A Section 508–conformant HTML version of this article is available at https://doi.org/10.1289/EHP16024.

Integrating Transcriptomic and Targeted New Approach Methodologies into a Tiered Framework for Chemical Bioactivity Screening

Jesse D. Rogers,[1,2] Joseph L. Bundy,[1] Joshua A. Harrill,[1] Richard J. Judson,[1] Katie Paul-Friedman,[1] and Logan J. Everett[1]

[1]Center for Computational Toxicology and Exposure, Office of Research and Development, US Environmental Protection Agency, Durham, North Carolina, USA
[2]Oak Ridge Institute for Science and Education, Oak Ridge, Tennessee, USA

Background: With thousands of chemicals in commerce and the environment, rapid identification of potential hazards is a critical need. Combining broad molecular profiling with targeted *in vitro* assays, such as high-throughput transcriptomics (HTTr) and receptor screening assays, could improve identification of chemicals that perturb key molecular targets associated with adverse outcomes.

Objectives: We aimed to link transcriptomic readouts to individual molecular targets and integrate transcriptomic predictions with orthogonal receptor-level assays in a proof-of-concept framework for chemical hazard prioritization.

Methods: Transcriptomic profiles generated via TempO-Seq in U-2 OS and HepaRG cell lines were used to develop signatures composed of genes uniquely responsive to reference chemicals for distinct molecular targets. These signatures were applied to 75 reference and 1,126 nonreference chemicals screened via HTTr in both cell lines. Selective bioactivity toward each signature was determined by comparing potency estimates against the bulk of transcriptomic bioactivity for each chemical. Chemicals predicted by transcriptomics were confirmed for target bioactivity and selectivity using available orthogonal assay data from the US Environmental Protection Agency ToxCast program. A subset of 37 selectively acting chemicals from HTTr that did not have sufficient orthogonal data were prospectively tested using one of five receptor-level assays.

Results: Of the 1,126 nonreference chemicals screened, 201 demonstrated selective bioactivity in at least one transcriptomic signature and 57 were confirmed as selective nuclear receptor agonists. Chemicals bioactive for each signature were significantly associated with orthogonal assay bioactivity, and signature-based points-of-departure were equally or more sensitive than biological pathway altering concentrations in 95.4% of signature-prioritized chemicals. Prospective profiling found that 18 of 37 (49%) chemicals without prior orthogonal assay data were bioactive against the predicted receptor.

Discussion: Our work demonstrates that integrating transcriptomics with targeted orthogonal assays in a tiered framework can support Next Generation Risk Assessment by informing putative molecular targets and prioritizing chemicals for further testing. https://doi.org/10.1289/EHP16024

Introduction

Thousands of chemicals have been registered in the United States for commercial nonfood or drug applications,[1] many of which do not have sufficient data for assessment of human or ecological risk. Traditional toxicity testing using intact animals often requires multiyear studies with high monetary and animal costs.[2] These requirements limit the available data necessary to efficiently assess the safety of chemicals in commerce. New approach methodologies (NAMs) have gained traction for the ability to screen chemicals at higher throughput and lower cost in comparison with traditional toxicity testing.[3,4] The US Environmental Protection Agency (US EPA) developed the ToxCast program for high-throughput chemical testing by screening thousands of chemicals across a variety of *in vitro* bioactivity assays.[5] Augmenting the capabilities to efficiently profile chemical effects *in vitro*, the US EPA and others are using high-throughput transcriptomics (HTTr) to broadly profile chemical-induced gene expression changes across multiple cell lines.

Combining broad-profiling and targeted NAMs into a decision framework could advance efforts to efficiently screen chemicals for key molecular targets related to human health effects. In a tiered framework proposed in the Next Generation Blueprint of Computational Toxicology at the US EPA,[6] chemicals would be initially screened via broad-profiling assays covering a range of biological pathways such as HTTr ("Tier 1"). Chemicals demonstrating bioactivity for a target or pathway would then be profiled using targeted assays for orthogonal confirmation ("Tier 2"), and organotypic assays related to potential health effects ("Tier 3") or traditional animal-based toxicity testing could be used as follow-up tests.

One challenge in integrating NAM-based data in such a framework is identifying key molecular targets from broad profiling data for confirmation with targeted assays.[7] Although targeted biological effects are frequently mined from transcriptomic data using gene sets related to signaling pathways, characteristic diseases, and other biological events,[8–10] such gene sets may not capture pathways modulated by xenobiotics in the assay or cell type used for *in vitro* screening. Another challenge lies in determining criteria for prioritizing chemicals based on both broad-profiling and targeted NAMs within a screening context. Multiple approaches could be used to combine HTTr and targeted screens, such as in tandem or in a tiered fashion, to improve confidence in the putative molecular targets that might be perturbed in various exposure scenarios.[11–14]

In this work, we employed a tiered framework to prioritize chemicals based on mechanism-of-action, starting with HTTr and followed by targeted testing. To do this, we developed a computational strategy to *a*) create transcriptomic signatures tailored to specific cell types that are indicative of changes in key molecular targets, *b*) apply these signatures to screening data for nonreference chemicals, and *c*) validate signature-level readouts with orthogonal assay end points from ToxCast to prioritize chemicals for further testing and detection of potential hazards. This work

Address correspondence to Jesse D. Rogers, 109 TW Alexander Dr., Durham, NC 27711, USA. Email: jessedrogers2@gmail.com

Supplemental Material is available online (https://doi.org/10.1289/EHP16024).

The authors declare they have no conflicts of interest related to this work to disclose.

Conclusions and opinions are those of the individual authors and do not necessarily reflect the policies or views of EHP Publishing or the National Institute of Environmental Health Sciences.

Received 16 August 2024; Revised 3 February 2025; Accepted 2 May 2025; Published 13 June 2025.

Note to readers with disabilities: *EHP* strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in *EHP* articles may not conform to 508 standards due to the complexity of the information being presented. If you need assistance accessing journal content, please contact ehpsubmissions@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.

demonstrates the utility of a tiered framework in both retrospective analyses and prospective screening for rapid chemical hazard prioritization and safety assessment.

Methods

Reference Chemical Selection

Reference chemicals annotated for individual molecular targets were compiled using the RefChemDB resource, which links over 40,000 chemicals to the modulation of gene products via semiautomated curation of 15 public sources.[15] Using the RefChemDB candidate reference chemical set (accordingly as in Judson et al.[15]) molecular target and mode of regulation (i.e., positive or negative) columns were concatenated into a single molecular target annotation. Each annotation in RefChemDB contains a numeric support value, defined as the number of public sources confirming bioactivity toward a target and mode. Chemical–target pairs were filtered for a support value of at least 5 (i.e., all chemicals must have at least five sources supporting bioactivity for the annotated target and mode). Underlying records that were duplicated in multiple public sources were only counted once toward each annotation's support value.[15]

Some groups of molecular targets had similar sets of annotated chemicals. For example, 10 of 18 chemicals linked to ESR1 or ESR2 agonism were linked to both paralogs. Few chemicals in RefChemDB are uniquely linked to a single paralog or protein within a tightly coupled process such as heterodimerization. Target annotations were therefore collapsed into clusters based on the similarity of annotated chemicals between any two targets as previously described.[16] In brief, chemicals annotated or not annotated for each target were encoded as binary bits, and Jaccard distances were calculated between all pairs of targets. Hierarchical clustering was performed to cluster targets by Jaccard distance via the *hclust* function in R statistical software (version 4.4.1; R Development Core Team). Clusters were chosen using a dendrogram cut height of 0.8 to optimally group target paralogs or families without grouping unrelated targets. Individual reference chemicals were assigned to a single cluster based on the highest number of individual targets linking the chemical to the cluster. Any ties between clusters were broken using the highest sum of support values for chemical–target annotations within each cluster. Reference chemicals were further filtered for those that were tested and had at least one concentration-responsive gene in HTTr screening data. Retained reference chemical groupings are referred to as "reference classes." A total of 14 reference classes had at least 3 different reference chemicals for a total of 75 unique reference chemicals (69 in U-2 OS, 66 in HepaRG). All reference classes are summarized in Figure S1 and detailed in Excel Table S1.

HTTr Screening Data

HTTr screens of 1,201 chemicals in U-2 OS and HepaRG cell lines[17] were used as source data for all transcriptomic analyses. Chemicals were selected to reflect those in the US EPA ToxCast chemical inventory including agricultural and industrial use compounds as well as a selection of pharmaceuticals. The U-2 OS cell line was chosen for its previous characterization and use in HTTr,[17] and the HepaRG cell line was chosen for its increased metabolic competency and higher similarity in expression profiles to primary human hepatocytes vs. previous cell lines. Both screens used the same experimental design and bioinformatics pipeline for initial derivation of transcriptional potency values.[17] In brief, dimethylsulfoxide (DMSO)-solubilized chemical stock solutions were provided frozen from the US EPA ToxCast chemical inventory management contractor (EvoTec; Princeton, New Jersey) and stored at −80°C prior to dose plate preparation. U-2 OS osteosarcoma and HepaRG hepatocyte cell lines were used for chemical exposure (see Excel Table S2 for cell culture details). Chemicals were tested at 8 nominal concentrations ranging from 0.03–100 μM using 0.5-log spacing with a final DMSO concentration of 0.05%. All exposures were conducted in triplicate using independent cell cultures. At 24 h after chemical exposure, cell lysates were prepared for TempO-Seq analysis as previously described.[18] The TempO-Seq human whole transcriptome v2 assay (hWTv2) with cell line–specific custom attenuation was performed by BioSpyder, Inc.[19]

Raw FASTQ files were processed using a custom bioinformatics pipeline as previously described,[18] including read alignment, sample quality control, differential expression analysis, and concentration–response modeling. Moderated $\log_2$ (fold change) values were generated from the *DESeq2* R package (version 1.24)[20] for all concentrations of each chemical–gene pair. Concentration–response modeling was then conducted using the *tcplfit2* R package (version 0.1.5) by fitting $\log_2$ (fold change) values to nine models (constant, hill, poly1-2, power, exp2-5) and selecting the model with the minimum Akaike Information Criterion.[21] Benchmark concentrations (BMCs) were computed for each chemical–gene pair based on benchmark response values representing departure from a null response distribution. The null distribution cutoff was defined as 1.349 times the standard deviation (SD) of responses from the lowest two concentrations of all test samples,[22] which assumes that these concentrations will be inactive for a large majority of the chemical test set.[23]

Key metrics from *tcplfit2* include continuous hitcalls, which represent confidence in winning model selection and that model responses exceed cutoffs, and normalized efficacy values (top_over_cutoff), which represent the signal-to-noise ratio.[21] A hitcall ≥ 0.9 was chosen to match previous HTTr and ToxCast screening studies.[24] A top_over_cutoff ≥ 1.5 was chosen empirically to limit the potential for noise when fitting individual genes. Concentration-responsive ("active") chemical–gene pairs were defined for downstream analyses using both criteria above. BMCs for active genes were bounded to be within 0.1 times the lowest concentration tested and 10 times the highest concentration tested after modeling to match previous signature-based analyses.[18,23]

RCAS Identification

Log_{10}-transformed BMCs were used as gene sensitivity estimates toward all reference chemicals, and a default value of 2.5 (0.5-log greater than the highest concentration tested) was assigned for inactive chemical–gene pairs. Using the reference classes identified above, a univariate analysis was conducted to identify genes that were uniquely sensitive toward one reference class vs. other classes (Figure S2). Only genes that were active for one or more reference chemicals were retained for analysis.

For each gene, a one-way analysis of variance (ANOVA) was used to test for differences in $\log_{10}$BMCs across reference classes. *p*-Values for all tested genes were converted to false discovery rate (FDR) using Benjamini-Hochberg multiple testing correction. Genes demonstrating an FDR ≤ 0.05 were tested for pairwise comparisons of reference classes using Tukey's post hoc tests with Benjamini-Hochberg correction. A gene selection step was then applied based on post hoc test results: for gene *g*, if *a*) the mean $\log_{10}$BMC for a single reference class *X* was significantly lower (FDR ≤ 0.05) than 70% of the other classes, and *b*) the mean $\log_{10}$BMC met this criterion for only reference class *X*, then gene *g* was assigned as uniquely potent for class *X*. The 70% threshold was chosen to maintain a threshold for specificity while increasing the pool of included genes. Each resulting gene set identified for a reference class and cell line is referred to as a reference class–

associated signature (RCAS). RCASs containing ≥10 genes were retained for further analysis to address sparse representation during signature concentration–response profiling to ensure interoperability with previous signature-based analyses.[17,18] RCASs were identified independently for U-2 OS and HepaRG screening data, which allowed signatures to capture uniquely sensitive transcriptional behavior per cell line. A total of three and two RCASs were identified from HepaRG and U-2 OS screening data, respectively, and RCAS-identified genes are detailed in Excel Table S3.

The FDR of all ANOVA or Tukey's post hoc tests during RCAS generation was tested to evaluate whether additional genes and/or RCAS could be detected. The number of genes in each RCAS using the current FDR ≤ 0.05 was compared to the number of genes if an FDR ≤ 0.1 were used for either ANOVA tests (Figure S3A) or post hoc tests (Figure S3B). Relaxation of either threshold added between 0 and 36 genes to each RCAS representing a maximum 6.4% increase in signature size, and no additional RCASs were identified when relaxing these criteria.

Concentration–response modeling of each RCAS in both cell lines was performed using the *httrpathway* R package (https://github.com/USEPA/CompTox-httrpathway) as previously described.[18] DESeq2-moderated $\log_2$ (fold change) values for all genes used in the gene-level concentration–response profiling above were used for signature scoring. Normalized enrichment scores (NES) for each chemical–concentration pair were computed for each RCAS via single sample gene set enrichment analysis (ssGSEA)[25] using the *GSVA* R package (version 1.50.0)[26] as previously described.[18,23] Concentration–response modeling was performed on NES values for each chemical-RCAS pair using *tcplfit2* (version 0.1.5). Active chemical–RCAS pairs were defined using a hitcall ≥ 0.9 and top_over_cutoff ≥ 1.5 to match the gene-level criteria above. The resulting BMCs generated per RCAS are hereafter referred to as an RCAS-derived point of departure (POD_{RCAS}).

RCAS Validation

To evaluate possible type I errors in RCAS concentration–response profiling results, permutation tests were conducted for each RCAS vs. 100 randomly generated signatures. Genes comprising all RCASs were first pooled together for each cell line. Genes were then randomly selected from the pool without replacement to generate random signatures. A random number of genes were assigned to each signature following the size distribution for RCAS in each cell line. Genes were replaced between each signature generation, which allowed genes to be used in multiple signatures. Concentration–response profiling of all reference chemicals in random signatures was performed as described above. Reference chemicals were divided into those matching or not matching each RCAS by target. Nonmatching reference chemicals were randomly down sampled to three times the number of matching reference chemicals for each RCAS to reduce class imbalances. The difference in means between matching and nonmatching $\log_{10}POD_{RCAS}$ was computed for all signatures. An empirical *p*-value was then computed for each RCAS as the probability that a random signature's difference in means is lower than that of the RCAS.

To further evaluate possible type I and type II error against screening data separate from signature development, RCAS were tested against HTTr screening data for an external dataset not used in RCAS identification. This dataset comprises 29 replicated concentration–response experiments for three chemicals each in U-2 OS and HepaRG cell lines. RARA agonist all-*trans*-retinoic acid, GR agonist dexamethasone, and TOP2A inhibitor etoposide were profiled in 8-point concentration–response experiments U-2 OS cells,[17] and AHR agonist benzo[a]pyrene, PXR agonist rifampicin, and PPARG agonist troglitazone were profiled in 7-point concentration–response experiments in HepaRG cells. Each chemical and replicate was profiled separately for RCAS in the matching cell line as described above, and active sample-RCAS pairs were defined using a hitcall ≥ 0.9 and top_over_cutoff ≥ 1.5.

ToxCast HTS Data

High-throughput screening (HTS) bioactivity data from the US EPA ToxCast program[5] was used as source data for all Tier 2 analyses. These data consist of hundreds of *in vitro* assay measurements for up to 9,000 chemicals in both single- and multiconcentration formats, and all multiconcentration measurements were processed using *tcpl* (version 2.0.2)[27] to produce binary hitcall values in addition to potency estimates based on the winning curve-fitting model between Hill, gain-loss, or constant models. Level 5 multiconcentration curve fit data for all chemicals tested in HTTr were downloaded from InvitroDB v3.4.[28] Active chemical–end point pairs were defined as having a positive hitcall and <3 quality control flags identified during level-6 processing.[29] Because BMCs were not calculated as part of InvitroDB v3.4, activity concentration at cutoff (ACC) values were used for all potency comparisons. These values are conceptually similar to BMCs as a quantitative estimate of the concentration at which assay responses exceed assay-dependent noise distribution cutoffs. In InvitroDB, cutoff values are determined for individual end points based on each detection technology as encoded in the database schema.[24] ToxCast assay end points are additionally annotated to the most relevant Entrez gene symbol and official names as a convenience mapping in InvitroDB to enable data interoperability, and these annotations were used to map assay end points to RCAS target annotations.

Orthogonal End Point Selection

Initial end points or orthogonal confirmation of each RCAS were selected by matching official gene symbols for each ToxCast end point to the gene symbol(s) of each RCAS. Assay end points that were annotated for the same target as an RCAS but measured a different biological process (e.g., corticosterone synthesis in comparison with an RCAS measuring glucocorticoid receptor activation) were removed from consideration. Unidirectional assay end points without biological relevance (e.g., negative signal direction in Attagene trans-FACTORIAL end points) were also removed.

A two-phase analysis was employed to select end points that distinguish bioactivity between in-class and out-of-class reference chemicals with low type I error. All reference chemicals annotated for each reference class in RefChemDB (not limited to those tested in HTTr) were used for each analysis. Chemicals were denoted as either matching or not matching each ToxCast end point based on their annotated target. Hitcalls of each ToxCast end point were first compared across reference chemicals. Fewer matching reference chemicals were tested in comparison with nonmatching reference chemicals in all end points, which resulted in imbalanced chemical sets. To mitigate possible effects of this imbalance on sensitivity, we retained ToxCast end points that contained a positive hitcall for at least 50% of matching reference chemicals and a negative hitcall for at least 75% of nonmatching reference chemicals. Nonmatching reference chemicals with a positive hitcall may exhibit differences in potency from that of matching chemicals. Therefore, end points that failed the criteria above were tested for differences in $\log_{10}ACC$ values between matching and nonmatching reference chemicals via Wilcoxon rank-sum tests with Benjamini-Hochberg correction. End points with significant differences in $\log_{10}ACC$ (FDR ≤ 0.05) were also retained

as orthogonal end points, and all other end points were removed from further analysis. Orthogonal end points for all RCAS-associated molecular targets are shown in Table 1.

Selectivity Assessments

Selectivity was defined as activity at a particular target at concentrations lower than those thought to elicit cytotoxicity and/or myriad nonspecific activities across pathways.[30] Chemicals were tested for selectivity of each RCAS and orthogonal ToxCast end point in comparison with estimates of nonselective activity. For HTTr data, a set of 11,037 gene sets or signatures derived from multiple public sources ("public signatures") were used for signature concentration–response profiling as previously described.[17] This collection is intended to provide broad coverage of molecular functions for use in determining a biological phenotype altering concentration (BPAC)[23] rather than for interpreting a cell line–specific mechanism. For ToxCast data, all measured end points were used to derive nonspecific activity estimates.

For HTTr data, the first statistical mode of $\log_{10}$BMC values from active public signatures ($Mode_{NS1}$) was estimated for each chemical using the density function in R. The standard deviation of $\log_{10}$BMC values from active public signatures (σ_{NS1}) was also determined for each chemical. For ToxCast data, the first statistical mode of $\log_{10}$ACC values ($Mode_{NS2}$) and standard deviation of $\log_{10}$ACC values (σ_{NS2}) were similarly determined for each chemical from all active end points. For chemicals with fewer than 10 active public signatures in HTTr or fewer than 10 active end points in ToxCast, $Mode_{NS1}$ and $Mode_{NS2}$ values were set to 0.5-log above the highest tested concentration because these chemicals were largely inert across pathways in each data stream. Nonselective PODs for Tier 1 and Tier 2 (POD_{NS1} and POD_{NS2}, respectively) were determined from mode and SD estimates: $POD_{NS} = Mode_{NS} - \sigma_{NS}$.

Chemical selectivity for each Tier was determined by comparing POD_{NS} estimates to RCAS-based PODs (POD_{RCAS}, defined as the $\log_{10}$BMC for each RCAS) or orthogonal end point PODs (POD_{ortho}, defined as the $\log_{10}$ACC for each orthogonal end point). Selective chemicals were defined for Tier 1 as $POD_{RCAS} < POD_{NS1}$ and for Tier 2 as $POD_{ortho} < POD_{NS2}$. Overall PODs reflecting the BPAC were additionally calculated for all chemicals using HTTr and ToxCast screening data. HTTr-based BPACs (POD_{BPAC1}) were determined as the fifth percentile of $\log_{10}$BMC values from all active public signatures for each chemical as described previously.[23] ToxCast-based BPACs (POD_{BPAC2}) were similarly determined as the fifth percentile of $\log_{10}$ACC values from all active end points. At least 10 active public signatures or 10 active ToxCast end points were required to calculate a BPAC for any chemical. Refer to Table 2 for descriptions of all abbreviated terms used for bioactivity and selectivity testing.

Prospective Profiling of Tier 1-Predicted Chemicals

For a subset of chemicals that were selective for one or more RCAS but had <2 orthogonal end points measured in ToxCast, prospective profiling was performed using one or more targeted *in vitro* assays. A total 34 test chemicals were selected by ranking the difference between POD_{RCAS} and POD_{NS1} values for each chemical. Three positive controls were selected from reference chemicals for each assay based on positive hitcalls for ToxCast end points with matching Entrez gene symbols. Three negative controls were conversely chosen based on negative hitcalls for matching end points and positive hitcalls for one or more nonmatching end points.

Table 1. Orthogonal assay end points from ToxCast high-throughput screening data used for tier 2 testing.

Abbreviation	Assay end point	Source	Target gene symbol(s)	Species	Cell line	Assay detection type	Detection technology
AHR1	ATG_Ahr_CIS_up	Attagene	AHR	Human	HepG2	Transcription factor activity	mRNA induction
AHR2	TOX21_AhR_LUC_Agonist	NCATS/NCGC	AHR	Human	HepG2	Transcription factor activity	Luciferase induction
GR1	ATG_GRE_CIS_up	Attagene	NR3C1	Human	HepG2	Transcription Factor Activity	mRNA induction
GR2	ATG_GR_TRANS_up	Attagene	NR3C1	Human	HepG2	Transcription factor activity	mRNA induction
GR3	TOX21_GR_BLA_Agonist_ratio	NCATS/NCGC	NR3C1	Human	HeLa	Transcription factor activity	Beta lactamase induction
RAR1	NVS_NR_hRARa_Agonist	Novascreen	RARA	Human	Cell-free	Receptor binding	TR-FRET
RAR2	ATG_DR5_CIS_up	Attagene	RARA/B/G	Human	HepG2	Transcription factor activity	mRNA induction
RAR3	TOX21_RAR_LUC_Agonist	NCATS/NCGC	RARA	House	C3H10T1/2	Transcription factor activity	Luciferase induction
RXR1	ATG_RXRb_TRANS_up	Attagene	RXRB	Human	HepG2	Transcription factor activity	mRNA induction
RXR2	ATG_RARg_TRANS_up	Attagene	RARG	Human	HepG2	Transcription factor activity	mRNA induction
RXR3	TOX21_RXR_BLA_Agonist_ratio	NCATS/NCGC	RXRA	Human	HEK293T	Transcription factor activity	Beta lactamase induction
RXR4	TOX21_TR_RXR_BLA_Agonist_ Followup_ratio	NCATS/NCGC	RXRA	Human	RXRa-UAS-bla HEK 293T	Transcription factor activity	Beta lactamase induction

Note: AHR, aryl hydrocarbon receptor; GR, glucocorticoid receptor; RAR, retinoic acid receptor; RXR, retinoid X receptor; TR-FRET, time-resolved fluorescence energy transfer.

Table 2. Description of terms used throughout the current study for bioactivity and selectivity testing.

Acronym	Term	Description
HTTr	High-throughput transcriptomics	Tier 1 screening assay for transcriptional bioactivity in one or more human-derived cell lines
HTS	High-throughput screening	Tier 2 screening assays for individual molecular targets in cell-free or cell-based platforms
RCAS	Reference class–associated signatures	Gene sets comprised of genes with uniquely sensitive expression toward reference chemicals with a common mechanism-of-action
Hitcall	Hitcall	Measure of confidence in chemical-induced bioactivity for a gene, signature or assay based on concentration–response profiling
Top_over_cutoff	Top over cutoff	Measure of relative signal strength compared with null distribution of assay responses
BMC	Benchmark concentration	The estimated nominal concentration in HTTr at which a gene or signature demonstrates altered expression in comparison with the null distribution of responses
ACC	Activity concentration at cutoff	The estimated nominal concentration in HTS at which an assay demonstrates altered responses in comparison with an assay-specific cutoff of null responses
BPAC	Biological pathway altering concentration	The estimated nominal concentration at which a chemical can significantly alter a biological pathway, defined as the fifth percentile of BMCs for active signatures in HTTr or the fifth percentile of ACCs for active assays in HTS
POD_{RCAS}	RCAS-derived point of departure	Tier 1 point of departure for an individual molecular target, defined as the BMC estimated for a given RCAS
POD_{NS1}	Nonselective Tier 1 point of departure	Tier 1 point of departure for nonselective gene expression, defined as the statistical mode minus 1 standard deviation of BMCs from a set of public signatures
POD_{BPAC1}	BPAC-derived point of departure: Tier 1	Tier 1 point of departure for overall transcriptional activity, defined as the BPAC estimated from HTTr
POD_{ortho}	Orthogonal end point–derived point of departure	Tier 2 point of departure for an individual molecular target, defined as the ACC estimated for an assay whose target matches each RCAS
POD_{NS2}	Nonselective Tier 2 point of departure	Tier 1 point of departure for nonselective responses in HTS assays, defined as the statistical mode minus 1 standard deviation of ACCs from all active assays
POD_{BPAC2}	BPAC-derived point of departure: Tier 2	Tier 2 point of departure for overall receptor-level activity, defined as the BPAC estimated from measured HTS assays
$POD_{Prospective}$	Prospective end point–derived point of departure	Tier 2 point of departure for chemicals screened in prospective orthogonal assays, defined as the ACC from each orthogonal assay

DMSO-solubilized chemical stock solutions were provided frozen from the US EPA ToxCast chemical inventory management contractor and stored at −80°C prior to dose plate preparation. Dose plates were shipped to Eurofins DiscoverX or Eurofins Panlabs frozen, and all assays were performed by Eurofins DiscoverX or Eurofins Panlabs (Table 3). Chemicals were tested at 5 nominal concentrations ranging from 0.3–30 μM with 0.5-log spacing and 2 technical replicates per concentration to maximize chemical coverage while still reliably performing concentration–response profiling. Data were received from vendors as control-normalized percent activation or inhibition values, and concentration–response modeling was conducted using *tcplfit2* (version 0.1.5)[21] per its integration with the ToxCast Pipeline.[24] Baseline median absolute deviations (BMADs) were determined per assay using responses from the lowest two concentrations of all chemicals tested, and response cutoffs were determined by the following ruleset: *a*) if 3× BMAD <20% activation or inhibition, the cutoff was set to 20%; *b*) if 3×BMAD >50% activation or inhibition, the cutoff was set to 50%; and *c*) if 20% <3× BMAD <50%, the cutoff was set to 3× BMAD. These bounds were established to account for the small and biased chemical selection of the study, in which low concentrations of preselected chemicals may induce large responses, while maintaining a minimum response requirement for bioactivity. All other parameters for *tcplfit2* were set to default values, and active concentration–response profiles were defined as hitcall ≥ 0.9 and top_over_cutoff ≥ 1.

Software and Data Availability

R scripts for reproducing all analyses are available on GitHub (https://github.com/USEPA/CompTox-HTTr-RCAS). All sequencing data are available via the Gene Expression Omnibus repository (accession numbers GSE274318 for U-2 OS and GSE284321 for HepaRG). HTS assay data are available from InvitroDB via download[28] or the US EPA CompTox Chemicals Dashboard (https://comptox.epa.gov/dashboard/).

Results

Data-Driven Signatures for Targeted Activity from Transcriptomic Screening Data

We first aimed to infer key molecular targets related to potential adverse outcomes from transcriptomic screening data. To accomplish this goal, we summarized gene expression data from published HTTr screens in U-2 OS and HepaRG cell lines[17] into gene sets based on patterns of potency among well-annotated reference chemicals (Figure 1A). We first identified 14 molecular targets from the RefChemDB semiautomated literature mining resource[15] comprising a total 75 reference chemicals (Figure S1; see Excel Table S1 for all targets and chemicals). Each target was required to have at least three reference chemicals that were both tested in the above screens and caused concentration-dependent changes in gene expression. Genes linked to each target were also evaluated for basal expression in vehicle control samples ($n = 432$ or 433 replicates for U-2 OS and HepaRG, respectively), and these genes were found to be expressed across samples in both cell lines (Figure S4).

From these 14 targets that met our criteria for testing, we developed 5 RCAS that had least 10 genes (range 24–540 genes) that were each uniquely sensitive for a single target (FDR ≤ 0.05 via one-factor ANOVA and Tukey's post hoc). From U-2 OS screening data, RCAS were identified for glucocorticoid receptor (GR; gene symbol NR3C1) and retinoic acid/retinoid X receptor (RAR/RXR) agonism. From HepaRG screening data, RCASs were identified for aryl hydrocarbon receptor (AHR) and RAR/RXR agonism as well human Ether-à-go-go-Related Gene (hERG) ion channel antagonism. A maximum of 18 genes were common to any pair of RCAS (Figure S5), and this set of overlapping genes belonged to the U-2 OS and HepaRG-based RCAS for RAR/RXR agonism. We investigated the specificity of individual genes assigned to each RCAS by comparing gene BMCs between reference chemicals annotated for any of the five RCAS

Table 3. Orthogonal end points from prospective HTS profiling of Tier 1–predicted chemicals.

Target	Assay end point	Source	Species	Cell line	Assay detection type	Detection technology	Chemicals tested	Cutoff applied (% activation)
AHR	ERF_DX_NR_hAHR	Eurofins DiscoverX	Human	U-2 OS	Receptor complex stabilization	Protein–protein interaction	12	24
GR	ERF_DX_NR_hGR	Eurofins DiscoverX	Human	CHO-K1	Receptor complex stabilization	Protein–protein interaction	4	50
RARG	ERF_PL_NR_COA_hRARG	Eurofins Panlabs	Human	Cell-free	Receptor coactivation	TR-FRET	4	22
RXRB	ERF_PL_NR_COA_hRXRB	Eurofins Panlabs	Human	Cell-free	Receptor coactivation	TR-FRET	4	20
hERG	ERF_PL_IC_hKCNH2	Eurofins Panlabs	Human	HEK293	Receptor binding	Scintillation counting	16	46

Note: AHR, aryl hydrocarbon receptor; GR, glucocorticoid receptor; hERG, Ether-à-go-go-related gene; HTS, high-throughput screening; RARG, retinoic acid receptor gamma; RXRB, retinoid X receptor beta; TR-FRET, time-resolved fluorescence energy transfer.

(Figure 1B,C). Gene BMCs for reference chemicals matching each RCAS were up to 5.5 orders of magnitude lower than genes assigned a different RCAS. Most genes were not active in nonmatching reference chemicals ($85 \pm 9.9\%$ and $88 \pm 21\%$ of RCAS-identified genes on average for HepaRG and U-2 OS RCAS, respectively). Hierarchical clustering of genes by BMC estimates yielded primary groupings of genes by their RCAS assignment.

RCAS generation did not identify a uniquely sensitive gene set for peroxisome proliferator–activated receptor (PPAR) agonism in the HepaRG cell line despite well-studied linkages to liver function.[31] We hypothesized that separating this family into references classes for PPAR subtypes PPARA, PPARG, and PPARD could improve the detection of uniquely sensitive genes, because each receptor can play different roles in metabolic pathways.[32,33] Of the reference chemicals annotated for PPAR agonism, six chemicals were found to have a majority support for PPARA, whereas PPARG and PPARD each had one chemical uniquely supported for each receptor (Figure S6A). We therefore generated RCASs using the PPARA agonist subset of reference chemicals in place of all PPAR agonists. Five genes were uniquely sensitive toward these chemicals (Figure S6B), including several that have been linked to PPARA-specific bioactivity (PDK4,[34] CYP4A11,[35] CREB3L3,[36] and HSDL2).[37] However, too few genes were identified to meet criteria for signature-based profiling even after limiting our analysis to PPARA agonists.

We next evaluated whether RCAS can detect target-specific bioactivity at the signature level. We profiled reference chemicals for each RCAS using ssGSEA[25] followed by concentration–response modeling to estimate signature-level BMCs,[38] which are hereafter referred to as RCAS-derived PODs (POD_{RCAS}). By using ssGSEA for signature scoring, signal can be detected from chemicals that induce weak but coordinated changes in gene expression that may not be observed from analysis of individual genes. We found that 89.5% of reference chemicals matching each RCAS were bioactive for their respective signatures, whereas only 10.0% of nonmatching chemicals were bioactive for each RCAS (Figure 1D). Matching reference chemicals also ranked among the lowest POD_{RCAS} estimates and highest top_over_cutoff values of active reference chemicals. Each RCAS derived from HepaRG data demonstrated more frequent bioactivity in nonmatching chemicals (between 16 and 18 of 120 nonmatching chemicals) than those derived from U-2 OS data (between 1 and 7 of 120 chemicals). This increased promiscuity could be due to an increase in reactive metabolites in HepaRG cells, but additional investigation would be needed to determine potential drivers of this behavior.

We next conducted permutation testing of POD_{RCAS} estimates in comparison with 100 signatures generated by random assignment of RCAS genes. All RCASs had a significantly greater difference in mean POD_{RCAS} between matching and nonmatching reference chemicals than random signatures (Figure 2), suggesting that potency-based assignment of genes can distinguish potencies between reference chemicals with limited type I error. We further tested RCASs against 174 held-out test samples linked to individual molecular targets and not used in RCAS development.[17] RCAS were bioactive for all 87 test samples with matching targets based on signature-level profiling results, and few to no test samples linked to different targets were found to be active (Figure S7A). The PPAR agonist troglitazone was active for the mismatching heparg_hERG_inhibitor RCAS in 22 of 29 replicates, but POD_{RCAS} estimates for these replicates ranged among the higher concentrations tested (4.8–85 μM) (Figure S7B). POD_{RCAS} estimates for all troglitazone samples were also higher than BPACs derived from the fifth percentile of active BMCs from a collection

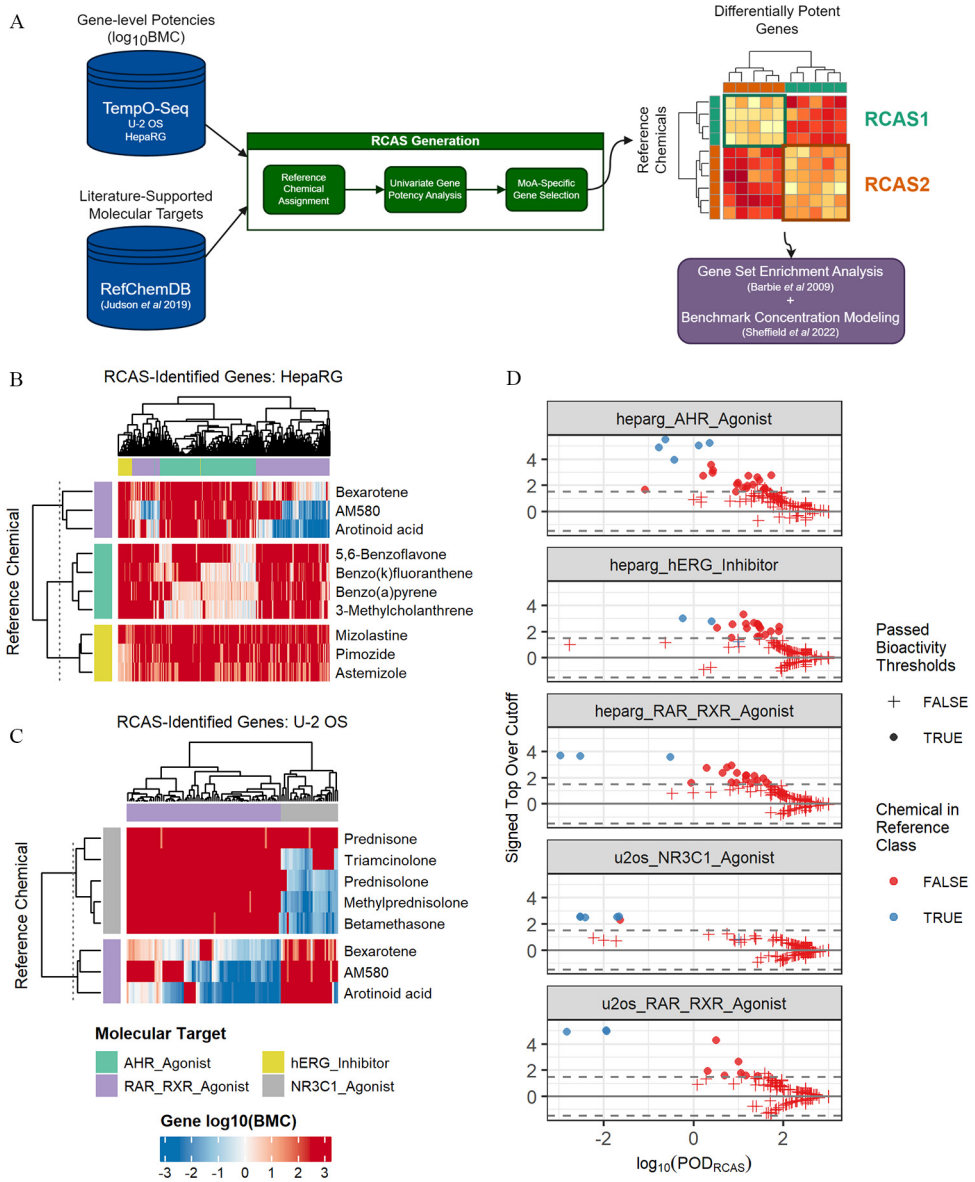

Figure 1. Development of Reference Chemical-Associated Signatures (RCAS) from high-throughput transcriptomics data. (A) Schematic of prior datasets and processing steps used to identify genes for each RCAS based on unique sensitivity for a single class of reference chemicals. Resulting clusters of genes (columns) represent hypothetical RCAS for two groups of reference chemicals (rows). (B,C) Comparison of gene BMCs for RCAS genes across reference chemicals used for identification in HepaRG (B) and U-2 OS (C) cell lines. Genes identified as part of any RCAS (columns) are shown for reference chemicals associated with any RCAS (rows). Inactive genes are designated by a default value of 2.5 (i.e., 0.5-log above the highest tested concentration). (D) RCAS concentration-response profiling estimates across reference chemicals. Chemicals annotated for the same molecular target as each RCAS (blue) and those annotated for a different target (red) are compared using POD_{RCAS} estimates (*x*-axis) and effect size (*y*-axis). Chemicals with passing bioactivity for each RCAS were determined using a hitcall ≥ 0.9 and top_over_cutoff ≥ 1.5 (dashed lines). See Excel Table S6 for all underlying data. Note: BMC, benchmark concentration; POD_{RCAS}, RCAS-derived point of departure; RCAS, reference class–associated signature.

of public signatures (Figure S7C), which suggests that inhibition of hERG was likely not the primary effect of troglitazone. Thiazolidinediones have additionally been linked to ion transport modulation[39–41] among broader effects on solute transport[42] and hepatotoxicity,[43] so it is not surprising that these samples demonstrated hERG-related bioactivity due to molecular crosstalk or general hepatotoxicity.

Tiered NAM-Based Decision Framework for Identifying Selectively Acting Perturbagens

We next aimed to investigate how mechanistic end points from HTTr could be integrated with targeted HTS assays in a chemical prioritization context. We developed a decision framework for identifying target-specific activity using both modalities as individual tiers (Figure 3A). Chemicals would first be tested for activity in Tier 1 (i.e., HTTr) using signature-level profiling of RCAS and resulting in a POD_{RCAS} estimate. Because activity for each RCAS could have been the result of nonselective perturbation at higher chemical concentrations,[30] we applied a selectivity test in which POD_{RCAS} estimates were compared to nonselective Tier 1 PODs (POD_{NS1}) estimated from a distribution of 11,037 signatures sourced from public repositories. Chemicals that demonstrate both activity ("Tier 1-active") and selectivity ("Tier 1-active and selective") for a target in Tier 1 were then similarly tested for activity and selectivity in Tier 2 using multiple

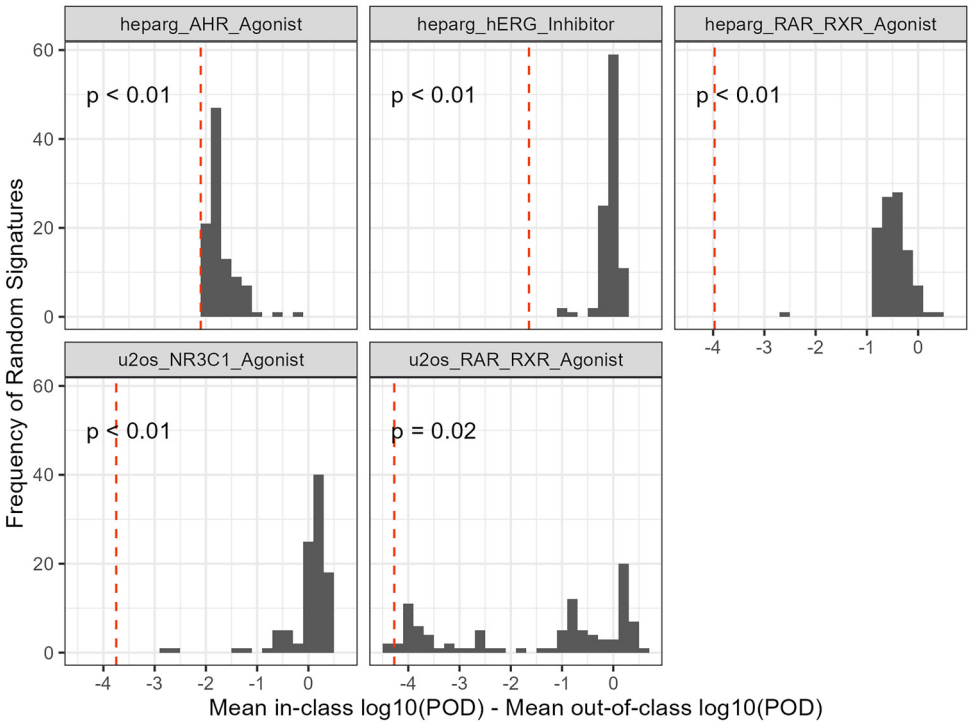

Figure 2. Comparison of RCAS signature predictions vs. random signature sets. Differences in mean $\log_{10}POD_{RCAS}$ estimates between positive reference chemicals (whose targets match the RCAS) and negative reference chemicals (whose targets do not match the RCAS) are shown for each RCAS (red dashed line) and for 100 randomly generated signatures (histogram). *p*-Values were computed as the probability that a random signature's difference in mean $\log_{10}POD_{RCAS}$ is lower than that for each RCAS. See Excel Table S7 for all underlying data. POD_{RCAS}, RCAS-derived point of departure; RCAS, reference class–associated signature.

orthogonal HTS assays from the US EPA ToxCast program. Chemicals that appeared active and selective toward a particular target in both tiers were considered candidate target modulators for further testing based on linked adverse outcomes.

We applied this framework to HTTr screening data in both U-2 OS and HepaRG cell lines[17] by using RCAS for Tier 1 tests and a set of high-confidence ToxCast end points matching each target for Tier 2 tests (Table 1; see Figure S8 for end point selection results). Of a total 1,201 chemicals screened in both cell lines, 228 chemicals were active and selective in Tier 1 for at least one RCAS, and 190 of these chemicals had existing data for orthogonal ToxCast end points in Tier 2 (Figure 3B). Of these, 64 chemicals were active and selective in both Tier 1 and Tier 2 tests for at least one RCAS and orthogonal ToxCast end point. Of these, 57 chemicals were not among the reference chemicals used for RCAS identification, and these were designated as candidate perturbagens that may elicit a targeted mechanistic response not widely identified in previous literature.

To assess RCAS performance in predicting orthogonal Tier 2 activity, we tested the association between all Tier 1-active chemicals and all Tier 2-active chemicals (Figure 3C, left). We found that chemicals active for each of three different RCAS were more likely to be active in at least one orthogonal Tier 2 end point. We similarly found that chemicals that were active and selective in each RCAS were more likely to be active and selective for at least one orthogonal end point in Tier 2 (Figure 3C, right), although odds ratios varied sizably across RCAS (Excel Table S4).

Evaluation of Inactive Chemicals in Tier 1 for Tier 2 PODs

One potential concern for the use of transcriptomics as a Tier 1 screen is the potential for false negative outcomes, i.e., chemicals labeled as inactive in Tier 1 that show bioactivity in Tier 2 end points. To investigate the potential for missed activity in Tier 1 tests, we identified chemicals with false negative outcomes as those that were negative in each RCAS (i.e., neither active nor selective) and either active only (i.e., active but not selective) or active and selective in at least one orthogonal ToxCast end point. Minimum PODs across orthogonal ToxCast end points (POD_{ortho}) were used to compare Tier 2 potency estimates between Tier 1 outcomes.

For chemicals negative in Tier 1 but active in Tier 2 (Figure 4A, red boxes), POD_{ortho} values were significantly higher than those for chemicals that were active in Tier 1 and Tier 2 in two of three RCAS tested (Figure 4A, blue boxes). For chemicals negative in Tier 1 but active and selective in Tier 2 (Figure 4B, red boxes), POD_{ortho} values were significantly higher than values for chemicals that were active in Tier 1 and both active and selective in Tier 2 for all RCAS tested (Figure 4B, blue boxes). These trends suggest that false negative outcomes in Tier 1 corresponded to higher bioactive concentrations in Tier 2 in comparison with chemicals that were active in both tiers. Higher bioactive concentrations often coincide with a burst of bioactivity in assays spanning many pathways and can be indicative of nonselective effects, such as cell stress or nonspecific receptor binding.[30] Although uncertainty remains as to why some chemicals appear as false negatives in Tier 1 tests, further investigation into overall bioactivity and/or cytotoxicity for these chemicals could identify sources of reduced sensitivity in Tier 1.

We additionally compared POD_{ortho} values based on Tier 1 outcome (i.e., negative, active but not selective, or active and selective). For multiple RCAS, false negative chemicals either trended toward higher POD_{ortho} values when active but not selective in Tier 2 (Figure S9A, red boxes) or had significantly higher POD_{ortho} values when active and selective in Tier 2 (Figure S9B, red boxes) vs. chemicals that were active but not selective or active and selective in Tier 1.

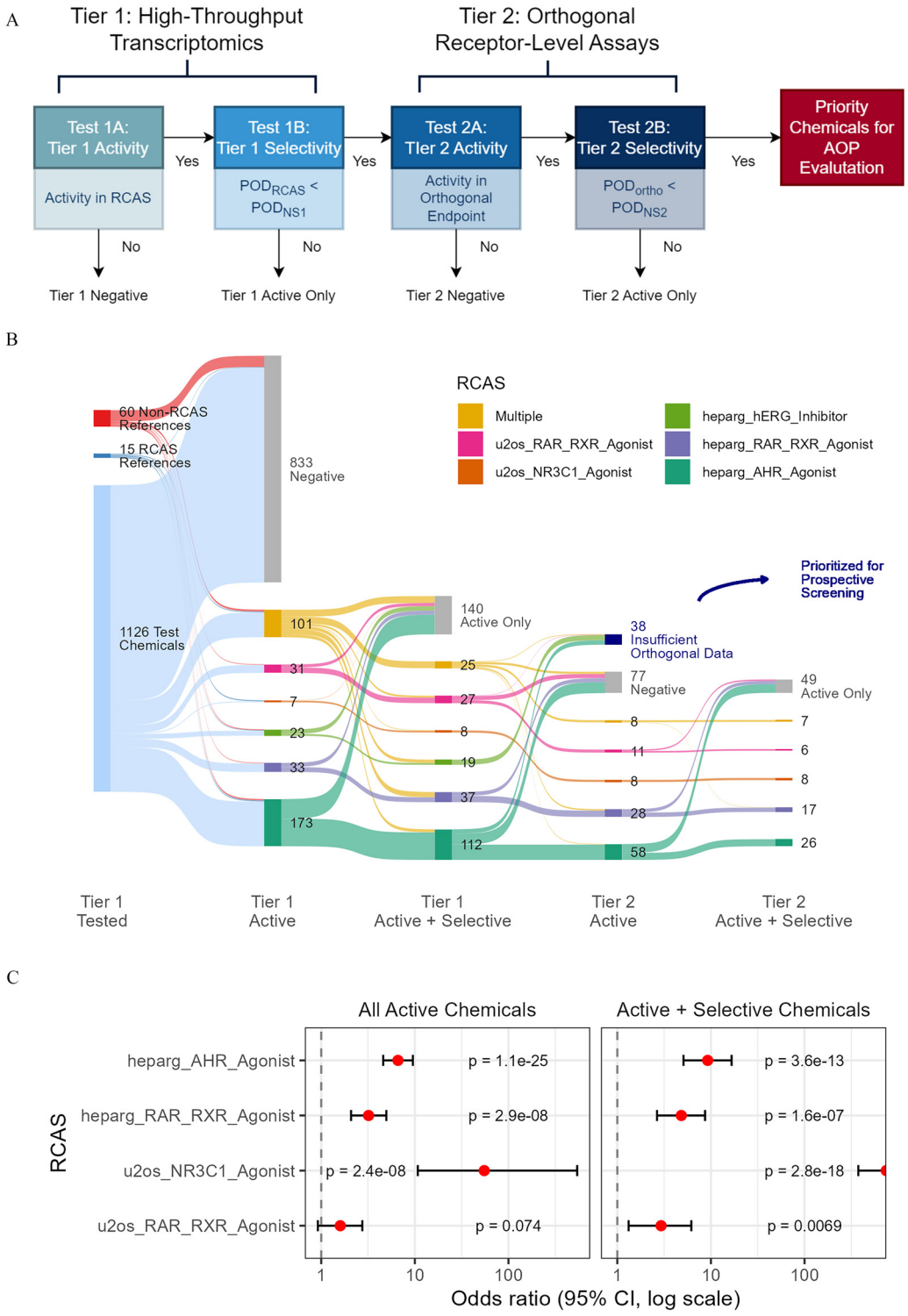

Figure 3. Tiered assessment framework for prioritizing selective molecular target perturbagens. (A) Decision framework used to prioritize chemicals across tiered transcriptomic and high-throughput screening assays. (B) Summary of tiered framework outcomes when applied to chemicals previously screened in HTTr. Chemicals passing successive steps were advanced to the next stage of prioritization, and final candidates were determined as candidate perturbagens for each RCAS. Chemical totals represent those chemicals passing tests for the individual or multiple RCAS at each step. (C) Association tests between Tier 1 and Tier 2 framework outcomes. Odds ratios that a Tier 1-active chemical would also be active in at least one Tier 2 orthogonal ToxCast end point (left) and that a Tier 1-active and selective chemical would also be active and selective in at least one Tier 2 orthogonal end point (right) are shown for each RCAS. Points extending beyond the *x*-axis range indicate an infinite odds ratio, i.e., no odds that a Tier 1-active and selective chemical was not active and selective in Tier 2. See Excel Table S8 for all underlying data. Note: BMC, benchmark concentration; HTTr, high-throughput transcriptomics; POD_{RCAS}, RCAS-derived point of departure; POD_{NS1}, Tier 1 nonselective point of departure; POD_{NS2}, Tier 2 nonselective point of departure; POD_{ortho}, Tier 2 orthogonal assay point of departure; RCAS, reference class–associated signature. $^{*}p \leq 0.05$ via Fisher's exact test for associations between Tier 1 and Tier 2 outcomes described above.

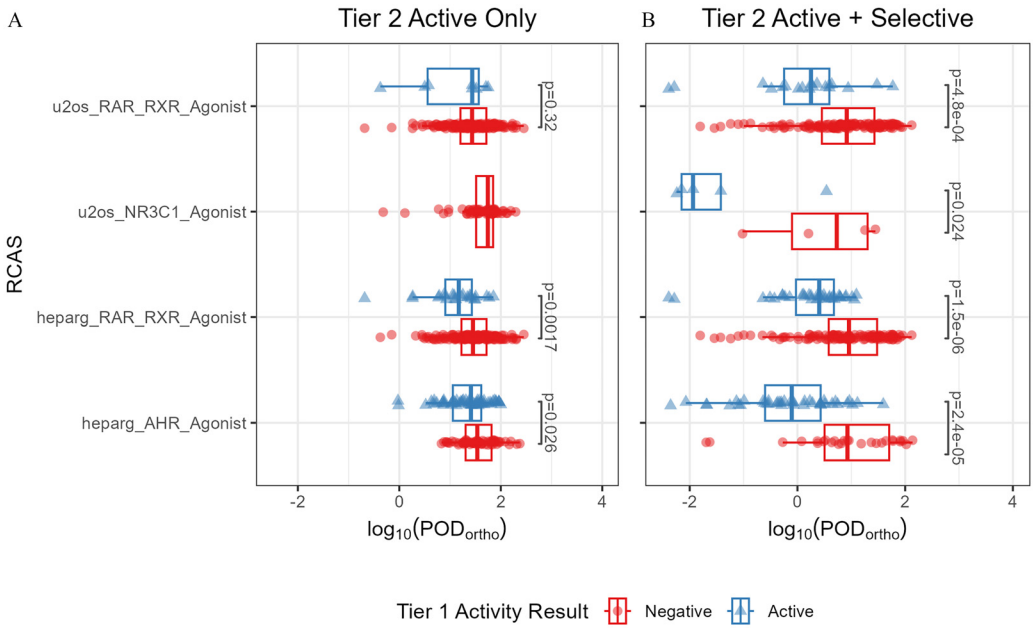

Figure 4. Comparison of Tier 2 potency estimates for chemicals grouped by Tier 1 bioactivity. Chemicals demonstrating either an active only outcome (A) or an active and selective outcome (B) in any Tier 2 orthogonal end point for each RCAS are grouped by their Tier 1 activity test outcome, regardless of selectivity. Statistical comparisons of POD_{ortho} estimates (*x*-axis) by Tier 1 activity outcome for each RCAS were conducted using Wilcoxon rank sum tests with Benjamini-Hochberg correction. Groupings are not shown when no chemicals matched the specified Tier 1 and Tier 2 outcomes (e.g., Tier 1-active only and Tier 2-active only for u2os_NR3C1_Agonist RCAS). All chemicals that were either active only or active and selective in Tier 2 orthogonal assays are displayed. Boxes represent 25th percentile, median, and 75th percentile values of POD_{ortho} estimates; whiskers represent 1.5× IQR of POD_{ortho} estimates. See Excel Table S9 for all underlying data. Note: BMC, benchmark concentration; IQR, interquartile range; POD_{ortho}, Tier 2 orthogonal assay point of departure; POD_{RCAS}, RCAS-derived point of departure; RCAS, reference class–associated signature.

Target-Specific Potency Estimates in Comparison with Overall Transcriptomic PODs

We next compared POD_{RCAS} estimates to PODs representing the HTTr-derived biological pathway altering concentration (POD_{BPAC1}) for chemicals prioritized in Tier 1. Of the chemicals with either active only or active and selective Tier 1 outcomes, we found that 95.4% of chemicals had a POD_{RCAS} within 1-log of the POD_{BPAC1} (Figure 5A). Pearson correlations between POD_{RCAS} and POD_{BPAC1} values were significant for all groups except chemicals selective for the "u2os_NR3C1_Agonist" RCAS, which had fewer chemicals for comparison.

We additionally compared POD_{RCAS} estimates to orthogonal Tier 2 potency estimates to evaluate the relative sensitivity of targeted estimates across tiers. Of chemicals with active only or active and selective outcomes in at least one orthogonal end point, 15.1% of chemicals had a POD_{ortho} 1-log lower than the POD_{RCAS} or more (Figure 5B). These differences in POD estimates were evident in a greater proportion of chemicals with an active and selective Tier 2 outcome (24.6% of chemicals).

Comparison of POD_{RCAS} values to PODs derived from all active Tier 2 end points in ToxCast (POD_{BPAC2}) resulted in weaker Pearson correlations vs. comparisons with POD_{BPAC1} values above (Figure S10), and 62.2% of chemicals had a POD_{RCAS} within 1-log of the POD_{BPAC2} or less. The assays in ToxCast encompass a variety of cell lines, detection technologies, concentration ranges, and biological functions, however,[5] so resulting POD_{BPAC2} estimates can be highly variable in comparison with the cell line-specific PODs derived from HTTr.

Structural Similarities and Known Activity of Framework-Identified Candidates

Of the 1,126 nonreference chemicals tested using the tiered framework above, 57 met criteria across Tiers 1 and 2 for selective AHR agonism (26 chemicals), GR agonism (8), or RAR/RXR agonism (23 via both HepaRG and U-2 OS RCAS). Sensitivity toward targeted assays varied by molecular target, because POD_{ortho} estimates for GR agonists (Figure 6B) were often lower than those for AHR and RAR/RXR agonists (Figure 6A,C,D). Dissimilarity in the ToxCast assay technologies' metabolic-capacity biological functions profiled may also account for differences in Tier 2 potency estimates. Further exploration would be needed to investigate differences in assay sensitivity.

Individual candidate perturbagens demonstrated similarities between their chemical structures and the structures of known classes of nuclear receptor modulators. Within candidate AHR agonists, 11 chemicals contained oxygenated polycyclic aromatic structures, including 2-aminoanthraquinone, 1,4-diaminoanthraquinone, and 1-naphthol (Figure 6A), indicating similarities to known AHR-mediated toxicants, such as oxygenated polycyclic aromatic hydrocarbons and anthraquinone dyes.[44–47] All four candidate GR agonists shared steroidal features and are either marketed synthetic steroids or have been reported to demonstrate minor glucocorticoid agonism[48] and immunosuppression[49] *in vitro* (Figure 6B). Several triazole and imidazole fungicides were identified as RAR/RXR agonists using RCASs from either cell line, including triflumizole, flusilazole, imazalil, prochloraz, and etoxazole, as well as several organochlorine pesticides, including aldrin, dieldrin, endrin, and lindane (Figure 6C,D). Chemicals within both structural groups have been linked to developmental skeletal defects[50,51] and neurotoxicity[52,53] in both *in vitro* and guideline rodent studies and are consistent with health effects associated with RAR disruption as compiled from *in vivo* and *in vitro* studies.[51]

Prospective Screening of Tier 1-Predicted Chemicals in HTS Assays

Tier 1 transcriptomics predicted 38 chemicals as active and selective perturbagens that did not have sufficient Tier 2 data for

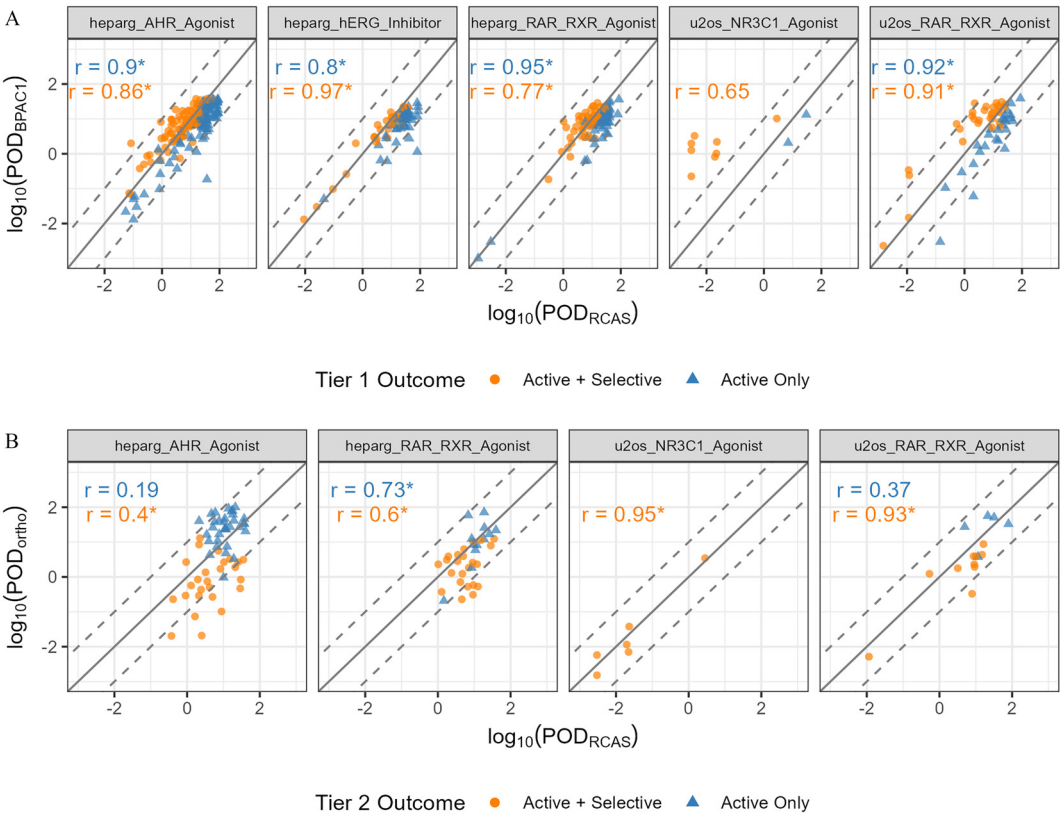

Figure 5. Comparison of RCAS-derived PODs to overall HTTr-derived PODs and orthogonal HTS-derived PODs. (A) Comparison of RCAS-derived potency estimates with POD_{BPAC1} values, determined as the fifth percentile of active signature BMCs estimated from publicly sourced signatures. Chemicals with either an active only or an active and selective Tier 1 outcome are displayed. (B) Comparison with POD_{ortho} values, determined as the minimum ACC from active orthogonal ToxCast end points. Only chemicals with an active and a selective Tier 1 outcome and either an active only or an active and selective Tier 2 outcome are displayed. Solid lines reflect equal PODs between estimates, and dashed lines reflect a difference of 1-log. Pearson correlation coefficients between PODs (r) are displayed by RCAS and tier outcome. See Excel Table S10 for all underlying data. Note: ACC, activity concentration at cutoff; BMC, benchmark concentration; BPAC, biological pathway altering concentration; HTTr, high-throughput transcriptomics; POD, point of departure; POD_{ortho}: Tier 2 orthogonal assay point of departure; POD_{RCAS}, RCAS-derived point of departure; RCAS, reference class–associated signature. $^{*}p \leq 0.05$ via Student's t-test for Pearson correlation coefficients between POD estimates described for each panel.

comparison (Excel Table S5). To test whether Tier 1 predictions could similarly prioritize these chemicals in comparison with chemicals with existing Tier 2 data, we profiled a subset of the 38 chemicals above using one of five prospective end points for transcription factor transactivation or ion channel inhibition (Table 3). Although chemicals chosen for GR and RAR/RXR end points were previously profiled in orthogonal ToxCast end points, the results here represent additional assay technologies not used in our retrospective analyses above and are included for comparison.

Of the 28 chemicals prospectively profiled for AHR and hERG modulation, 33.3% and 43.8% of chemicals were concentration responsive for AHR and hERG end points, respectively (Figure 7A). These hit rates are lower than those for chemicals that were active and selective in each RCAS and had existing orthogonal data in ToxCast (Figure S11). Chemicals with existing orthogonal data were tested in multiple targeted assays per target, however, so it is expected that these hit rates would be higher than those for chemicals prospectively profiled in only one assay. Hit rates were higher for chemicals tested for RARG and GR modulation (75% and 50%, respectively); caution should be used when interpreting these results due to smaller sample sizes in comparison with AHR and hERG.

We additionally compared POD_{RCAS} estimates for prospectively screened chemicals to PODs derived from the prospective end point assays ($POD_{Prospective}$) (Figure 7B). POD_{RCAS} values for chemicals active in any end point were significantly correlated with $POD_{Prospective}$ values ($p = 9.52 \times 10^{-6}$), and the root mean square deviation between PODs was approximately 1 order of magnitude. POD_{RCAS} values were within 1-log of the $POD_{Prospective}$ or less for 21 of 31 chemicals (67.7%) across all targets. The various detection technologies used for these prospective end points could contribute to differences in individual potency estimates vs. RCAS, however, and investigation with a larger chemical screen would be needed to fully evaluate differences in assay sensitivity.

Discussion

We investigated whether a tiered, NAM-based decision framework could identify chemicals selectively acting on key molecular targets by integrating HTTr and targeted HTS data streams. Our development and use of RCAS allowed for translation of broad profiling assay data to predictions of molecular target perturbation. Further combining RCAS predictions with potency estimates from HTS assays allowed for orthogonal validation and subsequent prioritization of candidate perturbagens. The approach used here was intended as an additional method to complement current derivations of transcriptomic PODs by providing mechanistic insight specific to the cell line tested.

Utility of Broad-Profiling NAMs for Identifying Mechanistic Behavior

Broad-profiling NAMs such as transcriptomics have been increasingly used to profile chemical interactions with molecular

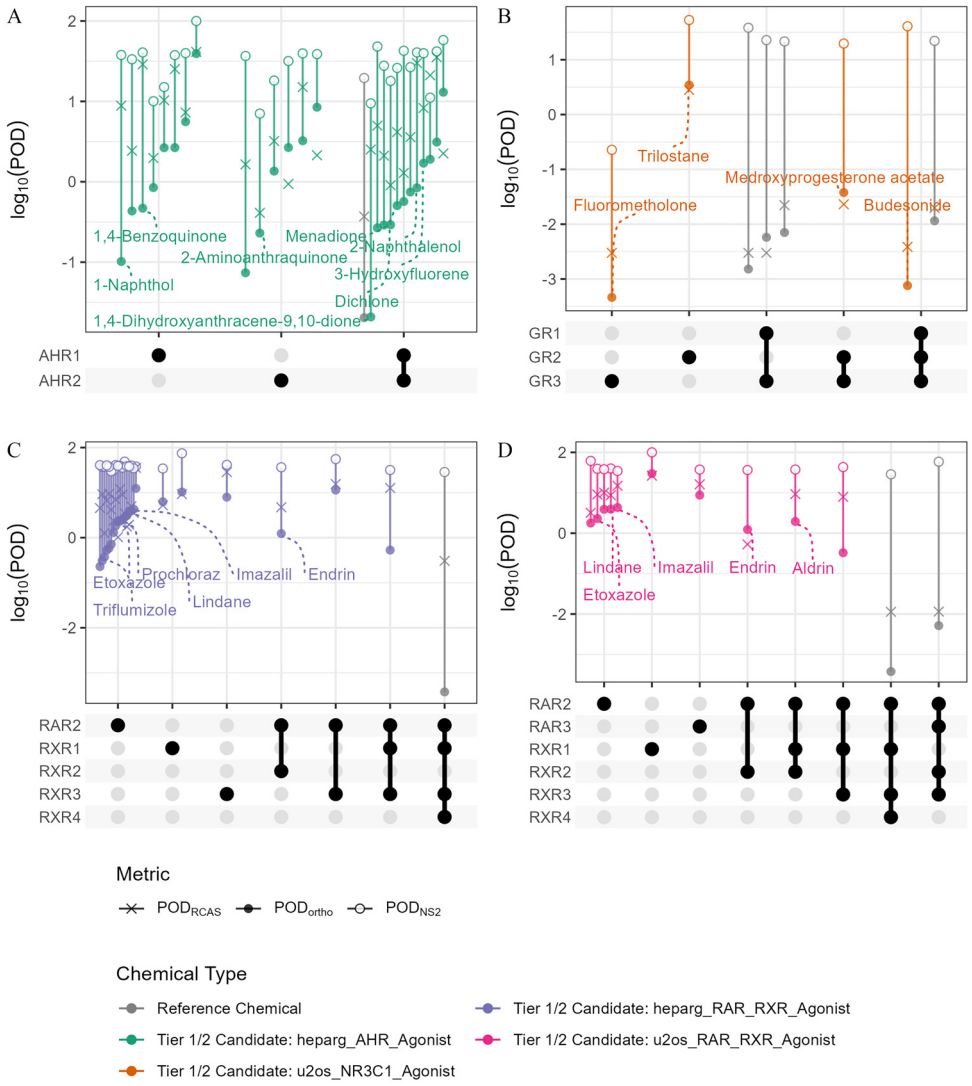

Figure 6. Candidate target perturbagens identified by tiered framework. Comparison of Tier one-half potency values for candidate nuclear receptor agonists based on RCAS for AHR in HepaRG (A), GR in U-2 OS (B), RAR/RXR in HepaRG (C), and RAR/RXR in U-2 OS (D). Chemicals displayed represent test chemicals passing all four framework steps in at least one RCAS and orthogonal end point. POD_{ortho} values represent the minimum ACC across active end points for each chemical. Chemicals are grouped by the combination of orthogonal end points in which a chemical was active and selective (*x*-axis, see Table 1 for end point names and assay details). Only end points with at least one passing chemical are shown. Nonreference chemicals with similar structural features within each panel (e.g., polycyclic aromatic structures) are annotated where possible. See Excel Table S11 for all underlying data. Note: ACC, activity concentration at cutoff; AHR, aryl hydrocarbon receptor; GR, glucocorticoid receptor; POD_{NS1}, Tier 1 nonselective point of departure; POD_{NS2}, Tier 2 nonselective point of departure; POD_{ortho}, Tier 2 orthogonal assay point of departure; POD_{RCAS}, RCAS-derived point of departure; RAR, retinoic acid receptor; RCAS, reference class–associated signature; RXR, retinoid X receptor.

targets relating to a variety of adverse outcomes.[54] Signature-based analyses have often been used to mine transcriptomic data for effects associated with cellular pathway modulation,[8,55] disease progression,[10] or other human health effects.[56] Many publicly sourced signatures, although useful for biological interrogation, may not be well-suited for detecting molecular targets for chemical hazard identification. First, many signatures reflect transcriptomic responses that are cell- or tissue-specific and may not reflect changes observed in the cell lines used for *in vitro* screening. Second, some signatures reflect broadly defined biological processes with low confidence for detecting individual chemical–target interactions.[57] Finally, various signature databases were developed based on other cellular responses such as cell signaling and may not be reflective of changes in gene expression.[58]

Our use of RCAS indicates that data-driven signatures for specific molecular targets and cellular contexts can be used to screen for targeted effects upon chemical exposure. Such analyses demonstrate the utility of HTTr for NAM-based screening beyond deriving nontargeted PODs as used previously.[59,60] Other studies have also linked chemicals to relevant targets by comparing gene expression profiles between DNA damage toxicants and nontoxicants,[61] by using transgenic *in vivo* models,[62] or via machine learning models[16] and coexpression networks.[63–65] Although these studies did not directly incorporate potency criteria for selecting signatures, such methods can detect perturbations that were not identified here. Combining signatures generated from multiple methods into a battery of targeted signatures could complement RCAS and expand the number of targets screened with confidence in HTTr.

From 14 initial molecular targets, only 5 RCASs were successfully identified and used for bioactivity screening. Molecular targets were required to have multiple well-annotated reference chemicals in RefChemDB that were also bioactive in HTTr. These criteria limit the possible targets to be mined from HTTr

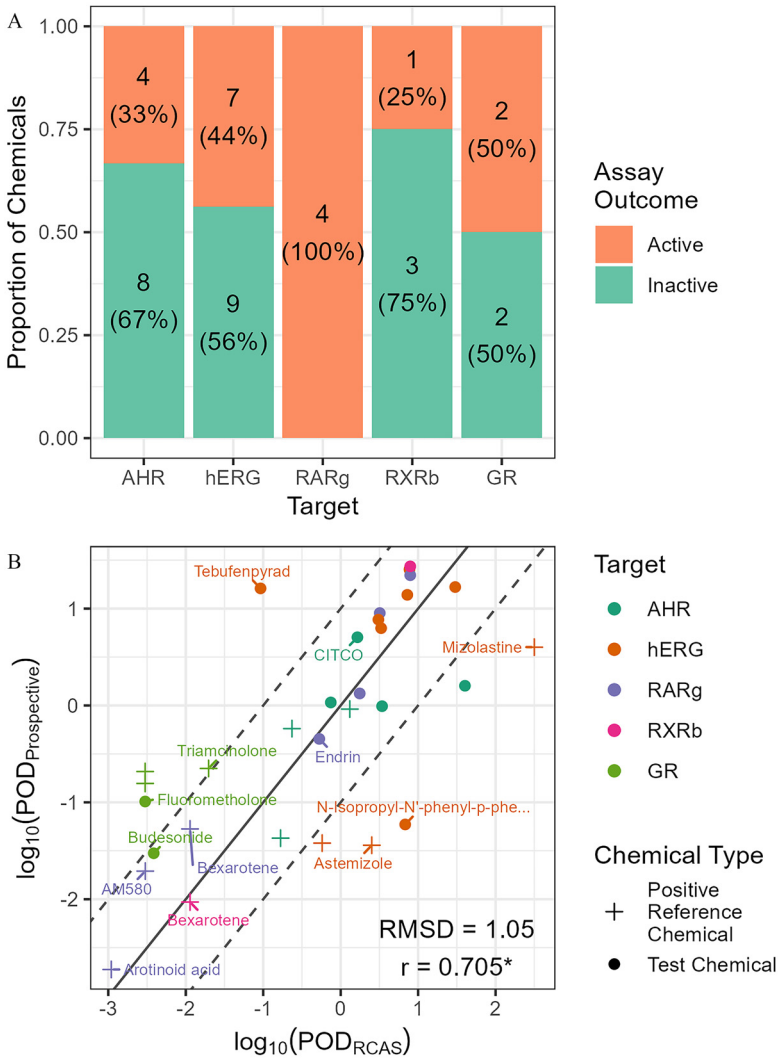

Figure 7. Prospective profiling of Tier 1-predicted candidates for molecular target modulation in receptor-level assays. (A) Activity classifications for chemicals tested in prospective orthogonal assays. Numbers indicate the total number of unique chemicals with each prospective end point and activity classification (B) Comparison of POD_{RCAS} estimates (*x*-axis) and $POD_{Prospective}$ estimates (*y*-axis) for active chemicals. The solid line reflects equal POD estimates, and dashed lines reflect a difference of 0.5-log. See Excel Table S12 for all underlying data. Note: POD, point of departure; POD_{ortho}, Tier 2 prospective assay point of departure; POD_{RCAS}, RCAS-derived point of departure; RCAS, reference class–associated signature; RMSD, root mean square deviation between PODs (shown in log-units). $^{*}p \leq 0.05$ via Student's *t*-test for Pearson correlation coefficients between POD_{RCAS} and $POD_{Prospective}$ estimates.

when existing datasets were not explicitly designed for mechanistic inference. An important aspect is that it should not be expected that all molecular targets produce a uniquely sensitive gene-expression signature in each cell line tested. Nonnuclear receptor targets such as G-coupled protein receptors induce changes in transcription via a complex network of cell signaling pathways, so any resulting changes in cell transcription could span many pathways and may only be detected at higher concentrations. Some molecular targets may also be more responsive in one cell line vs. another (e.g., GR agonism in U-2 OS vs. HepaRG), and the 24-h exposure duration used here could have additionally impacted the sensitivity of gene expression toward chemical perturbation. It is therefore plausible that the molecular targets tested here may demonstrate unique effects on gene expression in one or more other cell lines not yet profiled in HTTr. Optimal bioactivity coverage will likely involve screening across multiple cell lines,[17,66] but our analyses suggest that the methodology for generating RCAS can augment existing transcriptomic screening data by detecting selectively acting chemicals in the cell line(s) tested.

RCASs were also not identified for some targets where bioactivity was expected and reference chemicals were available (e.g., PPAR agonism in HepaRG cells). This lack of sensitivity could have been because of differences in cellular metabolism in comparison with other *in vitro* and *ex vivo* models. Although HepaRG cells retain key phase I and II metabolic enzymes in comparison with other immortalized cell lines,[67] such enzymes may be differentially expressed vs. primary human hepatocytes and liver biopsies,[68,69] and exposure duration could also affect the bioavailability of reactive metabolites.[70,71] We also found that PPARA agonists induced expected bioactivity in HepaRG cells for individual genes, but this small gene set did not meet criteria for signature-based profiling. This finding raises the question of whether signature-based analyses are equally sensitive toward all molecular targets. Expert-curated biomarkers have alternatively detected effects linked to adverse outcomes such as liver injury or tumorigenesis.[72,73] The method proposed for generating RCASs here should ultimately be employed within a larger toolbox of methods for signature-based profiling to provide sensitivity across a range of molecular targets for each cell line used for testing.

Tiered Decision Framework for Distinguishing Selectively Acting Perturbagens

The tiered framework presented here represented a proof-of-concept for integrating broad-profiling and targeted NAMs into a single workflow for chemical bioactivity screening. Others have shown that combining bioactivity estimates from multiple NAMs can improve confidence in Next Generation Risk Assessment scenarios including for consumer goods,[11,12] pharmaceuticals, and other chemical classes.[74] By incorporating HTTr as an initial screen for targeted bioactivity, we found that transcriptomics could improve the efficiency of chemical screening by predicting chemical–target interactions based on targeted gene sets, thereby reducing both the number of chemicals to be tested in targeted assays and the number of assays to test per chemical. Our transcriptomic testing strategy additionally identified several chemicals as potential targeted perturbagens for which few or no orthogonal measurements were available in ToxCast, and prospective screening of these chemicals revealed additional insight into possible mechanisms that might not have been profiled otherwise.

Our incorporation of selectivity criteria provided additional evidence that a chemical may perturb a target of interest as opposed to assay responses that cannot be distinguished from generalized cell stress or cytotoxicity.[30] The nonselective potency metric employed here provided an estimate of the central tendency of transcriptional responses rather than the lower bound, and this strategy allows for a higher sensitivity vs. BPAC-based comparisons that could reduce the incidence of false negative results. Active and selective chemicals from this framework could be candidates for fit-for-purpose assays designed to closely approximate human health effects,[6,75] because the targets explored here can be mapped to adverse outcomes such as developmental defects,[52] cardiotoxicity,[76] and tumorigenesis and immune function.[77,78] An important consideration is that chemicals with an active but nonselective outcome in Tiers 1 or 2 using the framework above should not necessarily be interpreted as nonhazardous chemicals. Such chemicals may act by multiple specific or nonspecific mechanisms not captured here, and any resulting alterations to the transcriptome could still result in adverse outcomes. These chemicals may merit further assessment for other putative hazards or estimation of a quantitative POD in absence of a defined mechanism.[6]

Sensitivity of NAM-Based Framework within Decision Contexts

Although the framework presented here was tailored toward chemical screening and prioritization, tiered integration of multiple NAMs may not be suitable for all decision-making scenarios. For example, performing transcriptomics and RCAS-like testing prior to targeted NAMs on data-poor chemicals in which no prior NAMs data are available may yield information on selected targets, but it would not necessarily inform stakeholders whether all relevant targets are covered. It would also be uncertain whether more cell lines are needed for comprehensive coverage in Tier 1. In this scenario, the use of a targeted NAM panel in tandem with transcriptomics would more likely provide comprehensive target coverage when testing data-poor chemicals.[79] Different assays may also be prioritized first within a tiered framework for maximizing positive or negative predictive power,[80] or assays may be run in tandem for estimation of systemic PODs that are protective of human health.[12,81]

Chemical structure–based predictions could be added to this framework to improve screening efforts or help translate to other decision contexts. *In silico* NAMs could serve as a "Tier 0" screen by providing structural alerts or predictions of human health effects such as endocrine disruption.[82–84] Additional incorporation of predicted internal concentrations via toxicokinetic models have been demonstrated and applied to NAM-based assessments via bioactivity–exposure ratios.[11,12,29] These estimates could further prioritize chemicals that are likely to exceed points of departure in various exposure scenarios.

Conclusions

Our investigation suggests that integrating transcriptomic and targeted NAMs into a framework can streamline the prioritization of chemicals for further toxicity testing with respect to specific molecular targets. The transcriptomic signatures presented here allowed for mechanistic interpretation in a specific cellular context in addition to quantitative POD estimation and thereby expanded the utility of HTTr screening with few to no added resources. The framework presented here provided a proof-of-concept for integrating *in vitro* NAMs in a tiered manner to support hazard screening efforts. Although an expanded evaluation of HTTr-predicted perturbagens and incorporation of *in silico* predictions will further build confidence in and expand the utility of tiered testing strategies, the current approach described here is readily applicable to the needs of large-scale toxicity testing.

Acknowledgments

The authors would like to acknowledge Madison Feshuk and Kurt Dunham for assistance with processing prospective profiling data.

The US EPA Office of Research and Development (ORD) provided funding for this research. This project was supported in part by an appointment to the Research Participation Program at the Center for Computational Toxicology and Exposure, US EPA, administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the US Department of Energy and US EPA.

The views expressed in this manuscript are solely those of the authors and do not represent the policies of the US Environmental Protection Agency. Mention of trade names or commercial products should not be interpreted as an endorsement by the US Environmental Protection Agency.

References

1. US EPA (US Environmental Protection Agency). About the TSCA Chemical Substance Inventory. https://www.epa.gov/tsca-inventory/about-tsca-chemical-substance-inventory [accessed 13 February 2024].
2. Hartung T. 2009. Toxicology for the twenty-first century. Nature 460(7252):208–212, PMID: 19587762, https://doi.org/10.1038/460208a.
3. NRC (US National Research Council). 2007. *Toxicity Testing in the 21st Century: A Vision and a Strategy*. Washington, DC: The National Academies Press, 216.
4. Krewski D, Andersen ME, Tyshenko MG, Krishnan K, Hartung T, Boekelheide K, et al. 2020. Toxicity testing in the 21st century: progress in the past decade and future perspectives. Arch Toxicol 94(1):1–58, https://doi.org/10.1007/s00204-019-02613-4.
5. Richard AM, Judson RS, Houck KA, Grulke CM, Volarath P, Thillainadarajah I, et al. 2016. ToxCast chemical landscape: paving the road to 21st century toxicology. Chem Res Toxicol 29(8):1225–1251, PMID: 27367298, https://doi.org/10.1021/acs.chemrestox.6b00135.
6. Thomas RS, Bahadori T, Buckley TJ, Cowden J, Deisenroth C, Dionisio KL, et al. 2019. The next generation blueprint of computational toxicology at the U.S. Environmental Protection Agency. Toxicol Sci 169(2):317–332, PMID: 30835285, https://doi.org/10.1093/toxsci/kfz058.
7. Brockmeier EK, Hodges G, Hutchinson TH, Butler E, Hecker M, Tollefsen KE, et al. 2017. The role of omics in the application of adverse outcome pathways for chemical risk assessment. Toxicol Sci 158(2):252–262, PMID: 28525648, https://doi.org/10.1093/toxsci/kfx097.
8. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. 2015. The molecular signatures database hallmark gene set collection. Cell Syst 1(6):417–425, PMID: 26771021, https://doi.org/10.1016/j.cels.2015.12.004.

9. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. 2017. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res 45(D1):D353–D361, PMID: 27899662, https://doi.org/10.1093/nar/gkw1092.
10. Piñero J, Ramírez-Anguita JM, Saüch-Pitarch J, Ronzano F, Centeno E, Sanz F, et al. 2020. The DisGeNET knowledge platform for disease genomics: 2019 update. Nucleic Acids Res 48(D1):D845–D855, PMID: 31680165, https://doi.org/10.1093/nar/gkz1021.
11. Baltazar MT, Cable S, Carmichael PL, Cubberley R, Cull T, Delagrange M, et al. 2020. A next-generation risk assessment case study for coumarin in cosmetic products. Toxicol Sci 176(1):236–252, PMID: 32275751, https://doi.org/10.1093/toxsci/kfaa048.
12. Middleton AM, Reynolds J, Cable S, Baltazar MT, Li H, Bevan S, et al. 2022. Are non-animal systemic safety assessments protective? A toolbox and workflow. Toxicol Sci 189(1):124–147, PMID: 35822611, https://doi.org/10.1093/toxsci/kfac068.
13. Rajagopal R, Baltazar MT, Carmichael PL, Dent MP, Head J, Li H, et al. 2022. Beyond AOPs: a mechanistic evaluation of NAMs in DART testing. Front Toxicol 4:838466, PMID: 35295212, https://doi.org/10.3389/ftox.2022.838466.
14. Zobl W, Bitsch A, Blum J, Boei JJWA, Capinha L, Carta G, et al. 2024. Protectiveness of NAM-based hazard assessment–which testing scope is required? ALTEX 41(2):302–319, https://doi.org/10.14573/altex.2309081.
15. Judson RS, Thomas RS, Baker N, Simha A, Howey XM, Marable C, et al. 2019. Workflow for defining reference chemicals for assessing performance of in vitro assays. ALTEX 36(2):261–276, PMID: 30570668, https://doi.org/10.14573/altex.1809281.
16. Bundy JL, Judson R, Williams AJ, Grulke C, Shah I, Everett LJ. 2022. Predicting molecular initiating events using chemical target annotations and gene expression. BioData Min 15(1):7, PMID: 35246223, https://doi.org/10.1186/s13040-022-00292-z.
17. Bundy JL, Everett LJ, Rogers JD, Nyffeler J, Byrd G, Culbreth M, et al. 2024. High-throughput transcriptomics screen of ToxCast chemicals in U-2 OS cells. Toxicol Appl Pharmacol 491:117073, PMID: 39159848, https://doi.org/10.1016/j.taap.2024.117073.
18. Harrill JA, Everett LJ, Haggard DE, Sheffield T, Bundy JL, Willis CM, et al. 2021. High-throughput transcriptomics platform for screening environmental chemicals. Toxicol Sci 181(1):68–89, PMID: 33538836, https://doi.org/10.1093/toxsci/kfab009.
19. Yeakley JM, Shepard PJ, Goyena DE, Vansteenhouse HC, McComb JD, Seligmann BE. 2017. A trichostatin A expression signature identified by TempO-Seq targeted whole transcriptome profiling. PLoS One 12(5):e0178302, PMID: 28542535, https://doi.org/10.1371/journal.pone.0178302.
20. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15(12):550, PMID: 25516281, https://doi.org/10.1186/s13059-014-0550-8.
21. Sheffield T, Brown J, Davidson S, Friedman KP, Judson R. 2022. tcplfit2: an R-language general purpose concentration–response modeling package. Bioinformatics 38(4):1157–1158, PMID: 34791027, https://doi.org/10.1093/bioinformatics/btab779.
22. Thomas RS, Allen BC, Nong A, Yang L, Bermudez E, Clewell HJ, et al. 2007. A method to integrate benchmark dose estimates with genomic data to assess the functional effects of chemical exposure. Toxicol Sci 98(1):240–248, PMID: 17449896, https://doi.org/10.1093/toxsci/kfm092.
23. Harrill JA, Everett LJ, Haggard DE, Bundy JL, Willis CM, Shah I, et al. 2024. Exploring the effects of experimental parameters and transcriptomic data modeling approaches on in vitro point-of-departure estimates. Toxicology 501:153694, PMID: 38043774, https://doi.org/10.1016/j.tox.2023.153694.
24. Feshuk M, Kolaczkowski L, Dunham K, Davidson-Fritz SE, Carstens KE, Brown J, et al. 2023. The ToxCast pipeline: updates to curve-fitting approaches and database structure. Front Toxicol 5:1275980, PMID: 37808181, https://doi.org/10.3389/ftox.2023.1275980.
25. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, et al. 2009. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature 462(7269):108–112, PMID: 19847166, https://doi.org/10.1038/nature08460.
26. Hänzelmann S, Castelo R, Guinney J. 2013. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 14(1):7, PMID: 23323831, https://doi.org/10.1186/1471-2105-14-7.
27. Filer DL, Kothiya P, Woodrow Setzer R, Judson RS, Martin MT. 2017. Tcpl: the ToxCast pipeline for high-throughput screening data. Bioinformatics 33(4):618–620. 2, PMID: 27797781, https://doi.org/10.1093/bioinformatics/btw680.
28. US EPA. 2015. Exploring ToxCast Data. https://www.epa.gov/comptox-tools/exploring-toxcast-data [accessed 13 February 2024].
29. Paul Friedman K, Gagne M, Loo L-H, Karamertzanis P, Netzeva T, Sobanski T, et al. 2020. Utility of in vitro bioactivity as a lower bound estimate of in vivo adverse effect levels and in risk-based prioritization. Toxicol Sci 173(1):202–225, PMID: 31532525, https://doi.org/10.1093/toxsci/kfz201.
30. Judson R, Houck K, Martin M, Richard AM, Knudsen TB, Shah I, et al. 2016. Analysis of the effects of cell stress and cytotoxicity on in vitro assay activity across a diverse chemical and assay space. Toxicol Sci 152(2):323–339, https://doi.org/10.1093/toxsci/kfw092.
31. Wang Y, Nakajima T, Gonzalez FJ, Tanaka N. 2020. PPARs as metabolic regulators in the liver: lessons from liver-specific PPAR-null mice. Int J Mol Sci 21(6):2061, PMID: 32192216, https://doi.org/10.3390/ijms21062061.
32. Feige JN, Gelman L, Michalik L, Desvergne B, Wahli W. 2006. From molecular action to physiological outputs: peroxisome proliferator–activated receptors are nuclear receptors at the crossroads of key cellular functions. Prog Lipid Res 45(2):120–159, PMID: 16476485, https://doi.org/10.1016/j.plipres.2005.12.002.
33. Sun J, Yu L, Qu X, Huang T. 2023. The role of peroxisome proliferator–activated receptors in the tumor microenvironment, tumor cell metabolism, and anticancer therapy. Front Pharmacol 14:1184794, PMID: 37251321, https://doi.org/10.3389/fphar.2023.1184794.
34. Roche TE, Hiromasa Y. 2007. Pyruvate dehydrogenase kinase regulatory mechanisms and inhibition in treating diabetes, heart ischemia, and cancer. Cell Mol Life Sci 64(7–8):830–849, PMID: 17310282, https://doi.org/10.1007/s00018-007-6380-z.
35. Savas Ü, Machemer DEW, Hsu M-H, Gaynor P, Lasker JM, Tukey RH, et al. 2009. Opposing roles of peroxisome proliferator–activated receptor alpha and growth hormone in the regulation of CYP4A11 expression in a transgenic mouse model. J Biol Chem 284(24):16541–16552, PMID: 19366684, https://doi.org/10.1074/jbc.M902074200.
36. Nakagawa Y, Satoh A, Yabe S, Furusawa M, Tokushige N, Tezuka H, et al. 2014. Hepatic CREB3L3 controls whole-body energy homeostasis and improves obesity and diabetes. Endocrinology 155(12):4706–4719, PMID: 25233440, https://doi.org/10.1210/en.2014-1113.
37. Samson N, Bosoi CR, Roy C, Turcotte L, Tribouillard L, Mouchiroud M, et al. 2024. HSDL2 links nutritional cues to bile acid and cholesterol homeostasis. Sci Adv 10(22):eadk9681, PMID: 38820148, https://doi.org/10.1126/sciadv.adk9681.
38. Harrill JA, Everett LJ, Haggard DE, Sheffield T, Bundy J, Willis CM, et al. Signature analysis of high-throughput transcriptomics screening data for mechanistic inference and chemical grouping. Toxicol Sci 202(1):103–122, PMID: 39177380, https://doi.org/10.1093/toxsci/kfae108.
39. Sunaga Y, Inagaki N, Gonoi T, Yamada Y, Ishida H, Seino Y, et al. 1999. Troglitazone but not pioglitazone affects ATP-sensitive K+ channel activity. Eur J Pharmacol 381(1):71–76, PMID: 10528136, https://doi.org/10.1016/s0014-2999(99)00539-7.
40. Lu L, Reiter MJ, Xu Y, Chicco A, Greyson CR, Schwartz GG. 2008. Thiazolidinedione drugs block cardiac KATP channels and may increase propensity for ischaemic ventricular fibrillation in pigs. Diabetologia 51(4):675–685, PMID: 18251006, https://doi.org/10.1007/s00125-008-0924-0.
41. Rusinova R, Herold KF, Sanford RL, Greathouse DV, Hemmings HC Jr, Andersen OS. 2011. Thiazolidinedione insulin sensitizers alter lipid bilayer properties and voltage-dependent sodium channel function: implications for drug discovery. J Gen Physiol 138(2):249–270, PMID: 21788612, https://doi.org/10.1085/jgp.201010529.
42. Beaudoin JJ, Bezençon J, Sjöstedt N, Fallon JK, Brouwer KLR. 2020. Role of organic solute transporter alpha/beta in hepatotoxic bile acid transport and drug interactions. Toxicol Sci 176(1):34–35, PMID: 32294204, https://doi.org/10.1093/toxsci/kfaa052.
43. Bouwmeester MC, Tao Y, Proença S, van Steenbeek FG, Samsom R-A, Nijmeijer SM, et al. 2023. Drug metabolism of hepatocyte-like organoids and their applicability in in vitro toxicity testing. Molecules 28(2):621, PMID: 36677681, https://doi.org/10.3390/molecules28020621.
44. Goodale BC, La Du J, Tilton SC, Sullivan CM, Bisson WH, Waters KM, et al. 2015. Ligand-specific transcriptional mechanisms underlie aryl hydrocarbon receptor–mediated developmental toxicity of oxygenated PAHs. Toxicol Sci 147(2):397–411, PMID: 26141390, https://doi.org/10.1093/toxsci/kfv139.
45. McCarrick S, Cunha V, Zapletal O, Vondráček J, Dreij K. 2019. In vitro and in vivo genotoxicity of oxygenated polycyclic aromatic hydrocarbons. Environ Pollut 246:678–687, PMID: 30616058, https://doi.org/10.1016/j.envpol.2018.12.092.
46. Chou P-H, Matsui S, Misaki K, Matsuda T. 2007. Isolation and identification of xenobiotic aryl hydrocarbon receptor ligands in dyeing wastewater. Environ Sci Technol 41(2):652–657, PMID: 17310736, https://doi.org/10.1021/es061500g.
47. Clergé A, Le Goff J, Lopez C, Ledauphin J, Delépée R. 2019. Oxy-PAHs: occurrence in the environment and potential genotoxic/mutagenic risk assessment for human health. Crit Rev Toxicol 49(4):302–328, PMID: 31512557, https://doi.org/10.1080/10408444.2019.1605333.
48. Bamberger CM, Else T, Bamberger AM, Beil FU, Schulte HM. 1999. Dissociative glucocorticoid activity of medroxyprogesterone acetate in normal human lymphocytes. J Clin Endocrinol Metab 84(11):4055–4061, PMID: 10566649, https://doi.org/10.1210/jcem.84.11.6091.
49. Marinello W, Feng L, Allen TK. 2020. Progestins inhibit interleukin-1β-induced matrix metalloproteinase 1 and interleukin 8 expression via the glucocorticoid

receptor in primary human amnion mesenchymal cells. Front Physiol 11:900, PMID: 32792990, https://doi.org/10.3389/fphys.2020.00900.
50. US EPA. Health Effects Support Document for Aldrin/Dieldrin. EPA 822-R-03-001. https://www.epa.gov/sites/default/files/2014-09/documents/support_cc1_aldrin-dieldrin_healtheffects.pdf [accessed 9 April 2024].
51. Pierro JD, Ahir BK, Baker NC, Kleinstreuer NC, Xia M, Knudsen TB. 2022. Computational model for fetal skeletal defects potentially linked to disruption of retinoic acid signaling. Front Pharmacol 13:971296, PMID: 36172177, https://doi.org/10.3389/fphar.2022.971296.
52. Harrill JA, Freudenrich T, Wallace K, Ball K, Shafer TJ, Mundy WR. 2018. Testing for developmental neurotoxicity using a battery of in vitro assays for key cellular events in neurodevelopment. Toxicol Appl Pharmacol 354:24–39, PMID: 29626487, https://doi.org/10.1016/j.taap.2018.04.001.
53. Knudsen TB, Martin MT, Kavlock RJ, Judson RS, Dix DJ, Singh AV. 2009. Profiling the activity of environmental chemicals in prenatal developmental toxicity studies using the U.S. EPA's ToxRefDB. Reprod Toxicol 28(2):209–219, PMID: 19446433, https://doi.org/10.1016/j.reprotox.2009.03.016.
54. Campos B, Colbourne JK, Brown JB, Viant MR, Biales AD, Gallagher K, et al. 2018. How omics technologies can enhance chemical safety regulation: perspectives from academia, government, and industry. Environ Toxicol Chem 37(5):1252–1259, PMID: 29697867, https://doi.org/10.1002/etc.4079.
55. Gillespie M, Jassal B, Stephan R, Milacic M, Rothfels K, Senff-Ribeiro A, et al. 2022. The reactome pathway knowledgebase 2022. Nucleic Acids Res 50(D1): D687–D692, PMID: 34788843, https://doi.org/10.1093/nar/gkab1028.
56. Cho E, Buick JK, Williams A, Chen R, Li H-H, Corton JC, et al. 2019. Assessment of the performance of the TGx-DDI biomarker to detect DNA damage-inducing agents using quantitative RT-PCR in TK6 cells. Environ Mol Mutagen 60(2):122–133, PMID: 30488505, https://doi.org/10.1002/em.22257.
57. Wadi L, Meyer M, Weiser J, Stein LD, Reimand J. 2016. Impact of outdated gene annotations on pathway enrichment analysis. Nat Methods 13(9):705–706, PMID: 27575621, https://doi.org/10.1038/nmeth.3963.
58. Mubeen S, Hoyt CT, Gemünd A, Hofmann-Apitius M, Fröhlich H, Domingo-Fernández D. 2020. The impact of pathway database choice on statistical enrichment analysis and predictive modeling. Front Genet 11:436, PMID: 31824580, https://doi.org/10.3389/fgene.2019.01203.
59. Farmahin R, Williams A, Kuo B, Chepelev NL, Thomas RS, Barton-Maclaren TS, et al. 2017. Recommended approaches in the application of toxicogenomics to derive points of departure for chemical risk assessment. Arch Toxicol 91(5):2045–2065, PMID: 27928627, https://doi.org/10.1007/s00204-016-1886-5.
60. Reardon AJF, Farmahin R, Williams A, Meier MJ, Addicks GC, Yauk CL, et al. 2023. From vision toward best practices: evaluating in vitro transcriptomic points of departure for application in risk assessment using a uniform workflow. Front Toxicol 5:1194895, PMID: 37288009, https://doi.org/10.3389/ftox.2023.1194895.
61. Li H-H, Hyduke DR, Chen R, Heard P, Yauk CL, Aubrecht J, et al. 2015. Development of a toxicogenomics signature for genotoxicity using a dose-optimization and informatics strategy in human cells. Environ Mol Mutagen 56(6):505–519, PMID: 25733355, https://doi.org/10.1002/em.21941.
62. Andersen ME, Cruzan G, Black MB, Pendse SN, Dodd D, Bus JS, et al. 2017. Assessing molecular initiating events (MIEs), key events (KEs) and modulating factors (MFs) for styrene responses in mouse lungs using whole genome gene expression profiling following 1-day and multi-week exposures. Toxicol Appl Pharmacol 335:28–40, PMID: 28951217, https://doi.org/10.1016/j.taap.2017.09.015.
63. Wijaya LS, Gabor A, Pot IE, van de Have L, Saez-Rodriguez J, Stevens JL, et al. 2024. A network-based transcriptomic landscape of HepG2 cells uncovering causal gene-cytotoxicity interactions underlying drug-induced liver injury. Toxicol Sci 198(1):14–30, PMID: 38015832, https://doi.org/10.1093/toxsci/kfad121.
64. Krämer S, Busch W, Schüttler A. 2020. A self-organizing map of the fathead minnow liver transcriptome to identify consistent toxicogenomic patterns across chemical fingerprints. Environ Toxicol Chem 39(3):526–537, PMID: 31820487, https://doi.org/10.1002/etc.4646.
65. Sutherland JJ, Webster YW, Willy JA, Searfoss GH, Goldstein KM, Irizarry AR, et al. 2018. Toxicogenomic module associations with pathogenesis: a network-based approach to understanding drug toxicity. Pharmacogenomics J 18(3):377–390, PMID: 28440344, https://doi.org/10.1038/tpj.2017.17.
66. De Abrew KN, Shan YK, Wang X, Krailler JM, Kainkaryam RM, Lester CC, et al. 2019. Use of connectivity mapping to support read across: a deeper dive using data from 186 chemicals, 19 cell lines and 2 case studies. Toxicology 423:84–94, PMID: 31125584, https://doi.org/10.1016/j.tox.2019.05.008.
67. Nelson LJ, Morgan K, Treskes P, Samuel K, Henderson CJ, LeBled C, et al. 2017. Human hepatic HepaRG cells maintain an organotypic phenotype with high intrinsic CYP450 activity/metabolism and significantly outperform standard HepG2/C3A cells for pharmaceutical and therapeutic applications. Basic Clin Pharmacol Toxicol 120(1):30–37, PMID: 27285124, https://doi.org/10.1111/bcpt.12631.
68. Gerets HHJ, Tilmant K, Gerin B, Chanteux H, Depelchin BO, Dhalluin S, et al. 2012. Characterization of primary human hepatocytes, HepG2 cells, and HepaRG cells at the mRNA level and CYP activity in response to inducers and their predictivity for the detection of human hepatotoxins. Cell Biol Toxicol 28(2):69–87, PMID: 22258563, https://doi.org/10.1007/s10565-011-9208-4.
69. Hammer H, Schmidt F, Marx-Stoelting P, Pötz O, Braeuning A. 2021. Cross-species analysis of hepatic cytochrome P450 and transport protein expression. Arch Toxicol 95(1):117–133, PMID: 33150952, https://doi.org/10.1007/s00204-020-02939-4.
70. Malinowska JM, Palosaari T, Sund J, Carpi D, Weber RJM, Lloyd GR, et al. 2023. Derivation of metabolic point of departure using high-throughput in vitro metabolomics: investigating the importance of sampling time points on benchmark concentration values in the HepaRG cell line. Arch Toxicol 97(3):721–735, PMID: 36683062, https://doi.org/10.1007/s00204-022-03439-3.
71. Ramirez-Hincapie S, Birk B, Ternes P, Giri V, Zickgraf FM, Haake V, et al. 2023. Application of high throughput in vitro metabolomics for hepatotoxicity mode of action characterization and mechanistic-anchored point of departure derivation: a case study with nitrofurantoin. Arch Toxicol 97(11):2903–2917, PMID: 37665362, https://doi.org/10.1007/s00204-023-03572-7.
72. Podtelezhnikov AA, Monroe JJ, Aslamkhan AG, Pearson K, Qin C, Tamburino AM, et al. 2020. Quantitative transcriptional biomarkers of xenobiotic receptor activation in rat liver for the early assessment of drug safety liabilities. Toxicol Sci 175(1):98–112, PMID: 32119089, https://doi.org/10.1093/toxsci/kfaa026.
73. Glaab WE, Holder D, He YD, Bailey WJ, Gerhold DL, Beare C, et al. 2021. Universal toxicity gene signatures for early identification of drug-induced tissue injuries in rats. Toxicol Sci 181(2):148–159. PMC], PMID: 33837425, https://doi.org/10.1093/toxsci/kfab038.
74. Burbank M, Gautier F, Hewitt N, Detroyer A, Guillet-Revol L, Carron L, et al. 2023. Advancing the use of new approach methodologies for assessing teratogenicity: building a tiered approach. Reprod Toxicol 120:108454, PMID: 37543254, https://doi.org/10.1016/j.reprotox.2023.108454.
75. Andersen ME, McMullen PD, Phillips MB, Yoon M, Pendse SN, Clewell HJ, et al. 2019. Developing context appropriate toxicity testing approaches using new alternative methods (NAMs). ALTEX 36(4):523–534, PMID: 31664457, https://doi.org/10.14573/altex.1906261.
76. Ridder BJ, Leishman DJ, Bridgland-Taylor M, Samieegohar M, Han X, Wu WW, et al. 2020. A systematic strategy for estimating hERG block potency and its implications in a new cardiac safety paradigm. Toxicol Appl Pharmacol 394:114961, PMID: 32209365, https://doi.org/10.1016/j.taap.2020.114961.
77. Kolluri SK, Jin U-H, Safe S. 2017. Role of the aryl hydrocarbon receptor in carcinogenesis and potential as an anti-cancer drug target. Arch Toxicol 91(7):2497–2513, PMID: 28508231, https://doi.org/10.1007/s00204-017-1981-2.
78. Gutiérrez-Vázquez C, Quintana FJ. 2018. Regulation of the immune response by the aryl hydrocarbon receptor. Immunity 48(1):19–33, PMID: 29343438, https://doi.org/10.1016/j.immuni.2017.12.012.
79. US EPA. 2022. *The New Chemicals Collaborative Research Program: modernizing the Process and Bringing Innovative Science to Evaluate New Chemicals Under TSCA*. https://www.epa.gov/system/files/documents/2022-10/White_Paper_New%20Chemicals%20Collaborative%20Research%20Program_BOSC_Final_24Oct2022.pdf [accessed 5 August 2024].
80. Felli JC, Leishman DJ. 2020. A Bayesian approach to toxicological testing. J Pharmacol Toxicol Methods 105:106898, PMID: 32735877, https://doi.org/10.1016/j.vascn.2020.106898.
81. Browne P, Paul Friedman K, Boekelheide K, Thomas RS. 2024. Adverse effects in traditional and alternative toxicity tests. Regul Toxicol Pharmacol 148:105579, PMID: 38309424, https://doi.org/10.1016/j.yrtph.2024.105579.
82. Mansouri K, Abdelaziz A, Rybacka A, Roncaglioni A, Tropsha A, Varnek A, et al. 2016. CERAPP: collaborative estrogen receptor activity prediction project. Environ Health Perspect 124(7):1023–1033, PMID: 26908244, https://doi.org/10.1289/ehp.1510267.
83. Mansouri K, Kleinstreuer N, Abdelaziz AM, Alberga D, Alves VM, Andersson PL, et al. 2020. CoMPARA: collaborative modeling project for androgen receptor activity. Environ Health Perspect 128(2):027002, PMID: 32074470, https://doi.org/10.1289/EHP5580.
84. Cronin MTD, Bauer FJ, Bonnell M, Campos B, Ebbrell DJ, Firman JW, et al. 2022. A scheme to evaluate structural alerts to predict toxicity–assessing confidence by characterising uncertainties. Regul Toxicol Pharmacol 135:105249, PMID: 36041585, https://doi.org/10.1016/j.yrtph.2022.105249.