which contain a single T base overhang. Adapter-ligated cDNA was amplified by polymerase chain reaction to increase the amount of sequence-ready library. During this amplification the polymerase stalls when it encounters a U base, rendering the second strand a poor template. Accordingly, amplified material used the first strand as a template, thereby preserving the strand information. Final cDNA libraries were analyzed for size distribution using the Agilent Technologies Bioanalyzer 1000 and Agilent DNA 1000 kit (Agilent, No. 5067-1504), quantitated by qPCR (KAPA SYBR Fast ABI Prism qPCR reagent kit, KAPA Biosystems, No. KK4835), then normalized to 2 nM in preparation for sequencing. The library products were ready for sequencing analysis via Illumina HiSeq 2000. A minimum of 40 million 50-bp paired-end reads were generated for each sample. The resulting Rx-TGx sequencing data have been deposited in NCBI's Gene Expression Omnibus (Barrett *et al.*, 2012) and are accessible through GEO Series accession number GSE144219.

SEQC toxicogenomics data. In addition, we downloaded and merged the RNA-Seq dataset from SEQC Toxicogenomic Study (GSE55347) (Gong *et al.*, 2014). The 116 RNA samples for this dataset were selected from DrugMatrix and include treatments with 27 compounds with 7 mechanisms of action (AHR, CAR/PXR, PPAR, DNA Damage, Cytotoxicity, ER, and HMGCOA). These data complement the internal data, which did not include many strong PPAR, ER, and P53 inducers. The raw downloaded data were processed and merged with the internal data using the same mapping and quantification algorithms.

Data processing and analyses. Genome alignment and gene quantitation was performed using OmicSoft Array Studio. Cleaned reads were aligned to the Rnor_6.0 genome reference using the OmicSoft Aligner with a maximum of 2 allowed mismatches. Gene level counts were determined by the RNA-Seq by expectation maximization algorithm as implemented in OmicSoft Array Studio and using Ensembl.R94 gene models.

FPKM normalization was applied to RNA-Seq counts followed by log-transformation with smoothing, ie, $\log_{10}(FPKM + 0.01)$ referred to as log*FPKM* hereafter. The relative log-ratios were calculated as mean differences between the treated and the average of the concurrent vehicle groups, each of which generally contained 4 animals. The group average log-ratios for each treatment versus corresponding study vehicles were used in the analysis. Heatmap genes were arranged using agglomerative hierarchical clustering using cosine distance.

We modeled gene expression changes, *G*, using a 10-factor linear regression model. All factors C_i (AHR, CAR, PXR, PPARα, ER, NRF1, NRF2, P53, and IIR) were calculated using an equal-weighted average of seed genes, with exception of the AHR factor (see Results section):

$$G = \sum \alpha_i C_i + \text{const} + \varepsilon,$$

where the loading coefficients, α_i were fitted for all well-expressed genes genome-wide. The NRF1 factor was adjusted 2-fold to achieve a comparable range of loading coefficients. Three criteria were used to select genes for inclusion in the respective signatures: (1) the large variance explained R^2 (or R-squared) (usually at least $R^2 > 0.7$) as well as (2) the dominant (largest) loading coefficients α_i among all loadings, and (3) $\alpha_i > 0.4$. The linear modeling was performed using MATLAB R2010b (Mathworks, Inc).

A signature score was evaluated as an average log-ratio of all selected component genes with equal weights, unless noted otherwise. A z score was also calculated to compare the signature score to the average variance of its component genes in all vehicle samples. Our criteria provided that the component genes were highly correlated and had similar induction range. Therefore, the z score calculated using an unbiased subset of 10 or more component genes proved to be most stable where possible (Supplementary Figure 1).

For cross-reference with ChIP-Seq data for AHR, PPARα, and NRF2 (Tamburino *et al.*, 2020), we required the regulatory potential score > 0.6 (Tang *et al.*, 2011), which was tighter than the default vicinity criteria but provided optimal specificity and sensitivity in the pool of our signatures. Statistical significance of gene set overlaps was evaluated using MSigDB 7.0 (Liberzon *et al.*, 2011).

RESULTS

Pathway Overview

The extensive RNA-Seq dataset that we assembled for this work included liver samples from 850 treated and control samples covering 120 different compounds. The Rx-TGx studies conducted internally were focused on compounds with a strong history of the presence or absence of hepatotoxicity. The data from the SEQC toxicogenomics dataset complemented the internal studies by adding more PPARα and ER inducers, as well as known genotoxic and cytotoxic compounds. The overall dataset covered vast chemical, pharmacological, and toxicological space.

We first assessed the expression changes in key ADMET genes and their relationships with canonical chemical inducers of xenobiotic nuclear receptors (aryl hydrocarbon receptor, AHR; constitutive androstane receptor, CAR; pregnane X receptor, PXR; peroxisome proliferator-activated receptor α, PPARα; estrogen receptor α, ER). We also assessed mediators of reactive metabolite-mediated stress responses (nuclear factor erythroid-

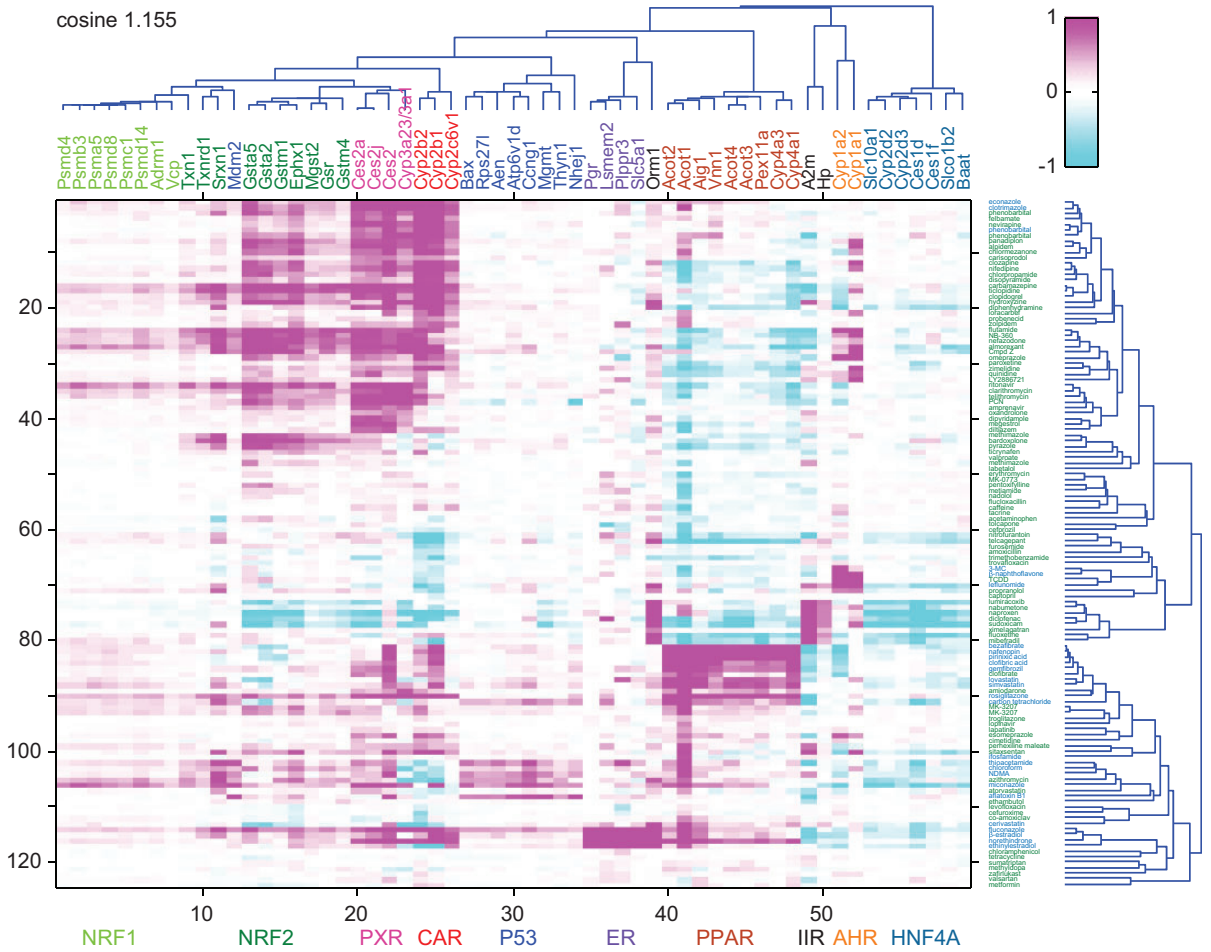

Figure 1. Regulation of ADMET gene expression. Select canonical response genes for AHR, CAR, PXR, PPARα, NRF2, NRF1, P53, and IIR are clustered using the agglomerative hierarchical method with cosine similarity as the distance measure. The height of the dendrogram branches indicates the distance metric between genes. The shown induction magnitude is calculated as log*FPKM* difference of gene expression between treated and the group median of corresponding study vehicle samples, averaged over 4 animals in each group. Only in this figure, low Cyp1a1 log*FPKM* values were raised to 0.7 before the log-ratio was calculated. The genes are colored according to purported mechanism. Treatment along y-axis was profiled internally and borrowed from SEQC toxicogenomics dataset as shown in green and blue colors, respectively.

2-like 1 and 2, NRF1 and NRF2; tumor suppressor P53), and transcriptional changes triggered following toll-like receptor pathway activation in liver reflecting a liver response to an acute innate immune system activation (IIR). The genes selected for these initial analyses were based on the most prominent literature consensus for each TF and were not intended to be a comprehensive set of xenobiotic response genes (Gotoh *et al.*, 2015; Nakata *et al.*, 2006; Rooney *et al.*, 2018; Xu *et al.*, 2005). Using agglomerative hierarchical clustering with cosine similarity as a distance measure across the entire 120-compound dataset, we confirmed that genes assigned to the same mechanisms in the literature had the highest gene-to-gene correlations and mostly clustered together (Figure 1), Indeed, the antioxidant response genes (glutathione metabolism genes, thioredoxin Txn1, sulfiredoxin Srxn1, etc.), P53 target genes (Aen, Bax, Mdm2, Ccng1, etc.), PXR/CAR-induced genes (Cyp3a23/3a1, Ces2c, Ces2a, Ces2j), CAR genes (Cyp2b1, Cyp2b2, and Cyp2c6v1), AHR genes (Cyp1a1, Cyp1a2), PPARα-induced genes (Acot1, Acot2, Aig1, Cyp4a1, etc.), ER genes (Pgr, Lsmem2, Plppr3, Slc5a1), and IIR genes (A2m, Orm1, etc.) all formed separate dendrogram branches.

This tight clustering by mechanism encouraged us to perform multivariate linear modeling to establish the more comprehensive list of specific rat liver signature genes and their relationships based on multisignature loading coefficients rather than binary correlations. In the following sections, genes were associated with and assigned to specific transcription factors based on dominant loading coefficients in the linear model. This modeling was performed genome-wide using the RNA-Seq data aligned to the current version of the rat genome to build the most *comprehensive* sets of highly transcription factor-specific biomarker genes in male rat liver.

AHR Signature

Cyp1a1, Cyp1a2, Cyp1b1, Ahrr, and Aldh3a1 are frequently referenced as AHR-induced genes (Manikandan and Nagini, 2018; Watson *et al.*, 2014). Using ChIP-Seq, we confirmed direct AHR binding at each of these genes except for Aldh3a1 in rat liver following TCDD treatments (Tamburino *et al.*, 2020). However, we found that only Cyp1a2 is well expressed in vehicle-treated rat liver samples with an average log*FPKM* value of 2.77. Cyp1b1, Aldh3a1, Ahrr had log*FPKM* values less than −1, which we considered unreliable for baseline biomarker quantitation and therefore removed them from further consideration. Although Cyp1a1 was low-expressed, it was detectable above background with an average log*FPKM* = −0.3, and a standard deviation of 0.5

in vehicle samples. Using qPCR data from a larger set of internal development compounds (Qin *et al.*, 2019), we examined the relationship between the expression levels of Cyp1a1 and Cyp1a2 and determined an orthogonal regression slope of 0.5 when Cyp1a1 log-ratio > 1 (Supplementary Figure 2), ie, Cyp1a1 was induced twice as much as Cyp1a2 at the higher end of their induction ranges. Therefore, to equalize contributions of both genes, we used the weighted average of $\frac{1}{3}$Cyp1a1 + $\frac{2}{3}$Cyp1a2 for scoring the AHR factor.

We next performed genome-wide linear modeling against this AHR factor and, concurrently, the other factors (CAR, PPAR, etc.; see Materials and Methods section), and did not find any genes in addition to Cyp1a1 and Cyp1a2 with $R^2 > 0.7$ or with an AHR loading coefficient > 0.4. Figure 2A shows that the AHR loadings for Cyp1a1 and Cyp1a2 were much higher than any other loadings. Ahrr and Aldh3a1 had the next highest AHR loadings but their R^2 was below 0.4. Of note, the Ugt1a family of genes, which have been proposed previously as AHR-responsive genes (Nakata *et al.*, 2006), did not exhibit large dominant AHR loading coefficients making them unacceptable as specific AHR biomarker genes in male rat liver. We therefore concluded with reasonable certainty that Cyp1a1 and Cyp1a2 were the most sensitive, specific, and reliably measurable components of the AHR response factor in rat liver.

The standard deviation of the AHR score in the baseline vehicle samples was calculated to be 0.28. The AHR score induction exceeded 2 standard deviations in 23 out of 120 treatments in the training set (19%). Figure 3A ranks compounds with top AHR score inductions. Flutamide, leflunomide, and TCDD, well-known potent AHR inducers, have AHR scores of approximately 2.0 ($z = 7$). We observed many AHR inducers with AHR scores barely exceeding 0.5. Although such induction tended to be statistically significant ($z = 2$, $p < .05$), the biological significance of these small and often transient inductions has been reasonably questioned (Hu *et al.*, 2007). We elsewhere address further the relation of biological thresholds including the importance of magnitude and duration of AHR induction to carcinogenic outcomes (Qin *et al.*, 2019).

CAR and PXR Signatures

Several subfamilies of cytochrome P450 enzymes, Cyp2b, Cyp2c, and Cyp3a, are believed to be induced by CAR and PXR receptors interchangeably (Manikandan and Nagini, 2018) and some compounds serve as dual CAR and PXR ligands (Moore *et al.*, 2000) possibly with different affinities. There is considerable debate in the literature on the possible interaction between these receptors (Tolson and Wang, 2010), perhaps owning to their close homology (Evans and Mangelsdorf, 2014). Our preliminary clustering (Figure 1) showed that these enzymes and certain carboxylesterase 2's form 2 distinct branches on the dendrogram, indicating that it might be possible to define 2 separate signatures and gauge relative affinity of compounds toward CAR or PXR using these signatures.

The 2 separate factors that we used in our multiple regression analysis were assigned as follows: Cyp2b1, Cyp2b2, Cyp2c6v1 were used as a CAR factor, Cyp3a23/3a1, Ces2c, Ces2a, Ces2j were used as a PXR+CAR factor. The plus symbol signifies that both transcription factors can contribute to the induction of these genes (see below). The mentioned P450 genes had very high baseline expression levels, log*FPKM* > 2. The carboxylesterases were slightly lower but still had log*FPKM* > 0. Thus, all these genes are expressed at baseline levels in rat liver that are easily detectable with conventional methods.

We used simple averages of the component genes as factors in the multiple regression analysis for these receptors. Figures 2B and 2C show that indeed these component genes obtain uniquely high loading coefficients for their respective factors. In addition, 2 UDP-glucuronosyltransferases Ugt2b10 and Ugt2a3 met our criteria of $R^2 > 0.7$ for inclusion in CAR and PXR+CAR signatures, respectively. Indeed, Ugt2a3 is reported as PXR-inducible in the literature (Court *et al.*, 2008). No other genes in this genome-wide analysis had CAR or PXR+CAR loading coefficients above 0.4 with $R^2 > 0.7$. Much like the AHR signature, the CAR and PXR signatures contain a very small number of genes with large loading coefficients (see Supplementary Table 1). This suggests that these receptors have a unique role and only induce transcription of the aforementioned drug metabolizing enzymes. They also do not directly impact the stress response genes covered by other transcription factors (see below).

We further examined the degree of independence of these signatures. Figures 4A and 4B compare the AHR, CAR, and PXR+CAR scores calculated as the average induction of the component genes. TCDD is an example of a compound that induces only the AHR score but not the PXR+CAR score, whereas PCN induces the PXR+CAR score but not the AHR score. Similarly, phenobarbital has high CAR score and no AHR score (Supplementary Table 2). Therefore, we could conclude that AHR and PXR+CAR inductions are independent and the responsible transcription factors do not cross-target the signature genes. Simultaneous inductions are, of course, possible when a ligand has affinity to multiple xenobiotic receptors as in the case of flutamide. On the contrary, when we compared CAR and PXR+CAR scores, we could not find compounds that induce the CAR score exclusively. Phenobarbital, a purported CAR ligand, induces both CAR and PXR+CAR scores. However, PCN, a purported PXR ligand, strongly induces only the PXR+CAR score with minimal effect on the CAR score. Due to this asymmetry and the observed dependence between these scores, they were named to indicate which transcription factors contribute to their induction. That is, the CAR score is impacted by contributions from CAR transcription factor activation, whereas the CAR+PXR score represents the additive contributions from both receptors. The relationship in Figure 4B suggests it is possible to gain a rough estimate of the individual PXR tone of a compound using the formula *PXR score* = *PXR+CAR score* − ½*CAR score*. For example, telithromycin had much higher PXR+CAR score than CAR score and, therefore, is mainly a PXR ligand similar to PCN.

The standard deviation of CAR and PXR+CAR scores in the vehicle samples was found to be 0.20 and 0.17, respectively. Statistically significant ($p < .05$) induction, with these scores above 2 standard deviations, was observed in 45 (38%) treatments for the CAR signature and 55 (46%) treatments for the PXR+CAR signature. The top-ranked compounds for both signatures are shown in Figures 3B and 3C. Phenobarbital was the top-ranked CAR compound with a CAR score of 2.2 ($z = 10$). PCN was the top ranked PXR+CAR compound with a score of 1.5 ($z = 8$) closely followed by telithromycin and ritonavir.

PPARα and SREBP2 Signatures

The PPARα receptor plays a critical role in fatty acid metabolism and induces profound gene expression changes following activation by certain endogenous fatty acids as well as some xenobiotics (Kersten, 2014; Rakhshandehroo *et al.*, 2010). Genes induced by PPARα include the Cyp4a subfamily, Acot family, Colq peptide among others. PPAR targets tend to be well expressed in vehicle samples with log*FPKM* generally > 1, likely

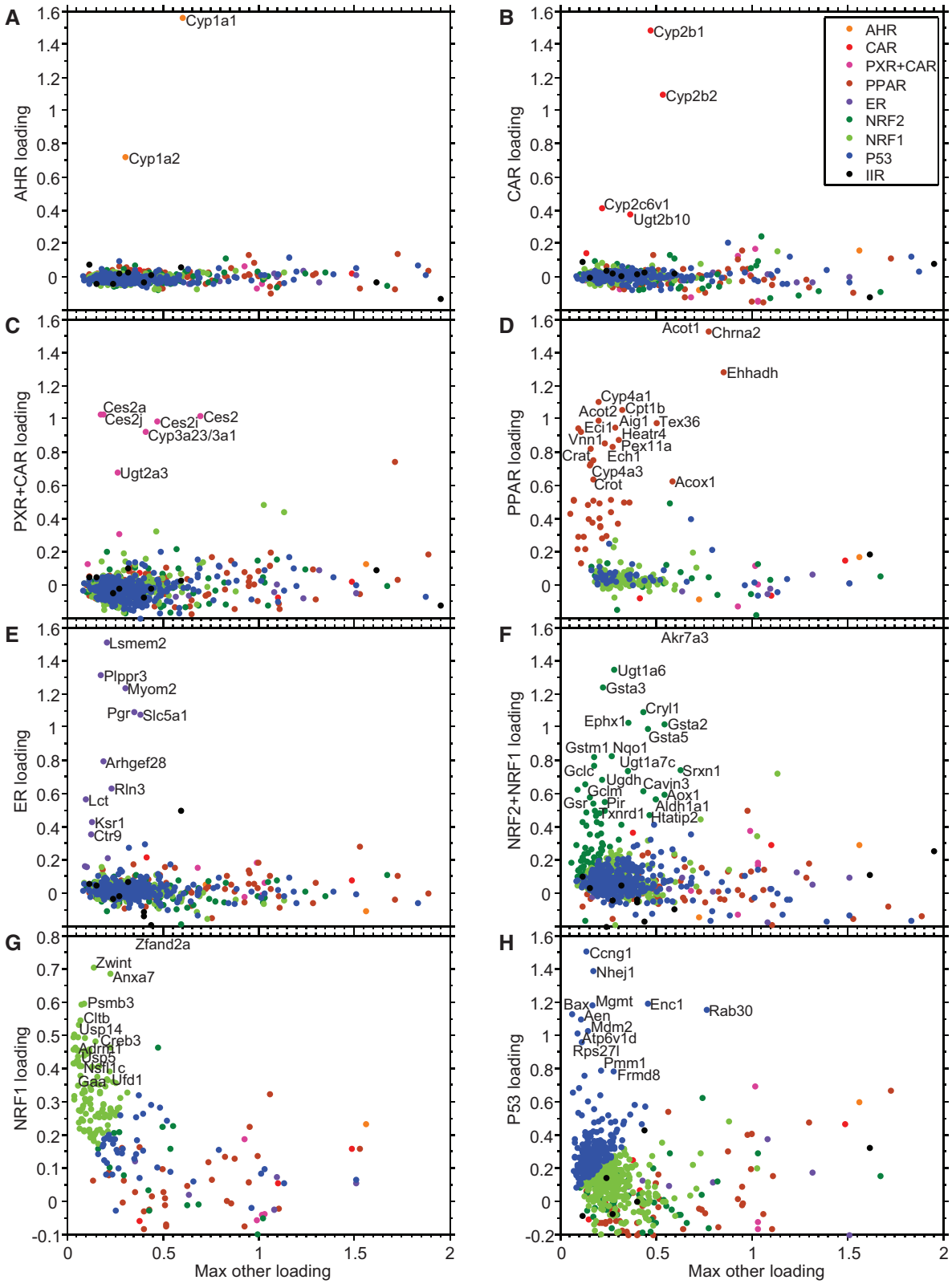

Figure 2. Loading coefficients from linear modeling. In panels A though H, the loading coefficients for AHR, CAR, PXR+CAR, PPAR, ER, NRF2+NRF1, NRF1, P53 are shown against another loading coefficient with maximum absolute value. The AHR, CAR, PXR+CAR, NRF2+NRF1, and P53 panels show genes with $R^2 > 0.7$, The PPAR and NRF1 panels use tighter $R^2 > 0.8$. Genes with highest loading coefficients are also labeled with their gene symbols. Colors indicate the factor with the highest loading for each gene.

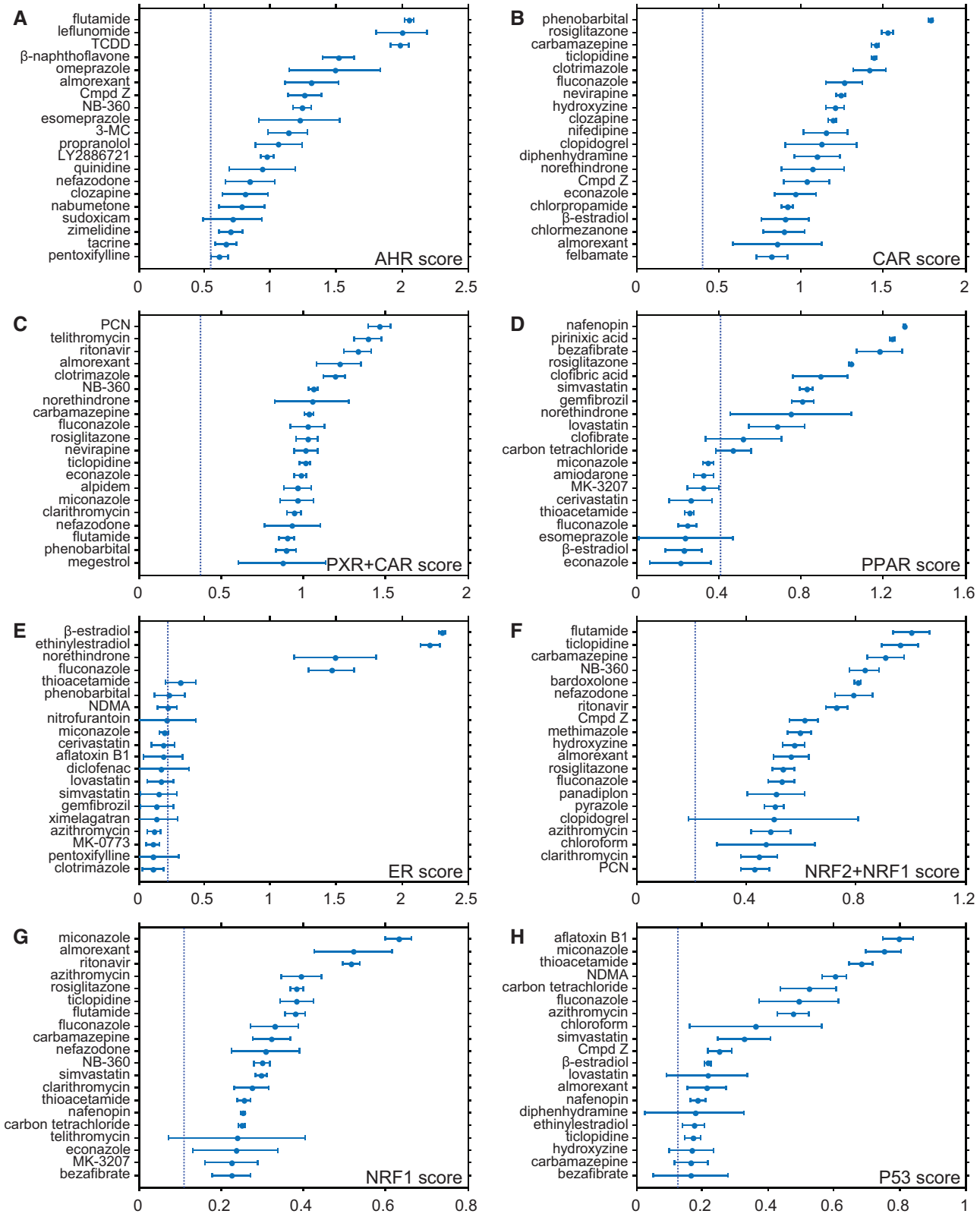

Figure 3. Signature scoring for top 20 compounds. Each score is calculated as an induction average of all genes selected for a given signature with equal weights, expect AHR was calculated as $\frac{1}{3}$Cyp1a1 + $\frac{2}{3}$Cyp1a2. The error bars show ± standard deviation among the studied animals. The dashed line represents 2 standard deviations of the average baseline expression of the signature genes in all available vehicle samples. Each panel shows only the 20 top-scoring compounds.

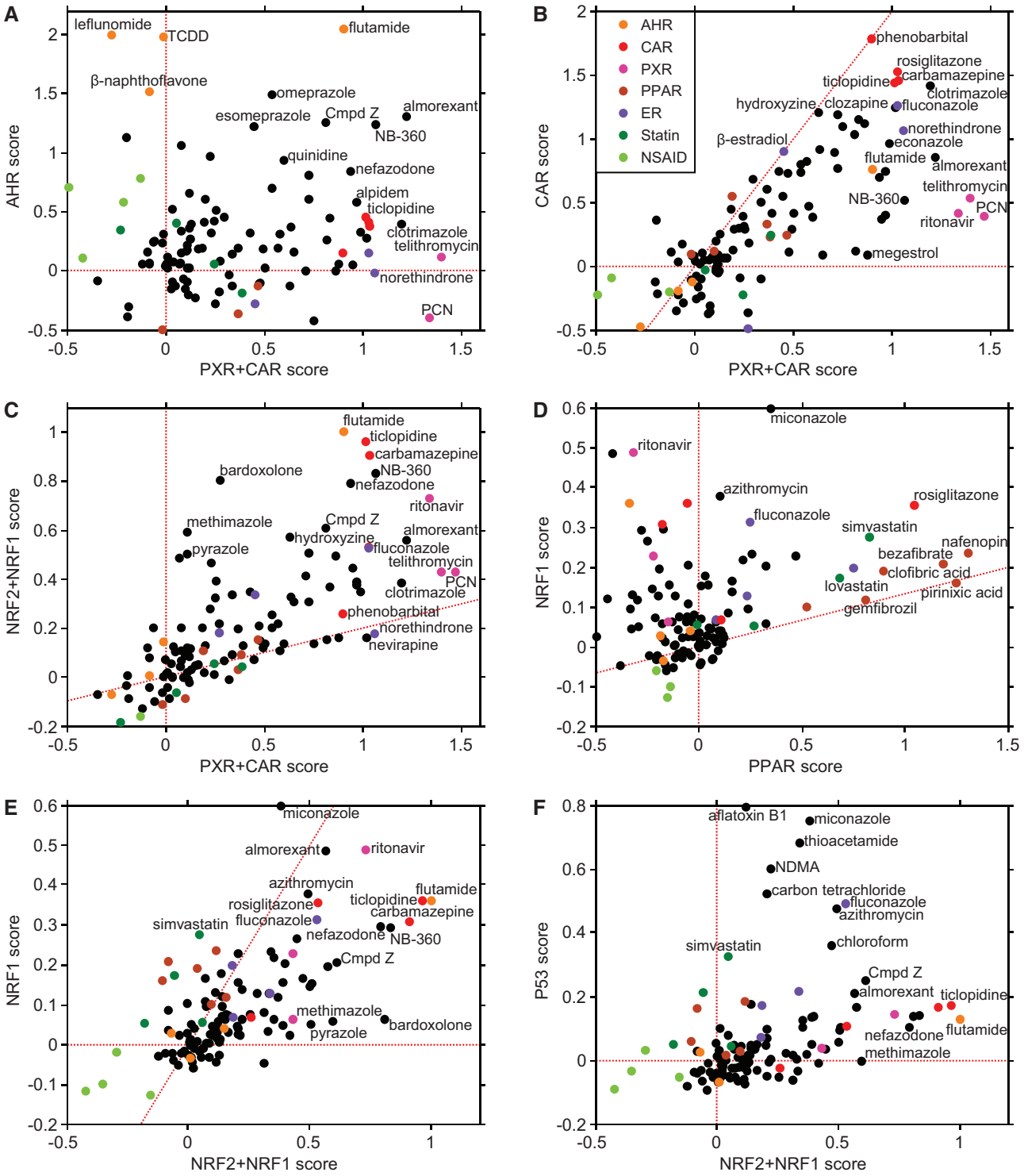

Figure 4. Relationships between the signature scores. The scatter plots show relationship of the PXR+CAR and AHR (A), PXR+CAR and CAR (B), NRF2+NRF1 and PXR+CAR (C), NRF1 and PPAR (D), NRF1 and NRF2+NRF1 (E), P53 and NRF2+NRF1 (F) signature scores for individual compounds. Some compounds are colored according to their high induction rank. The highest scoring compounds are also labeled. The slant of the red dotted lines signifies the envelope range of observed values, with canonical inducers aligned along the lines.

representing normal tone from endogenous fatty acid ligands, and are thus easily detectable with conventional methods. Our preliminary clustering (Figure 1) showed that they form a distinct well-defined branch on the dendrogram apart from other ADMET genes.

The seed for the PPARα factor in our multiple regression analysis included the following genes: Cyp4a1, Cyp4a2, Cyp4a3, Acot1, Acot2, Acot3, Acot4, Crat; and the average expression of these genes was used in the analysis. Genome-wide modeling identified 22 additional genes that satisfied the standard criteria

of PPAR loading coefficient > 0.4 and R^2 > 0.7, for example, Chrna1, Aig1, Vnn1, Ehhadh, Ech1, and Colq (Figure 2D). The large size of this signature is a testament to the broad role PPARα plays in fatty acid oxidation and CoA recycling. Notably, ChIP-Seq revealed direct PPARα-binding proximal to all 29 of these PPARα signature genes (Tamburino *et al.*, 2020). The full set of identified PPARα signature genes is listed in Supplementary Table 1.

The standard deviation of PPARα scores in the vehicle samples was found to be 0.17. The PPARα score in the treated samples exceeded a statistically significant ($p < .05$) threshold of 2 standard deviations in only 19 (16%) of treatments tested in the set of 120 compounds. Figure 3D plots the top ranked PPARα inducers and is enriched with fibrates and statins. The top-ranked compound was nafenopin with PPARα induction score of 1.2 ($z = 7$). Frequently, in 18 (14%) treatments, the PPARα score also exhibited statistically significant downregulation.

Besides PPARα activation by statins, in treatments of this drug class we observed strong upregulation of numerous genes in the mevalonate/cholesterol pathway which is known to be under tight feedback control by SREBP2 in response to cholesterol changes (DeBose-Boyd, 2008; Sato, 2010; Ye and DeBose-Boyd, 2011). To model this effect, which is not considered toxicologic, we seeded an additional SREBP2 factor with a few canonical cholesterol synthesis genes Cyp51, Msmo1, Sqle, Mvk, Idi1, and Lss (Mazein *et al.*, 2013). In the genome-wide modeling, we found 30 genes with SREBP2 loading > 0.4 and R^2 > 0.75, listed in Supplementary Table 1. We confirmed that all but 6 genes in this signature belong to either SREBP targets or cholesterol metabolism genes from multiple sources collected in MSigDB. The SREBP2 score exceeded a threshold of 2 standard deviations, which was 0.16 in the vehicle samples; in 16 (13%) compounds and only statins had the SREBP2 score above 0.78 ($z > 4.5$). As a part of this signature, upregulation of Acly and Acss2, which produce acetyl-CoA from citrate and acetate, reached more than 10-fold in atorvastatin treatment.

ER Signature

Estrogen effects are rarely studied in nonreproductive organs. A few ER-induced genes in male rat liver were recently identified in an analysis of the Open TG-GATEs dataset (Rooney *et al.*, 2018), where Ctr9, Rgs3, and Lifr were identified among the most responsive genes. We confirmed the specificity of these genes and identified additional liver-responsive genes using data from estrogen-treated male rats from the independent SEQC toxicogenomics dataset integrated into our analysis. The specific induction of Pgr, Lifr, and Citr9 by hormonal treatments is corroborated by literature evidence of the association of these genes with ER signaling (Ahlbory-Dieker *et al.*, 2009; Zeng and Xu, 2015). We used the average of these gene inductions as a seed ER factor, and multiple regression analysis identified some additional genes that satisfied our criteria of ER loading coefficient > 0.4 and R^2 > 0.7 (Figure 2E). Notably, we identified induced expression of progesterone receptor Pgr in the male liver samples. A full list of genes with high loading coefficients is included in Supplementary Table 1.

Many of the ER-induced genes had very low baseline expression with log*FPKM* < −1 and, therefore, very large fold changes. All estrogens in our dataset, β-estradiol, ethinylestradiol, and norethindrone, received a very high ER score (Figure 3E). Fluconazole also obtained a high ER score perhaps reflecting its proposed indirect effect on altering steroid metabolism and subsequent systemic exposures of endogenous estrogen and testosterone (Zarn *et al.*, 2003).

NRF2 and NRF1 Signatures

Oxidative stress responses in liver are managed by the Keap1-NRF2 system sensing of electrophilic metabolites (Kobayashi and Yamamoto, 2005) as well as the recently recognized proteasome-NRF1 system that controls degradation of damaged proteins (Bugno *et al.*, 2015; Sha and Goldberg, 2014). Figure 1 shows 2 closely correlated dendrogram branches containing NRF2-associcated glutathione *S*-transferases, thioredoxin, sulfiredoxin, and other detoxification genes on one branch, and several NRF1-associated proteasomal subunits, Vcp, and Adrm1 on the other. The correlation between the genes in both branches is very strong. Yet, the induction of antioxidant genes is in general much stronger than that of the proteasomal genes. We modeled these pathways separately and introduced 2 factors, NRF2+NRF1 and NRF1 (see below).

The NRF2+NRF1 seed antioxidant factor was initially calculated as an average of Ephx1, Gsr, Gsta2, Gsta5, Mgst2, Gstm1, Gstm4, and Txnrd1. The genome-wide linear modeling identified additional genes with loading coefficients > 0.4 and R^2 > 0.7 including Ugdh, Pir, Cryl1, Aox1, Ugt1a6, Ugt1a7c, etc. (Figure 2F and Supplementary Table 1). Notably, ChIP-Seq revealed direct NRF2-binding proximal to the majority of the NRF2+NRF1 signature genes (Tamburino *et al.*, 2020). However, we chose to call this biomarker NRF2+NRF1 because in NRF2 knockout animals, whereas the NRF2 activator bardoxolone can no-longer induce these genes, certain reactive metabolite generating compounds (eg, ticlopidine) can still induce them. We thus hypothesize that either NRF2 or NRF1 can activate these genes (Monroe *et al.*, 2020), despite the lack of ChIP-grade antibodies for NRF1 to test this directly. The full set of resulting NRF2+NRF1 signature genes is listed in Supplementary Table 1. The size of this signature is much larger than those of the above nuclear receptors and covers a broad spectrum of ROS chemistry.

The baseline expression of antioxidant genes was high with log*FPKM*>1. The standard deviation of NRF2+NRF1 signature scores in vehicle samples was 0.1. The NRF2+NRF1 scores in the treated samples exceeded a statistically significant ($p < .05$) threshold of 2 standard deviations very frequently, in 65 (52%) treatments. About half of the compounds in this set have well-documented DILI liabilities and were either withdrawn or received label warnings (Monroe *et al.*, 2020). Therefore, it was not surprising that many of them might produce reactive metabolites to induce antioxidant responses as reflected in our signature score. Figure 3F shows the top-ranking treatments in our dataset with flutamide and ticlopidine at the top having NRF2+NRF1 scores of 1.2 ($z = 8$). We further determined the relationship of the NRF2+NRF1 signature score and liver reactive metabolism burden to daily dose and clinical DILI risk elsewhere (Monroe *et al.*, 2020).

The seed NRF1 factor was calculated as an average of Psmb3, Psmc1, Psma5, Psmd4, Psmd8, Psmd14, Adrm1, Vcp. Genome-wide modeling identified 56 additional genes with NRF1 loading coefficient > 0.4 and R^2 > 0.8. This included almost all proteasome subunits, Usp5, Usp14, Abhd4, and many other genes related to proteasome function. The full set is listed in Supplementary Table 1. These genes were not induced by bardoxolone (Figure 4E), and their induction by compounds capable of forming chemically reactive metabolites was retained in NRF2 knockout studies (Monroe *et al.*, 2020), justifying using NRF1 (and not NRF2) in the naming of this signature. Despite the NRF1 and the NRF2+NRF1 signatures representing distinct biological responses to metabolic stress, there was clear correlation between them (R^2 = 0.45, Figure 4C), suggesting that reactive metabolites almost inevitably induce proteasomal stress via an increase in adducted proteins.

The baseline expression of all proteasome subunits, Vcp, and Adrm1 was high, log*FPKM*> 1, but their low (about 2-fold) dynamic range make them challenging for biomarker assessment. The standard deviation of the NRF1 signature score was 0.06. The NRF1 scores in the treated samples exceeded a statistically significant ($p < .05$) threshold of 2 standard deviations frequently, in 64 (52%) treatments. Again, frequent induction of proteasome activity is not surprising within this compound set as it was enriched in drugs suspected or shown to be associated with human DILI and many shown or suspected to form protein adducts in liver. Figure 3G shows the top ranking of the treatments in our dataset with almorexant and ritonavir at the top with a NRF1 score of 1.2 ($z = 8$).

With these signatures defined, we were able to study their interactions with the nuclear receptor signatures. As shown in Figure 4C, high induction of PXR+CAR genes resulted in some consistent low level of a NRF2+NRF1 signature response. We reasoned that at low levels of PXR+CAR induction, oxygen free radical generation was not sufficient to a trigger NRF2 activation, but with strong PXR inducers, excessive generation of oxygen free radicals known to result from "leaky microsomes" (Dostalek *et al.*, 2007) might contribute significantly to NRF2+NRF1 response. An even more striking interaction could be seen between the PPAR and NRF1 signatures (Figure 4D). Strong PPAR inducing compounds seemed to inevitably induce a low level of proteasomal-NRF1 stress as presented by the NRF1 signature score. An explanation could be that the increased PPAR activity led to excessive protein transacylation and glycation by fatty acid acyl-CoA thioesters and acyl-glucuronides (Skonberg *et al.*, 2008). Interestingly however, a high PPAR score was not associated with significant NRF2+NRF1 score increases, quite likely because of Keap1-NRF2 not sensing this form of reactive chemistry.

P53 Signature

The P53 driven transcriptional DNA damage response is of considerable interest within the toxicogenomic community, and several P53-responsive gene signatures have been reported and reviewed recently (Auerbach, 2016; Fischer, 2017). The complementary SEQC toxicogenomics dataset integrated in our analysis contained several classic genotoxic and carcinogenic compounds such as aflatoxin B1, NDMA, and thioacetamide. Based on previously published signatures, the seed P53 factor included Mdm2, Ccng1, Aen, Bax, Nhej1, Apex1. In addition (Figure 2H and Supplementary Table 1), the genome-wide linear modeling identified Apt6v1d, Rps27l, and several other genes that serve as good DNA damage biomarkers and are known to be direct P53 targets (Auerbach, 2016; Fischer, 2017). The signature score based on these genes demonstrated high sensitivity in identifying known genotoxic compounds in our dataset, ranking aflatoxin B1, thioacetamide, and NDMA at the top of the list (Figure 3H).

The P53 signature genes had moderate log*FPKM*> 0.5 and low standard deviations in base line samples, making them good biomarker genes. Induction of the P53 signature above 2 standard deviations was observed in 30 (25%) treatments, but only 6 compounds reached a z score of 4, including all the known genotoxicants. As shown in Figure 4F, P53 score was largely independent of NRF2+NRF1 score. In fact, the genotoxic compounds appeared to escape NRF2+NRF1 detection with very unremarkable scores, indicating the importance of establishing a distinct P53 signature threshold for harder electrophiles that could be expected to be more relevant to carcinogenicity risk prediction in future studies relative to the softer electrophiles detected by the Keap1-NRF2 system.

IIR and HNF4α Suppression

Inflammatory responses both protect from and contribute to drug-induced damage of liver function. Liver can robustly induce the tissue expression and excretion into plasma of the rat acute phase plasma proteins, A2m and Orm1 (Tugendreich *et al.*, 2006) and indeed these proteins in plasma and their transcripts in liver tissue can serve as biomarkers of the liver response to systemic inflammatory stress. We noticed that these genes were induced by some treatments in our dataset, particularly the NSAIDs. This induction may be indirect, and for NSAIDs it could result from their known gastrointestinal toxicity in rodents and the release of endotoxin into the blood stream (Tugendreich *et al.*, 2006). Local liver cytotoxicity can also induce universal tissue injury signatures reflecting a component consisting of innate inflammation (Glaab *et al.*, 2018). In this study, we observed induction of genes related to interferon signaling and antigen processing, as well as neutrophil and lymphocyte infiltration to some compounds. Although our current dataset was not sufficient to train a comprehensive and mechanistic inflammation model covering these various factors, we included a single IIR factor in our linear model to exclude nonspecific association of immune-related genes to the other gene signatures.

This abridged IIR factor seed was calculated as an average expression of A2m and Orm1. The Orm1 gene induction value was doubled to compensate for much stronger dynamic range induction capacity in A2m. The genes that satisfied the regression criteria of inflammation loading above 0.4 and $R^2 > 0.7$ were, for example, neutrophil-expressed Lcn2 and Nos2 as well as fibrinogen subunits (Supplementary Table 1). Thus, in addition to acute phase response genes, the IIR factor in our model accounted for correlated neutrophil infiltration as evidenced by Lcn2 upregulation.

The baseline standard deviation for the IIR signature was very low 0.06 and the induction above 2 standard deviations was observed in 25 (20%) of treatments, with NSAID treatments ranking at the top with remarkable z scores above 10 (Figure 5A). Importantly, we observed that the IIR signal was associated with uncharacteristic downregulation of CAR, PXR+CAR, PPAR, and NRF2+NRF1 responses, but not AHR, ER, NRF1, or P53 (Supplementary Table 2, Figure 1). For example, the NSAIDs and some other treatments had especially large negative PXR+CAR and NRF2+NRF1 scores (Figs. 5B and 5C).

Besides these signatures, we found that several other important ADMET genes, including Cyp2d2, Cyp2d3, Slc10a1, Slco1b2, and Baat, were strongly downregulated by the NSAIDs and several other treatments. Our experimental observations match the gene expression changes in HNF4α knockout and knockdown studies (Kamiyama *et al.*, 2007; Lu *et al.*, 2010) or deletions of HNF4α-binding sites (Geier *et al.*, 2008). Using these genes as a seed for an additional factor in our linear model we defined a signature with loading > 0.4 and $R^2 > 0.7$ (Supplementary Table 1) and confirmed its significant overlap with curated HNF4α targets in MSigDB (9/29 genes, $FDR = 5 \times 10^{-12}$) (Ohguchi *et al.*, 2008). Figure 5D summarizes the strong association of negative HNF4α score with NSAID treatments without comparable upregulation in any other treatments. Therefore, we hypothesize that disruption of the HNF4α function, which maintains high basal expression of each of these genes (Gotoh *et al.*, 2015), could be responsible for the observed effects. Possible mechanisms of HNF4α regulation have been reviewed recently (Lu, 2016).

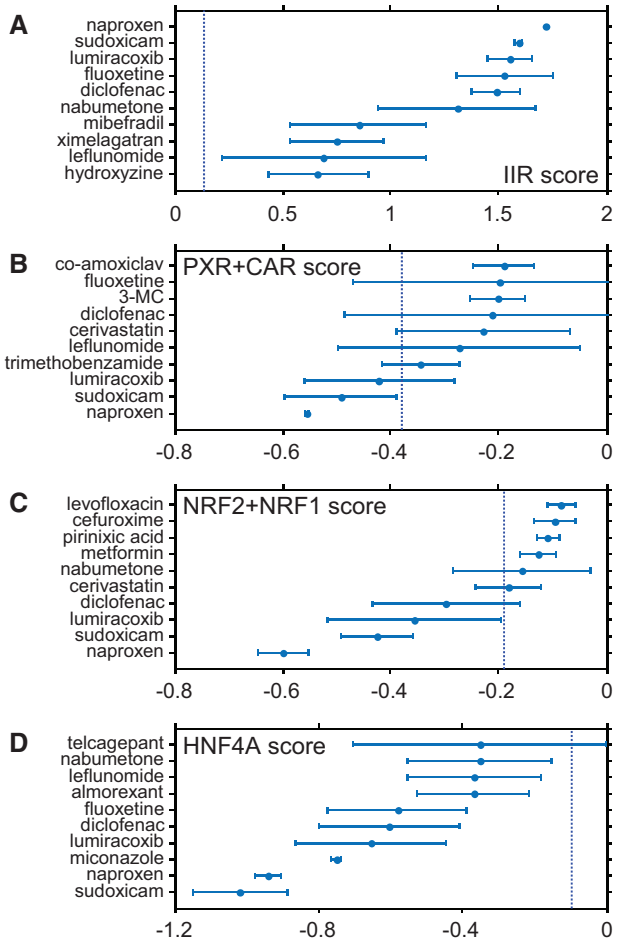

Figure 5. Suppression of signature scores by NSAID and some other treatment. Each score is calculated as an induction or suppression average of all genes selected for a given signature with equal weights. The 10 lowest negative scores in PXR+CAR, NRF2+NRF1, and HNF4α signatures are shown.

Preclinical Development Experience

The compounds tested in the internal and external datasets that were used for signature identification, as described above, were intentionally selected and enriched in chemistry with confidence for affecting or not affecting relevant liver biology and toxicology of targeted interest. To obtain a more "real world" perspective on the frequency and behavior of these signatures across a more diverse and less biased representation of typical small molecules proposed as candidates for entry to drug development, signature scores were calculated from historical rat liver transcriptional data collected among internal discovery compounds just prior to candidate selection for over 10 years between 2008 and 2018. A total of 664 different compounds were assessed in samples collected 24 h after 4 or 7 days of q.d. dosing at multiple dose levels relevant to efficacy and to toxicity assessment. As these data were collected with various low- to mid-density qPCR array platforms over this 10-year timeframe, not all genes from all signatures were measured in all studies. In these cases, scores were calculated using the subset of signature genes measured in the highest number of studies, and only those studies where the complete subset was measured were included in the analyses. Specifically, for AHR, 663 of the 664 compounds had data for both Cyp1a1 and Cyp1a2 and were considered. For CAR, 658 compounds had data for Cyp2b1 and Cyp2b2, and signature scores were calculated for these compounds using these 2 genes. Similarly, PXR+CAR scores were calculated for 661 compounds using Ces2 and Cyp3a23/3a1; PPAR scores for 648 compounds using Acox1, Cyp4a3, and Cyp4a1; NRF2+NRF1 scores for 267 compounds using Akr7a3, Entpd5, Ephx1, Gclc, Gclm, Gsr, Gsta2, Gsta3, Gsta5, Gstm1, Gstm4, Htatip2, and Txnrd1; NRF1 for 376 compounds using Blvrb, Ran, and Zwint; ER for 109 compounds using only Slc5a1; P53 for 33 compounds using Bax, Mgmt, Ccng1, Atp6v1d, Mdm2, and Nhej1; and IIR for 379 compounds using Lcn2, A2m, Orm1, Hpx, Kng1, and Lbp. The qPCR data were normalized using an extended set of resistant endogenous control genes, including Hnrnpul1, Inpp5a, Ddx47, Pum1, Srrm1, Tlk2, Gusb, Rab35, Tmem183a, Rchy1, Tmed4, and 18S, for increased measurement accuracy (see Discussion below on minimal recommended signature sizes).

A statistically positive signature response was considered at 2-fold the mean of the standard deviation across the vehicle control samples from all corresponding studies ($z > 2$). Using this statistical significance threshold as a metric, 23% (153/663) of compounds exhibited some level of an AHR signature response, 29% (193/658) a CAR response, 44% (291/661) a PXR+CAR response, 10% (68/648) a PPARα response, 57% (153/267) a NRF2+NRF1 response, 48% (181/376) a NRF1 response, 2% (2/109) an ER response, 21% (7/33) a P53 response, and 35% (133/379) an IIR in liver (Supplementary Figure 3). Thus, the prevalence over these signatures among preclinical development compounds was overall quite similar to the set of commercial compounds used for training. For example, the NRF2+NRF1 response was the most commonly observed in both sets with some response seen in the majority of compounds. The largest difference between the compound sets was a higher incidence of the IIR signature among the preclinical development compounds compared with the training set, likely as we intentionally excluded known rat liver toxicants from the training set. Conversely, ER induction was also rarely observed among the preclinical candidates. Although these metrics represent the point of departure ($z > 2$) at 1 or more tested doses/days, the biological implications and any associated liabilities would depend on the magnitude and duration of induction above the point of departure at relevant test doses as compared with the doses and tissue exposures that will be achieved in the target prediction species.

This 664-compound set was also used to confirm the relationships between the signatures (Supplementary Figure 4). As noted in the Rx-TGx data (Figures 1 and 5), some negative association was observed between the IIR signature and CAR, PXR+CAR, PPARα, and the NRF2+NRF1 signatures. In addition, there is some degree of correlation between the NRF2+NRF1 signatures with the CAR and PXR+CAR, signatures likely representing the indirect biological associations expected when drug metabolism enzymes induced by nuclear receptors may amplify the generation of reactive metabolites to trigger NRF2 and NRF1 pathways as discussed above.

Other Datasets

The biomarker signatures developed in this work were applied and tested in the other large toxicogenomic datasets: DrugMatrix and Open TG-GATEs. We confirmed that the signature component genes maintained strong mutual correlation and reliably identified strong nuclear receptor ligands and known toxicants (Supplementary Table 2). Indeed, a few coexpression modules reported recently (Sutherland *et al.*, 2018) contain 50%–80% of genes used in our signatures: specifically modules 26m ~ PPARα, 43 ~ ER, 42m ~ NRF2+NRF1, 7m ~ NRF1,

205 ~ P53, 50 ~ IIR, 46m ~ SREBP2, 6m ~ HNF4A. Importantly, this demonstrates the cross-platform robustness of our transcriptional signature biomarker strategy in the context of being able to leverage older microarray data. However, the Rat230_2 microarray platform (GPL1355) was poorly suited for monitoring CAR and PXR+CAR signatures because it lacked critical genes from these signatures.

DISCUSSION

For biomarker gene expression panel (signature) identification and optimization we leveraged RNA-Seq data from male rat livers following q.d. oral administration of 120 different compounds with or without associated DILI risk in the clinic. We studied diverse but hepatocyte-specific transcriptional responses in relatively homogeneous liver samples, which minimized confounding effects of cell composition and morphology. The signatures were trained via multifactor linear regression modeling and coexpression with canonical targets and paradigm activating compounds for the specific regulatory receptors and mediators. These signatures were refined based on specificity of response across the dataset, magnitude of response to paradigm compounds, fitness as an analytical biomarker including sufficient baseline expression and stability. In addition, for AHR, PPARα, and NRF2, signature genes demonstrated proximal and specific TF binding at their promoters (Tamburino *et al.*, 2020). Although we have derived and presented the comprehensive gene sets meeting the inclusion criteria for each signature presented here, we recognize that all genes may not be necessary for accurate measurement. Pragmatically, when deploying routine qPCR platforms, we investigated robustness of the z score calculation using randomized subsets of genes from the large signatures. We determined that vehicle standard deviation, average induction and, consequently, z score become sufficiently stable as long as 10 or more genes from a large signature are selected for scoring (Supplementary Figure 1). This is not surprising because our signature genes were selected with extremely high mutual correlation (due to $R^2 > 0.7$) and comparable induction magnitude (due to loading > 0.4). For signatures containing < 10 regulated genes, the complete signature set should be used for best accuracy.

Critically, we were able to discern the signatures of xenobiotic sensors (AHR, CAR, PXR, PPARα, and ER) from the signatures of the stress mediators (NRF1, NRF2, P53). These sensor signatures contained very small numbers of specific genes dedicated to induction of metabolic enzymes, with the exception of PPARα which regulates a larger set of genes involved in fatty acid biology. However, the stress mediators generally induce larger numbers of detoxification and damage repair genes. With the optimized set of core signature sets we did not see evidence of direct cross-targeting between these broad categories. We hypothesized that any interaction between these signatures is due to alterations of endogenous biochemical metabolites and other indirect effects. For example, we observed SREBP2-mediated upregulation of Acly and Acss2 in cholesterol-lowering statin treatments. This should presumably increase cytosolic acetyl-CoA and stimulate *de novo* fatty acid synthesis (Kain *et al.*, 2015; Pietrocola *et al.*, 2015), which would, in turn, explain activation of PPARα by statins. An opposite decrease in endogenous fatty acids would manifest itself as the frequently observed suppression of the PPARα score.

The precise mechanisms underlying the associations of the xenobiotic sensor induction signatures following AHR, CAR, PXR, PPARα, and ER activation with liver toxicity and/or carcinogenicity are unclear and remain topics of active research. These are concluded however to be triggered by downstream effects following receptor activation as models engineered with defective receptors are more resistant. For the stress mediators NRF1, NRF2, and P53, on the other hand, phenotypic outcomes are explained by upstream biochemical insult. Engineered models with defective stress mediators that fail to trigger protective mechanisms are more susceptible.

IIR signals inform on insults activating one or more of a number of confounding toxicologic mechanisms ranging from vascular injury, multiorgan inflammation, GI perforation, which exert pressure to generally suppress a broad range of liver gene expression (Nguyen *et al.*, 2015). Future experiments will be needed to better understand mechanistically how these signals suppress gene expression in liver samples. We hypothesize this might be due to disruption of HNF4α activity, although changes to the cellular makeup of liver in response to immune signals might also contribute to a general decrease in hepatocyte gene levels in the bulk tissue samples profiled.

It is crucial to discern toxicological biomarker responses that are directly mechanistically informative of triggered preclinical or clinical phenotypic outcomes from those that may be intended to be biologically benign or beneficial adaptations. Either may be specific to the preclinical study species and irrelevant to humans. It is plausible that certain molecular responses intrinsic to the compound or its metabolites will protect and result in no toxicologic phenotype in the test species, but rather are indicative of toxicologic propensity that may manifest in less adaptive species such as humans, as we posit for NRF1 and NRF2 (Monroe *et al.*, 2020). Through accurate interpretation of mechanistic biomarker signals, and the establishment of thresholds of meaningful biomarker alterations, confident conclusions can be made. Such well-benchmarked tools will help to confidently dissociate levels of higher risks from benign levels of biomarker signals. These would be particularly impactful for liver as DILI is the most common adverse event resulting in denial, withdrawal or restriction of new pharmaceuticals (Stevens and Baker, 2009; Waring *et al.*, 2015).

Furthermore, the liver has been shown to be the most frequent site of tumor formation for human pharmaceuticals in 2-year rat carcinogenicity studies (Sistare *et al.*, 2011), the longest and most resource intensive toxicology study conducted in drug development. Just 6 months of chronic dosing could be informing future tumor risk potential in rat liver or distant sites via multistep mechanisms beginning with molecular initiating events in the liver (Rooney *et al.*, 2018; Sistare *et al.*, 2011). Improved identification and understanding of the molecular mechanisms triggering such predictive histologic risk factor alterations in liver is expected to enhance prediction of rat carcinogenicity study outcomes and inform human relevance. Among the human pharmaceuticals presenting with tumors in such 2-year rat studies, the overwhelming majority receive labels indicating that the rat tumors occur through mechanisms that are of questionable human relevance, or may likely be considered human irrelevant (Alden *et al.*, 2011; Friedrich and Olejniczak, 2011).

Here, we have specifically focused on establishing and optimizing a core set of mechanistic transcriptional signature biomarkers and initiating an assessment of the performance of these signatures. The signatures we chose are associated with canonical xenobiotic nuclear receptors (AHR, CAR, PXR, PPARα, ER), mediators of a reactive metabolite-mediated stress response (NRF1, NRF2, P53), and activation of a systemic IIR.

Sustained induction of AHR is associated with human risk for carcinogenicity, immune toxicity, developmental, cardiometabolic, dermatologic, and pulmonary toxicity (Linden *et al.*, 2010). CAR, PXR, PPARα inductions are associated with carcinogenicity risk in rodents generally not considered relevant to humans (Misra *et al.*, 2013; Yoshinari, 2019). Sustained ER activation is associated with human relevant tumorigenesis and developmental toxicities (Baldissera *et al.*, 2016). NRF1 and NRF2 are associated with proteasomal and electrophilic oxidative stress that can derive from chemically reactive metabolites associated with DILI (Monroe *et al.*, 2020) , and P53 activation can be associated with a DNA damage response to hard electrophilic metabolites of procarcinogens (Auerbach, 2016; Fischer, 2017). Because repetitive exposures of cytotoxic doses of nongenotoxic agents may also activate P53 resulting in cell cycle arrest to allow for tissue repair, it may be challenging to discriminate direct DNA damage at toxic drug doses. We have assessed the dose and time relationship of the P53 DNA damage response gene signature with a gene signature that we have identified as a universal biomarker of drug-induced tissue injury (Glaab *et al.*, 2018) and have indeed noted, for example, that such evidence for P53 activation can be seen concurrent with tissue injury even with short-term treatment at doses with agents such as carbon tetrachloride, monocrotaline, and methapyrilene that are not considered genotoxic but are believed to be carcinogenic secondary to chronic tissue toxicity (unpublished data). However, significant P53 DNA damage response is more frequently independent of drug-induced tissue injury, for example, with acetamidofluorene, cisplatin, and nitrosodiethylamine.

The quantitative toxicogenomic framework described here has evolved over time and has been used in our company at increasing stages of breadth over the past 10 years for internal decision making during early preclinical development. Signatures are used to identify dose-response relationships informing on human relevance potential, including for example, certain mechanisms of carcinogenicity, liver injury, and systemic innate immune activation. With broader collaboration and additional qualification, this initial mechanistic molecular biomarker framework could serve as an aligned reference for the wider pharmaceutical industry and regulatory agencies to account for commonly observed benign and potentially detrimental adaptations in response to drug exposure. Implementing such an approach early in preclinical development would complement and enhance the value of conventional toxicology study endpoints used later in drug development and would also provide a deeper mechanistic understanding of toxicological outcomes.

SUPPLEMENTARY DATA

Supplementary data are available at *Toxicological Sciences* online.

ACKNOWLEDGMENTS

We thank Daniel J. Holder for his advice, Jennifer K. Cho and Andrew Albright for helping with management of this large project, and our colleagues for their feedback and interest. The authors acknowledge the team of toxicology and pathology department scientists responsible for conducting the rat studies from which these data were generated.

DECLARATION OF CONFLICTING INTERESTS

The authors are employees of Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc., Kenilworth, NJ, USA.

REFERENCES

Afshari, C. A., Hamadeh, H. K., and Bushel, P. R. (2011). The evolution of bioinformatics in toxicology: Advancing toxicogenomics. *Toxicol. Sci.* **120(Suppl. 1)**, S225–237.

Ahlbory-Dieker, D. L., Stride, B. D., Leder, G., Schkoldow, J., Trolenberg, S., Seidel, H., Otto, C., Sommer, A., Parker, M. G., Schutz, G., *et al.* (2009). DNA binding by estrogen receptor-alpha is essential for the transcriptional response to estrogen in the liver and the uterus. *Mol. Endocrinol.* **23**, 1544–1555.

Alden, C. L., Lynn, A., Bourdeau, A., Morton, D., Sistare, F. D., Kadambi, V. J., and Silverman, L. (2011). A critical review of the effectiveness of rodent pharmaceutical carcinogenesis testing in predicting for human risk. *Vet. Pathol.* **48**, 772–784.

Auerbach, S. S. 2016. *In vivo* signatures of genotoxic and nongenotoxic chemicals. In *Toxicogenomics in Predictive Carcinogenicity* (R. S. Thomas and M. D. Waters, Eds.), pp. 113–153. Royal Society of Chemistry, Cambridge.

Baldissera, V. D., Alves, A. F., Almeida, S., Porawski, M., and Giovenardi, M. (2016). Hepatocellular carcinoma and estrogen receptors: Polymorphisms and isoforms relations and implications. *Med. Hypotheses* **86**, 67–70.

Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., Marshall, K. A., Phillippy, K. H., Sherman, P. M., Holko, M., *et al.* (2012). NCBI GEO: Archive for functional genomics data sets—update. *Nucleic Acids Res.* **41**, D991–995.

Bugno, M., Daniel, M., Chepelev, N. L., and Willmore, W. G. (2015). Changing gears in NRF1 research, from mechanisms of regulation to its role in disease and prevention. *Biochim. Biophys. Acta* **1849**, 1260–1276.

Court, M. H., Hazarika, S., Krishnaswamy, S., Finel, M., and Williams, J. A. (2008). Novel polymorphic human UDP-glucuronosyltransferase 2A3: Cloning, functional characterization of enzyme variants, comparative tissue expression, and gene induction. *Mol. Pharmacol.* **74**, 744–754.

DeBose-Boyd, R. A. (2008). Feedback regulation of cholesterol synthesis: Sterol-accelerated ubiquitination and degradation of HMG CoA reductase. *Cell Res.* **18**, 609–621.

Dostalek, M., Brooks, J. D., Hardy, K. D., Milne, G. L., Moore, M. M., Sharma, S., Morrow, J. D., and Guengerich, F. P. (2007). *In vivo* oxidative damage in rats is associated with barbiturate response but not other cytochrome P450 inducers. *Mol. Pharmacol.* **72**, 1419–1424.

Evans, R. M., and Mangelsdorf, D. J. (2014). Nuclear receptors, RXR, and the Big Bang. *Cell* **157**, 255–266.

Fischer, M. (2017). Census and evaluation of p53 target genes. *Oncogene* **36**, 3943–3956.

Friedrich, A., and Olejniczak, K. (2011). Evaluation of carcinogenicity studies of medicinal products for human use authorised via the European centralised procedure (1995–2009). *Regul. Toxicol. Pharmacol.* **60**, 225–248.

Ganter, B., Tugendreich, S., Pearson, C. I., Ayanoglu, E., Baumhueter, S., Bostian, K. A., Brady, L., Browne, L. J., Calvin, J. T., Day, G. J., *et al.* (2005). Development of a large-scale chemogenomics database to improve drug candidate selection and to understand mechanisms of chemical toxicity and action. *J. Biotechnol.* **119**, 219–244.

Geier, A., Martin, I. V., Dietrich, C. G., Balasubramaniyan, N., Strauch, S., Suchy, F. J., Gartung, C., Trautwein, C., and

Ananthanarayanan, M. (2008). Hepatocyte nuclear factor-4alpha is a central transactivator of the mouse Ntcp gene. *Am. J. Physiol. Gastrointest. Liver Physiol.* **295**, G226–233.

Glaab, W. E., Holder, D., He, Y. D., Gerhold, D. L., Bailey, W. J., Beare, C., Erdos, Z., Lane, P., Michna, L., Muniappa, N., *et al.* (2018). Universal toxicity gene signatures for early identification of drug-induced tissue injuries in rats. *Toxicol. Suppl. Toxicol. Sci.* **162**, PS1265.

Gong, B., Wang, C., Su, Z., Hong, H., Thierry-Mieg, J., Thierry-Mieg, D., Shi, L., Auerbach, S. S., Tong, W., and Xu, J. (2014). Transcriptomic profiling of rat liver samples in a comprehensive study design by RNA-Seq. *Sci. Data* **1**, 140021.

Gotoh, S., Ohno, M., Yoshinari, K., Negishi, M., and Kawajiri, K. 2015. Nuclear receptor-mediated regulation of cytochrome P450 genes. In *Cytochrome P450* (P. R. Ortiz de Montellano, Ed.), pp. 787–812. Springer, Cham.

Hu, W., Sorrentino, C., Denison, M. S., Kolaja, K., and Fielden, M. R. (2007). Induction of Cyp1a1 is a nonspecific biomarker of aryl hydrocarbon receptor activation: Results of large scale screening of pharmaceuticals and toxicants *in vivo* and *in vitro*. *Mol. Pharmacol.* **71**, 1475–1486.

Igarashi, Y., Nakatsu, N., Yamashita, T., Ono, A., Ohno, Y., Urushidani, T., and Yamada, H. (2015). Open TG-GATEs: A large-scale toxicogenomics database. *Nucleic Acids Res.* **43**, D921–927.

Kain, V., Kapadia, B., Misra, P., and Saxena, U. (2015). Simvastatin may induce insulin resistance through a novel fatty acid mediated cholesterol independent mechanism. *Sci. Rep.* **5**, 13823.

Kamiyama, Y., Matsubara, T., Yoshinari, K., Nagata, K., Kamimura, H., and Yamazoe, Y. (2007). Role of human hepatocyte nuclear factor 4alpha in the expression of drug-metabolizing enzymes and transporters in human hepatocytes assessed by use of small interfering RNA. *Drug Metab. Pharmacokinet.* **22**, 287–298.

Kersten, S. (2014). Integrated physiology and systems biology of PPARalpha. *Mol. Metab.* **3**, 354–371.

Kobayashi, M., and Yamamoto, M. (2005). Molecular mechanisms activating the NRF2-Keap1 pathway of antioxidant gene regulation. *Antioxid. Redox Signal.* **7**, 385–394.

Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdottir, H., Tamayo, P., and Mesirov, J. P. (2011). Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740.

Linden, J., Lensu, S., Tuomisto, J., and Pohjanvirta, R. (2010). Dioxins, the aryl hydrocarbon receptor and the central regulation of energy balance. *Front. Neuroendocrinol.* **31**, 452–478.

Lu, H. (2016). Crosstalk of HNF4alpha with extracellular and intracellular signaling pathways in the regulation of hepatic metabolism of drugs and lipids. *Acta Pharm. Sin. B* **6**, 393–408.

Lu, H., Gonzalez, F. J., and Klaassen, C. (2010). Alterations in hepatic mRNA expression of phase II enzymes and xenobiotic transporters after targeted disruption of hepatocyte nuclear factor 4 alpha. *Toxicol. Sci.* **118**, 380–390.

Manikandan, P., and Nagini, S. (2018). Cytochrome P450 structure, function and clinical significance: A review. *Curr. Drug Targets* **19**, 38–54.

Mazein, A., Watterson, S., Hsieh, W. Y., Griffiths, W. J., and Ghazal, P. (2013). A comprehensive machine-readable view of the mammalian cholesterol biosynthesis pathway. *Biochem. Pharmacol.* **86**, 56–66.

Misra, P., Viswakarma, N., and Reddy, J. K. (2013). Peroxisome proliferator-activated receptor-alpha signaling in hepatocarcinogenesis. *Subcell. Biochem.* **69**, 77–99.

Monroe, J. J., Tanis, K. Q., Nguyen, T., Podtelezhnikov, A. A., Mahotka, S., Lynch, D., Evers, R., Palamanda, J., Pippert, T., Johnson, T., *et al.* 2020. Application of a rat liver drug bioactivation transcriptional response assay early in drug development that informs chemically reactive metabolite formation and potential for drug induced liver injury. *Toxicol. Sci.*

Moore, L. B., Parks, D. J., Jones, S. A., Bledsoe, R. K., Consler, T. G., Stimmel, J. B., Goodwin, B., Liddle, C., Blanchard, S. G., Willson, T. M., *et al.* (2000). Orphan nuclear receptors constitutive androstane receptor and pregnane x receptor share xenobiotic and steroid ligands. *J. Biol. Chem.* **275**, 15122–15127.

Nakata, K., Tanaka, Y., Nakano, T., Adachi, T., Tanaka, H., Kaminuma, T., and Ishikawa, T. (2006). Nuclear receptor-mediated transcriptional regulation in phase I, II, and III xenobiotic metabolizing systems. *Drug Metab. Pharmacokinet.* **21**, 437–457.

Nguyen, T. V., Ukairo, O., Khetani, S. R., McVay, M., Kanchagar, C., Seghezzi, W., Ayanoglu, G., Irrechukwu, O., and Evers, R. (2015). Establishment of a hepatocyte-kupffer cell coculture model for assessment of proinflammatory cytokine effects on metabolizing enzymes and drug transporters. *Drug Metab. Dispos.* **43**, 774–785.

Ohguchi, H., Tanaka, T., Uchida, A., Magoori, K., Kudo, H., Kim, I., Daigo, K., Sakakibara, I., Okamura, M., Harigae, H., *et al.* (2008). Hepatocyte nuclear factor 4 alpha contributes to thyroid hormone homeostasis by cooperatively regulating the type 1 iodothyronine deiodinase gene with GATA4 and Kruppel-like transcription factor 9. *Mol. Cell. Biol.* **28**, 3917–3931.

Pietrocola, F., Galluzzi, L., Bravo-San Pedro, J. M., Madeo, F., and Kroemer, G. (2015). Acetyl coenzyme A: A central metabolite and second messenger. *Cell Metab.* **21**, 805–821.

Qin, C., Aslamkhan, A. G., Pearson, K., Tanis, K. Q., Podtelezhnikov, A., Frank, E., Pacchione, S., Pippert, T., Glaab, W. E., and Sistare, F. D. (2019). AHR activation in pharmaceutical development: Applying liver gene expression biomarker thresholds to identify doses associated with tumorigenic risks in rats. *Toxicol. Sci.* **171**, 46–55.

Qin, C., Tanis, K. Q., Podtelezhnikov, A. A., Glaab, W. E., Sistare, F. D., and DeGeorge, J. J. (2016). Toxicogenomics in drug development: A match made in heaven? *Expert Opin. Drug Metab. Toxicol.* **12**, 847–849.

Rakhshandehroo, M., Knoch, B., Muller, M., and Kersten, S. (2010). Peroxisome proliferator-activated receptor alpha target genes. *PPAR Res.* **2010**, 1–20.

Rooney, J., Hill, T., 3rd, Qin, C., Sistare, F. D., and Corton, J. C. (2018). Adverse outcome pathway-driven identification of rat liver tumorigens in short-term assays. *Toxicol. Appl. Pharmacol.* **356**, 99–113.

Sato, R. (2010). Sterol metabolism and SREBP activation. *Arch. Biochem. Biophys.* **501**, 177–181.

Sha, Z., and Goldberg, A. L. (2014). Proteasome-mediated processing of NRF1 is essential for coordinate induction of all proteasome subunits and p97. *Curr. Biol.* **24**, 1573–1583.

Sistare, F. D., Morton, D., Alden, C., Christensen, J., Keller, D., Jonghe, S. D., Storer, R. D., Reddy, M. V., Kraynak, A., Trela, B., *et al.* (2011). An analysis of pharmaceutical experience with decades of rat carcinogenicity testing: Support for a proposal to modify current regulatory guidelines. *Toxicol. Pathol.* **39**, 716–744.

Skonberg, C., Olsen, J., Madsen, K. G., Hansen, S. H., and Grillo, M. P. (2008). Metabolic activation of carboxylic acids. *Expert Opin. Drug Metab. Toxicol.* **4**, 425–438.

Stevens, J. L., and Baker, T. K. (2009). The future of drug safety testing: Expanding the view and narrowing the focus. *Drug Discov. Today* **14**, 162–167.

Sutherland, J. J., Webster, Y. W., Willy, J. A., Searfoss, G. H., Goldstein, K. M., Irizarry, A. R., Hall, D. G., and Stevens, J. L. (2018). Toxicogenomic module associations with pathogenesis: A network-based approach to understanding drug toxicity. *Pharmacogenomics J.* **18**, 377–390.

Tamburino, A. M., Podtelezhnikov, A. A., Monroe, J. J., Glaab, W. E., Sistare, F. D., Gustafson, E., and Tanis, K. Q. (2020). Application of ChIP-Seq to refine transcriptional biomarker signatures of drug induced liver adaptation and toxicities. In preparation.

Tang, Q., Chen, Y., Meyer, C., Geistlinger, T., Lupien, M., Wang, Q., Liu, T., Zhang, Y., Brown, M., and Liu, X. S. (2011). A comprehensive view of nuclear receptor cancer cistromes. *Cancer Res.* **71**, 6940–6947.

Te, J. A., AbdulHameed, M. D., and Wallqvist, A. (2016). Systems toxicology of chemically induced liver and kidney injuries: Histopathology-associated gene co-expression modules. *J. Appl. Toxicol.* **36**, 1137–1149.

Thomas, R. S., and Waters, M. D. 2016. *Toxicogenomics in Predictive Carcinogenicity*. Royal Society of Chemistry, Cambridge.

Tolson, A. H., and Wang, H. (2010). Regulation of drug-metabolizing enzymes by xenobiotic receptors: PXR and CAR. *Adv. Drug Deliv. Rev.* **62**, 1238–1249.

Tugendreich, S., Pearson, C. I., Sagartz, J., Jarnagin, K., and Kolaja, K. (2006). NSAID-induced acute phase response is due to increased intestinal permeability and characterized by early and consistent alterations in hepatic gene expression. *Toxicol. Pathol.* **34**, 168–179.

Waring, M. J., Arrowsmith, J., Leach, A. R., Leeson, P. D., Mandrell, S., Owen, R. M., Pairaudeau, G., Pennie, W. D., Pickett, S. D., Wang, J., *et al.* (2015). An analysis of the attrition of drug candidates from four major pharmaceutical companies. *Nat. Rev. Drug Discov.* **14**, 475–486.

Watson, J. D., Prokopec, S. D., Smith, A. B., Okey, A. B., Pohjanvirta, R., and Boutros, P. C. (2014). TCDD dysregulation of 13 AHR-target genes in rat liver. *Toxicol. Appl. Pharmacol.* **274**, 445–454.

Xu, C., Li, C. Y., and Kong, A. N. (2005). Induction of phase I, II and III drug metabolism/transport by xenobiotics. *Arch. Pharm. Res.* **28**, 249–268.

Ye, J., and DeBose-Boyd, R. A. (2011). Regulation of cholesterol and fatty acid synthesis. *Cold Spring Harb. Perspect. Biol.* **3**, pii: a004754.

Yoshinari, K. (2019). Role of nuclear receptors PXR and CAR in xenobiotic-induced hepatocyte proliferation and chemical carcinogenesis. *Biol. Pharm. Bull.* **42**, 1243–1252.

Zarn, J. A., Bruschweiler, B. J., and Schlatter, J. R. (2003). Azole fungicides affect mammalian steroidogenesis by inhibiting sterol 14 alpha-demethylase and aromatase. *Environ. Health Perspect.* **111**, 255–261.

Zeng, H., and Xu, W. (2015). Ctr9, a key subunit of PAFc, affects global estrogen signaling and drives ERalpha-positive breast tumorigenesis. *Genes Dev.* **29**, 2153–2167.