Sensitive image-based chromatin binding assays using inducible ERα to rapidly characterize estrogenic chemicals and mixtures

Adam T. Szafran, Maureen G. Mancini, Fabio Stossi, Michael A. Mancini
mancini@bcm.edu

Szafran et al., iScience 25, 105200
October 21, 2022 © 2022
https://doi.org/10.1016/j.isci.2022.105200
Sensitive image-based chromatin binding assays using inducible ERα to rapidly characterize estrogenic chemicals and mixtures

Adam T. Szafran,1 Maureen G. Mancini,1 Fabio Stossi,1,2 and Michael A. Mancini1,2,3,*

SUMMARY
The United States Environmental Protection Agency (EPA) has been pursuing new high throughput in vitro assays to characterize endocrine disrupting chemicals (EDCs) that interact with estrogen receptor signaling. We characterize two new PRL-HeLa cell models expressing either inducible C-terminal (iGFP-ER) or N-terminal (iER-GFP) tagged estrogen receptor-α (ERα) that allows direct visualization of chromatin binding. These models are an order of magnitude more sensitive, detecting 87 - 93% of very weak estrogens tested compared to only 27% by a previous PRL-HeLa variant and compares favorably to the 73% detected by an EPA-developed computational model using in vitro data. Importantly, the chromatin binding assays distinguished agonist- and antagonist-like phenotypes without activity specific assays. Finally, analysis of complex environmentally relevant chemical mixtures demonstrated how chromatin binding data can be used in risk assessment models to predict activity. These new assays should be a useful in vitro tool to screen for estrogenic activity.

INTRODUCTION
Currently, the rates of cancer, cardiovascular disease, chronic respiratory diseases, and diabetes mellitus are increasing faster than predicted by genetic drift. Although a number of factors are thought to play a role in this phenomenon, increased human exposure to environmental chemicals is likely a key contributor (Gore et al., 2015; Zoeller et al., 2012). Sources for these chemicals include industrial release, pesticide and herbicide use, environmental contamination because of accidental release, and more recently, increased use of pharmaceuticals and personal care products (Ezechias et al., 2016; Gore et al., 2015; Safe, 1995). One class of compounds of particular concern are endocrine disrupting compounds, any exogenous chemical or mixture of chemicals that interferes with normal hormone action (Gore et al., 2015). There is a substantial body of scientific work showing that endocrine disrupting chemicals (EDCs) are associated with a negative impact on animal and human health, interfering with normal female and male reproductive tract development, fertility, metabolism, and increasing the risk for endocrine related cancers (Gore et al., 2015; Safe, 1995).

Societal concern over environmental contaminants and EDCs in the US is reflected in the number of governmental regulations to understand potential exposure, the effects of that exposure, and how to minimize future exposure risks. The Environmental Protection Agency (EPA) established the Endocrine Disruptor Screening Program (EDSP) to fill a specific mandate to identify potential EDCs, their mechanism of activity and potential effects on human health using a two-tiered approach based on established in vitro and in vivo assays (Browne et al., 2015; US EPA, 2018a). More recently, a collaborative effort between the EPA, the National Toxicology Program (NTP) at the National Institute of Environmental Health Sciences (NIEHS), National Center for Advancing Translational Sciences (NCATS), and the Food and Drug Agency (FDA) established the Tox21 and ToxCast programs that emphasize the use of high throughput in vitro screening approaches to increase the exploration of the chemical landscape and identify the mechanisms of chemical biological activity. It was thought that the generation of highly diverse datasets would allow the EPA to increase the use of computational models to improve the prediction of chemical in vivo effects and prioritization of chemicals for further study.

The Tox21 program has examined over 10,000 chemicals using 75 cell-based assays (Gore et al., 2015; US EPA, 20118b). From this data, new computational models for EDC estrogenic activity have emerged.
Judson et al. have published an ab initio computational model based on 18 in vitro assays that was effective at identifying estrogenic, anti-estrogenic, and non-specific properties of chemicals (Judson et al., 2015). Using chemical sets defined by Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), the Organisation for Economic Co-operation and Development (OECD), or a literature search for qualifying in vivo studies, the integrated model has shown to have between 86 and 93% accuracy in classifying in vivo and in vitro estrogenic responses, suggesting the model could be useful as a potential alternative to EDSP tier 1 assays (Browne et al., 2015). More recently it has been shown that subsets of in vitro assays, with some including a few as four of the original assays, are able to correctly classify the estrogenic activity of a set of reference chemicals (Judson et al., 2017).

Included within the set of 18 in vitro assays is a chromatin binding assay based on a human cell line containing a stable, microscopically visible, multi-copy integration (promoter array) of the estrogen receptor-responsive prolactin promoter-enhancer unit (PRL-HeLa) (Judson et al., 2015). This cell line, GFP-ER:PRL-Hela, constitutively-expresses estrogen receptor-α (ERα) with an N-terminal green fluorescent protein (GFP) fusion and allows the direct quantitation of ERα DNA binding, recruitment of coregulators, epigenetic alterations, chromatin remodeling and transcriptional regulation in response to estrogenic ligands (Ashcroft et al., 2011; Berno et al., 2008; Sharp et al., 2006). When combined with automated image analysis solutions, this cell line was used to discriminate and classify the mechanistic effects of E2, ER antagonists, BPA and closely related BPA analogs (Mukherjee et al., 2020; Stossi et al., 2014, 2016; Szafran et al., 2017a; Szafran and Mancini, 2014). However, the Judson et al. computational subset models rarely selected the chromatin binding assay that is available from this cell model. The authors suggested that this may be because of reduced sensitivity and difficult analyzing the results (Judson et al., 2017).

Recently, an updated version of the chromatin binding assay using a PRL-HeLa cell line with inducible GFP-ER expression was used to characterize EDCs effects on ERα intra-nuclear mobility (Bolt et al., 2021). Here, we report the characterization of two updated inducible PRL-HeLa in vitro chromatin binding assays for identifying estrogenic activity in high throughput assays. The iER-GFP:PRL-HeLa and iGFP-ER:PRL-HeLa cell lines contain a stable integration of doxycycline-inducible ERα with either C-terminal (iER-GFP) or N-terminal (iGFP-ER) fusion to GFP. We examine the effects of GFP placement on sensitivity to ligands, ERα expression levels, and ligand dependent recruitment of cofactors and RNA Pol II to the promoter array as compared to the original GFP-ER cell line. We sought to understand the ability to detect and classify estrogenic activity from a set of 45 reference compounds defined in Judson et al. (2015). Finally, we use the iER-GFP cell line to test complex, designed mixtures and their individual components, predict an active mixture effects using concentration-response modeling, and compare predicted values to observed responses.

RESULTS
Generation of inducible GFP-ERα and ERα-GFP cell models
The PRL-HeLa cell line contains a multi-copy (~100) integration of a regulatory cassette (containing Pit-1 and ERα binding sites) derived from the prolactin promoter linked to a dsRed2ski fluorescent protein reporter (Sharp et al., 2006). This generates a localized nuclear region with a high density of estrogen response elements (EREs) that appears, when GFP-tagged ERα is expressed, as a bright intra-nuclear fluorescent spot/array indicative of ERα chromatin binding following ligand treatment (Ashcroft et al., 2011; Sharp et al., 2006). Prior efforts to generate a PRL-HeLa stably expressing GFP-ER yielded the GFP-ER:PRL-HeLa cell model which must be maintained in the presence of the ERα antagonist hydroxytamoxifen and has demonstrated lower sensitivity to ER agonists and EDCs than assays based on endogenous ERα expression, leading to its exclusion from simplified computational models of estrogenic EDC activity despite inclusion in the full model (Ashcroft et al., 2011; Judson et al., 2015, 2017; Stossi et al., 2014). To determine if the reduced sensitivity was because of the constitutive expression of ERα or the placement of the GFP-tag, we generated two new cell models based on the PRL-HeLa cell line but containing either inducible GFP-ER (N-terminus tag) or inducible ER-GFP (C-terminus tag). Human ERα was cloned into either the pEGFP-N1 or pEGFP-C1 expression vector which placed a GFP fusion protein on either the N-terminus or C-terminus of the receptor with a short amino acid spacer sequence (Figure S1). The fusion proteins were next cloned into a viral vector that included a tetracycline-inducible promoter. After the parental PRL-HeLa cells were stably transduced, single cell cloning or flow-based sorting was used to generate the final iGFP-ER:PRL-HeLa and iER-GFP:PRL-HeLa cell lines.
Characterization of cell model response

Without doxycycline, both the iGFP-ER and iER-GFP cell lines demonstrate minimal ERα expression, with low levels of background expression of the dsRed2skl2 reporter mRNA (Figure S2). On induction with doxycycline for 20 h, the iGFP-ER and iER-GFP cell models have similar or lower ERα expression (88 and 45%, respectively) compared to the original GFP-ER cell model; however, both have ERα expression levels that are higher than the endogenous levels observed in T47D and MCF7 cell lines (Figure S2).

To establish baseline (anti-) estrogen sensitivity, GFP-ER, iGFP-ER, and iER-GFP cells were treated with 10 nM of 17β-estradiol (E2) or 10 nM 4-hydroxytamoxifen (4HT) for 2 h and compared with untreated samples. Consistent with the original GFP-ER cell model, E2 and 4HT induced a prominent nuclear GFP spot/array in the new iGFP-ERα and iERα-GFP cell lines indicating ERα recruitment to the integrated PRL promoter array (Figure 1A). In general, E2 induced arrays appeared larger in the iER-GFP cells whereas 4HT treatment resulted in arrays with a smaller size.

To quantify these differences, previously described image analysis tools were used to measure nuclear and array features in 60 replicate wells per treatment observed across 3 independent experiments with each sample containing a minimum of 500 cells (Figure 1B, Table S3, Figure S3) (Szafran and Mancini, 2014). In the absence of ligand, the iER-GFP cells exhibited a significantly higher fraction of cells with visible arrays (FCA) compared to iGFP-ER and GFP-ER cells. However, the level of chromatin binding (CB, see Methods for calculation) at detected arrays is similar between cell lines (Figure 1C). When treated with either E2 or 4HT, chromatin binding (CB), and the contributing intensity ratio (IR) measurement, increases in all cell lines but is significantly higher in the iGFP-ER and iER-GFP cell lines compared to the GFP-ER cell line (Figures 1C and 1D). As an indicator of chromatin remodeling, the iER-GFP cells were found to have significantly larger arrays with E2 and significantly smaller arrays with 4HT treatments compared to the GFP-ER and iGFP-ER cells (Figure 1E). Finally, E2 and 4HT treatments altered levels of nuclear GFP signal (Figure 1F). This effect was greatest in the iER-GFP cells with E2 treatment resulting in a 22.5 ± 6.2% decrease and 4HT treatment resulting in a 54.3 ± 18.0% increase in nuclear GFP levels compared to untreated cells after 2 h of ligand exposure. Overall, the inducible cell models responded to both the ERα agonist (E2) and antagonist (4HT) in a manner similar to the original GFP-ER cell model, however, the magnitude of the observed responses were greater and suggests a more sensitive cell model for detecting estrogenic activity.

Inducible ERα expression alters estrogenic sensitivity

To determine baseline sensitivity to estrogens, all cell models were treated with a dose series of either E2 or 4HT ranging from 1 pM to 100 nM for 2 h, fixed, imaged, and analyzed. E2-induced chromatin binding in both the iGFP-ER (EC50 = 0.23 ± 0.09 nM, p < 0.01) and iER-GFP (EC50 = 0.14 ± 0.04 nM, p < 0.01) cells were similar and significantly more sensitive than in original GFP-ER cells (EC50 = 4.05 ± 0.86 nM) (Figure 2A). 4HT treatment produced a similar pattern in chromatin binding sensitivity, with both the iGFP-ER (0.28 ± 0.02 nM) and iER-GFP (0.14 ± 0.02 nM) EC50 values significantly (p < 0.01) lower than those observed in the GFP-ER cells (3.30 ± 0.53 nM) (Figure 2B).

To further understand the E2-induced recruitment of ERα to the PRL locus, we examined the induction of transcriptional activity using mRNA FISH specific to the dsRed2skl reporter protein encoded into the prolactin promoter array (Figure 2C). Again, we observe that both iGFP-ER and iER-GFP cells again had similar EC50 values, left-shifted compared to GFP-ER cells. However, the maximal induced dsRed2skl FISH signal produced from the array in the iER-GFP cells (234,008 ± 1329 IU, 11.4-fold increase above untreated, p < 0.01) was significantly more than that observed from either the iGFP-ER (84,564 ± 628 IU, 5.0-fold increase above untreated) or the GFP-ER (39,685 ± 2338 IU, 3.2-fold increase above untreated) cells. These results suggest that avoidance of constitutive exogenous ERα expression using doxycycline-inducible ERα may significantly increase estrogenic sensitivity in the engineered cell models; moreover, C-terminal placement of the GFP-tag in the iER-GFP cell model is associated with increased E2-induced transcriptional activity at the PRL-locus.

Recruitment of cofactors and RNA polymerase II

To probe the mechanism of the increased transcriptional activity associated with the iGFP-ER cell model, both GFP-ER and iER-GFP cells were treated with 10 nM E2 or 4HT for 2 h and examined for recruitment of selected co-activators and active RNA polymerase II (Figure 3). E2 treatment resulted in strong recruitment of SRC-3 in
both cell lines, however, recruitment was significantly higher in iER-GFP cells compared to GFP-ER cells. This pattern was mirrored by active RNA Pol II recruitment. As expected, treatment of iER-GFP cells with 4HT resulted in significantly lower recruitment SRC-1, SRC-3, and active RNA polymerase II levels at the PRL-array compared to cells treated with E2, consistent with an anti-estrogen effect in the iER-GFP and GFP-ER cell lines. 

![Image of cell lines and treatment conditions](image)

**Figure 1. Initial characterization of ligand response in cell lines expressing inducible GFP-ER and ER-GFP**

GFP-ER:PRL-HeLa and doxycycline induced iGFP-ER:PRL-HeLa and iER-GFP:PRL-HeLa cells were treated with vehicle (DMSO), 10 nM 17β-estradiol (E2), or 10 nM 4-hydroxytamoxifen (4-HT) for 2 h, fixed, and imaged using a Vala Sciences IC200 equipped with a 20×0.75 objective.

(A) Contrast enhanced images of the GFP signal, nuclear segmentation (green lines), and array segmentation (red, offset 5 pixels from actual array segmentation) representing statistically average cells from each cell line and treatment condition. Scale bar indicates 10 µM.

(B–F) Boxplots of the fraction of cells with an array (FCA), chromatin binding (CB), intensity ratio (IR), array size (Size), and GFP average nuclear intensity (NucGFP) generated from >500 cells per sample per treatment condition. Box indicates 25th-75th percentile with line indicating population median, the empty circle the mean, and outliers by full circles.
models. Time-course studies ranging from 10 min to 2 h did not find differences between GFP-ER and iER-GFP cells in timing of cofactor recruitment (data not shown), suggesting the observed differences in transcriptional activity at the PRL-array in iER-GFP cells may be related to enhanced levels of SRC-3 recruitment.

Detection of estrogenic activity in EPA reference chemicals
The ability for newly-developed cell models to detect estrogenic activity was further characterized using a reference chemical library defined by the EPA (Figure 4 and Table S1). This library consists of 45 chemicals with well-defined in vivo estrogenic properties. To further diversify the reference chemicals, they were supplemented with 6 additional chemicals known to act as inhibitors of estrogen activity (i.e., ERα antagonist/antiestrogens). All chemicals were tested by exposing cells for 2 h and a 6-point dose-response curve ranging from 1 nM to 100 μM. Analysis of chromatin binding (CB) in GFP-ER, iGFP-ER, and iER-GFP cell lines offers robust detection of chemical induced receptor activity with Z₀-prime values, a measure for assay quality, for detecting E2 or 4HT induced activity averaging 0.76 or greater for all cell lines (Morelock et al., 2005; Szafran et al., 2020). Using an E2-normalized AUC threshold of 0.1, all three cell lines detected 100% of moderate/strong agonists. However, the GFP-ER cell model only detected activity in 87.5% of weak agonist and 26.7% of very weak agonists, indicative of the reduced sensitivity observed by others (Judson et al., 2017). The chemistries with the lowest detection rates were phenol- and phthalate-based chemicals (Figure S4). In contrast, the iGFP-ER and iER-GFP cell models detected 100% of weak agonists and at least 86.7% of very weak agonists, matching or exceeding the performance of the full EPA-18 model. Due to only one or two false negatives when using the inducible cell models, there is no clear pattern as to what chemistries may be more prone to detection failure. All PRL-HeLa cell models detected a low level of estrogenic activity from a subset of compounds classified as inactive, something not observed with the more complex EPA-18 model. All PRL-HeLa models detected 100% of antagonists tested.

To understand how the chemical structures of the EPA reference chemicals may inform observed assay responses, we used the NR-ToxPred tool to predict (anti-)estrogenic activity of tested compounds. The NR-ToxPred tool uses the Nuclear Receptor Activity (NuRA) dataset and machine-learning based models to predict agonist- and antagonist-like binding to ERα based on chemical structures (Ramaprasad et al., 2022). The predicted likelihood of binding values tends to decrease as in vivo activity decreases (Figure S5A). Many chemicals reported to have very weak or no in vivo estrogenic activity have lower prediction values. Analysis of the false negatives observed with the EPA_18 model and the PRL-HeLa cell models indicate a distribution between chemicals predicted to have high and low binding to ERα (Figures SSB–SSE). 4-nonylphenol, a chemical with high predicted binding, was scored as a false negative in the GFP-ER and iGFP-ER chromatin binding assays (Figures SSC and SSD). Procymidone, a chemical with very low predicted ERα binding, was scored as a false negative in both inducible chromatin binding assays (Figures SSD amd SSE). In terms of false positives, the inducible cell models scored progesterone, spironolactone, and
corticosterone as positive for estrogenic activity, steroids with high predicted binding to ERα but not reported to have in vivo estrogenic activity. In addition, the synthetic imidazole ketoconazole, a chemical with moderate predicted binding to ER, was scored as a false positive in iGFP-ER and iER-GFP assays.

When overall performance of identifying estrogenic activity is examined, the inducible PRL-HeLa cell lines showed an improvement compared to the original constitutive GFP-ERα cell line (Table 1). The reference EPA-18 model has an accuracy, sensitivity, and specificity of 0.91, 0.89, and 1.0, respectively. The GFP-ER model, with its lower sensitivity, has an accuracy, sensitivity, and specificity of 0.69, 0.7, and 0.67. In contrast, the iGFP-ER and iER-GFP cell models have higher accuracy (0.87 and 0.89) due to higher sensitivity (0.95 and 0.98), but with poorer specificity (0.58). When iGFP-ER and iER-GFP cell modes are compared to the 16 individual agonist assays used in the EPA-18 model, accuracy was only exceeded by protein fragment complex assays utilizing Erβ fragments; sensitivity was not exceeded; and specificity was exceeded by assays of multiple types (Figure S6). Due to the robust response observed with ‘inactive’ steroids in the reference library, attempts to increase specificity of the iGFP-ER and iER-GFP cell models by increasing the threshold used for ‘active’ chemicals resulted in decreased sensitivity and overall lower accuracy.

Overall correlation between the EPA-18 model, GFP-ER, iGFP-ER, and iER-GFP exceeded 0.8 across all comparisons (Figure 5). Correlation was highest among chemicals classified as agonists, exceeding 0.9 (Figure S7A). Correlation of responses with antagonists was low between the EPA-18 and PRL-HeLa cell lines (~0.30 to 0.25), likely due to the EPA-18 model bias toward agonist detection and PRL-HeLa based models are measuring antagonist compound’s ability to induce ERα DNA binding, not directly antagonize the activity of E2 (Figure S7B). Finally, the correlation between inactive compounds was generally low (~0.20 to 0.13), driven primarily by the unexpected ERα chromatin binding activity induced by phenobarbital sodium and, as expected, the low estrogenic activity of non-ER steroids such as progesterone, spironolactone, and corticosterone (Figure S7C).

**Phenotypic clustering of agonist and antagonist responses**

We have previously used unsupervised hierarchical clustering to visualize the distinct responses induced in the GFP-ER cell line by estrogenic agonists and antagonists (Stossi et al., 2014; Szafran et al., 2017a). These differences were further confirmed by the successful development of machine-learning classification models (Mukherjee et al., 2020). To determine if similar information can be extracted from responses in the iGFP-ER and iER-GFP cells, hierarchical clustering was applied using a limited feature set describing the observed array phenotype (intensity ratio (IR), and size) and nuclear levels of GFP-tagged ERα. Data from active compounds was normalized to internal E2 controls and range normalized using a Z score calculation for each cell line. Clustering was accomplished using group averages and Euclidean distance, with a total of 6 clusters identified for each cell line (Figure 6). In both inducible cell lines, similar pattern of responses emerges. There is one (iER-GFP, Figure 6B) or two (iGFP-ER, Figure 6A) clusters that contain nearly all agonists regardless of their known level in vivo activity. The few remaining agonists fall into isolated single-compound clusters or cluster with antagonist due to the relative intensity of visible arrays. The
remaining clusters largely consists of antagonists with relatively small visible arrays, with individual clusters differentiated by their effects on the nuclear level of GFP-tagged ERα and the relative intensity of the visible arrays. A similar separation of agonists and antagonists into separate clusters was also observed in the analysis of the GFP-ER cell data (Figure S8), however fewer active compounds were analyzed due to lower sensitivity. In addition, the predominant determinant of membership in either an agonist- or antagonist-rich cluster was the relative effect of a compound on nuclear GFP levels.

Analysis of complex mixtures

Next, the ability to use the iER-GFP cell line to characterize estrogenic activity of complex mixtures of chemicals was examined. The mixtures evaluated here have been previously used in risk assessment studies based on a panel of functional and cytotoxic assays (Hsieh et al., 2021). These mixtures were designed to be diverse, with concentrations determined by either EC50 values found in the ToxCast database, estimated exposure levels as defined by ExpoCast, point-of-departure (POD) values determined from in vivo studies, or from RfDs themselves. Mixtures were tested using an assay design in which each component is in constant proportion to one another across mixture dilutions ranging from 200-fold to 20,000,000-fold. Data was normalized using a parallel dilution series of 100 μM E2 and AUC calculated. Two mixtures, AC50 Mx and Abdo mix, were found to have at least 10% of the activity observed with E2 (Figure 7A). Although other mixtures had low levels of activity (0-10%), these did not meet the 10% threshold previously established to define active samples (Judson et al., 2015).
Next, the individual chemicals used to assemble the most active mixture (AC50 mix) were examined. iER-GFP cells were treated with a dose series of each chemical ranging between 1 nM and 100 μM for 2 h and the level of ER-GFP chromatin binding was determined. Five active chemicals achieving at least 10% of the activity of E2 were identified (Figure 7B). These included 3 chemicals (o,p’-DDT, methoxychlor, dicofol) with known weak or very weak estrogenic activity in vivo and two (ethion, endosulfan) with unknown in vivo activity.

To simplify further analysis of the AC50 mix mixture response, only the active individual chemicals were considered in subsequent analysis, reducing a mixture from 42 chemicals to five. At the lowest dilution (1:200) the mixture had 15% less activity than o,p’-DDT, the most efficacious chemical in the mixture (Figure 7C). However, at dilutions 1:20,000 and greater, the mixture had 28 to 55% more activity than predicted for o,p’-DDT at the concentration at which it was present in the mixture. To understand how components of the mixture might be interacting with ER-GFP in the chromatin binding response, the effect summation (ES), independent action (IA), and generalized concentration addition (GCA) prediction models were calculated using either all 5 components or the 2 most active components (Figure 7C). As expected, at the highest dilutions the models converge and are within the error range of the observed AC50 mix mixture response. However, at lower dilutions (i.e., higher chemical concentrations) the models diverge greatly with the ES model exceeding the observed results. The 5- and 2-component IA model tend to overestimate the observed effects compared to the GCA models. For the GCA models, the 5-component model tends to better predict the intermediate dilutions and the 2-component model accurately predicts the mixture response observed with the lowest dilutions. This is consistent with chemicals acting through a common pathway; and the GCA model best accounts for the antagonistic effects of partial agonists at high concentrations.

Finally, the response phenotype of the active mixtures and individual chemicals at the highest tested concentration were determined and compared against a set of estrogenic agonists, weak agonists, very weak agonists, and antagonists (Figure 7D). Similar to the analysis of the EPA reference chemicals, the response phenotype was determined by the intensity ratio (IR) of the identified array compared to the nucleoplasm, size of the array (Size), and overall nuclear GFP intensity (NucGFP). The AC50 Mix and Abdo mixtures were

| Table 1. Performance based validation of single cell models measuring potential estrogen receptor (ER) activities in a supplemented EPA reference chemical set |
|-----------------------------------------------|
| **Assay** | **EPA_18** | **GFP-ER Model** | **iGFP-ER Model** | **iER-GFP Model** |
|----------|------------|------------------|------------------|------------------|
| Assay Quality | Z-Prime | Z-Prime | Z-Prime | Z-Prime |
| Chromatin Binding (E2) | N/A | 0.76 ± 0.04 | 0.89 ± 0.01 | 0.87 ± 0.02 |
| Chromatin Binding (4HT) | N/A | 0.81 ± 0.04 | 0.86 ± 0.04 | 0.88 ± 0.02 |
| Detection By Class | Count (%) | Count (%) | Count (%) | Count (%) |
| Mod/Strong Agonist | 6 (100) | 6 (100) | 6 (100) | 6 (100) |
| Weak Agonist | 8 (100) | 7 (87.5) | 8 (100) | 8 (100) |
| Very Weak Agonist | 11 (73.3) | 4 (26.7) | 13 (86.7) | 14 (93.3) |
| Inactive | 0 (0) | 4 (33.3) | 5 (41.7) | 5 (41.7) |
| Antagonist | 6 (54.5) | 11 (100) | 11 (100) | 11 (100) |

Overall Performance
- # true pos: 31, 28, 38, 39
- # true neg: 12, 8, 7, 7
- # false pos: 0, 4, 5, 5
- # false neg: 4, 12, 2, 1
- Accuracy: 0.91, 0.69, 0.87, 0.88
- Sensitivity: 0.89, 0.7, 0.95, 0.98
- Specificity: 0.67, 0.7, 0.58, 0.58
- balanced accuracy: 0.945, 0.685, 0.765, 0.78

*Z*-prime values >0.5 indicate high assay quality.
found to be members of a cluster consisting of weak and very weak agonists. This cluster also contained 4 of the 5 individual active chemicals from the AC50 mix. o,p'-DDT appears in a separate cluster consisting of estrogenic agonists and weak agonists, consistent with its known in vivo activity. Again, estrogenic antagonists formed their own distinct clusters distinct from the agonists in the analysis. This suggest that the active AC50 mix and Abdo mixtures most likely act as weak or very weak agonists.

**DISCUSSION**

We have updated the well-established image-based high throughput GFP-ER:PRl-HeLa chromatin binding assay for detecting estrogenic activity by generating two new monoclonal cell lines, iGFP-ER:PRl-HeLa and iER-GFP:PRl-HeLa, in which the fluorescently tagged ERα expression is inducible with doxycycline treatment. This allows us to maintain the cell lines in the absence of ERα expression, eliminating the need for low concentrations of 4HT to inhibit a constitutively-expressed ERα and maintain PRL-HeLa cell growth. The inducible cell lines maintain the same endpoints as the original chromatin binding assay, allowing for the direct observation of multiple mechanistic steps of ERα genomic signaling. These include (anti-)estrogen-induced chromatin binding and remodeling at an ERE-rich locus, coregulator recruitment, transcriptional activity, and ERα protein stabilization/degradation. Importantly, we observed a significant left-shift in estrogenic sensitivity in both inducible ERα cell lines, with EC50 values of ~150-200 p.m. compared to ~4 nM observed in the original non-inducible GFP-ERα cell line. These EC50 values are more similar to the ~6-10 p.m. EC50 values for E2 mediated transcription of endogenous GREB1 and MYC genes in MCF7 cells and results observed in other single cell and bulk population assays (Dahlman-Wright et al., 2006; Stossi et al., 2020).

Despite the well-established use of fluorescently-tagged nuclear receptors to understand the mechanisms by which they act, there is minimal published work directly comparing the effect of N- versus C-terminal placement of the fluorescent protein. Here, we used the monoclonal PRL-HeLa cell line to generate two derivatives expressing either the N-terminal fusion of GFP-ER or the C-terminal fusion of ER-GFP. We did not observe a notable difference in sensitivity of visible array formation/chromatin binding to E2, 4HT or a panel of 51 reference chemicals. In terms of promoter DNA interactions, this suggest that placement of the fluorescent tag does not generate functional differences in ligand affinity to the binding pocket.
within the C-terminal ligand binding (LBD) domain of ERα (Dahlman-Wright et al., 2006; Hewitt and Korach, 2018).

However, we found E2 treatment in the C-terminally linked ER-GFP cell line generated significantly more transcriptional activity compared to N-terminally linked GFP-ER cell lines. This was associated with larger visible array sizes and enhanced recruitment of the SRC-3 co-activator protein. The N-terminal A/B domain of ERα is largely unstructured, allowing it to modulate interactions with ligands, DNA, and coregulator proteins (Hewitt and Korach, 2018). Recently, cryo-EM experiments have shown that the AF1 region within the A/B domain resides in closeness, proximity to the LBD and its interaction with SRC-3 (Hewitt and Korach, 2018). It is plausible that the presence of the N-terminal GFP tag may either impart abnormal structure to the A/B domain or cause steric interference between the ERα LBD and SRC-3 leading to decreased transcription. In addition, the higher ERα expression observed in the GFP-ERα cell lines compared to the ERα-GFP cell line may contribute to squelching of transcriptional activity due to competition for a limited pool of transcription factors within the cell (Meyer et al., 1989). However, studies examining endogenous
Figure 7. Analysis of complex mixtures using the iER-GFP cell line

Induced iER-GFP cells were exposed to (A) dilutions (1:200 to 1:200,000) of complex designed mixtures or (B) dose series (1 nM–100 μM) of individual chemicals for 2 h and ERα chromatin binding was measured per cell per sample. See acronym explanations and description of the design mixtures in Table S1. AUC values were determined and normalized to DMSO and E2 controls. When known, EPA in vitro model predictions (EPA_18) are shown. Responses below 0.1 were considered inactive (shaded region). (C-Top panel) Dose-response curves of the active components of the AC50 Mix mixture (lines) and the observed AC50 Mix response (open circles) in the ray dilution series. (C-Bottom panel) Observed AC50 Mix response compared against effect summation (ES), independent action (IA), and generalized concentration addition (GCA) model predictions generated using either all 5 or the 2 most active mixture components. (D) Unsupervised hierarchical clustering of response phenotypes of reference chemicals, active mixtures, or active individual mixture chemicals. Heatmap color scale indicates Z-score relative to independent E2 control samples run in parallel. Experimental data represents the mean ± SE of 4 samples.
single allele transcription of the GREB1 gene in MCF7 cells did not find evidence of squelching occurring at higher ERα expression levels (Stossi et al., 2020).

The growing use of high throughput in vitro assays and computational models to predict in vivo estrogenic activity has greatly accelerated the understanding of how thousands of man-made chemicals present in the environment may impact health. A model based on 18 in vitro assays, including the original GFP-ER:PRL-HeLa chromatin binding assay, was validated using an in vitro and in vivo reference set to predict estrogenic activity of individual chemicals (Browne et al., 2015; Judson et al., 2015). Attempts to streamline the number of assays and maintain accuracy identified accurate subgroups with as few as four in vitro assays (Judson et al., 2017). The GFP-ER:PRL-HeLa chromatin binding assay was generally excluded from the smaller subgroups based on sensitivity. The results with the present study suggest that this limitation is not an issue with the inducible GFP-ER and ER-GFP cell lines. With an E2-sensitivity mirroring endogenous ERα in a breast cancer cell line, screens performed with either the iGFP-ER or iER-GFP cell lines achieved at least 95% sensitivity for detecting estrogenic chemicals using the previously established AUC and threshold values. The AUC values for estrogenic chemicals agreed (Pearson’s correlation >0.84) with those predicted by models based on multiple assays. The low number of false negatives in the current study precludes full understanding of specific chemistries poorly detected using the inducible PRL-HeLa assays. However, analysis using the NR-ToxPred tool to predict receptor binding suggest some chemicals with very weak in vivo activity are more likely to bind poorly to ER. The screen design of the EPA reference chemicals, which used a fixed concentration range of 1 nM–100 μM, may have resulted in false negatives due the limited range of concentrations tested and low ligand affinity.

An important feature of the chromatin binding assay based on the PRL-HeLa cell line is that detection of estrogenic activity is not based on ER-mediated transcriptional regulation. Because ERα agonists, selective estrogen modulators (SERMs) and pure ERα antagonists such as Fulvestrant (ICI 182 780) induce ERα DNA-binding, a single iGFP-ER or iER-GFP PRL-HeLa chromatin binding assay can detect both agonist and antagonist ERα ligands and effectively distinguish them using unsupervised clustering approaches based on the multiple mechanistic or phenotypic endpoints available in the image-based assay (Ashcroft et al., 2011; Stossi et al., 2014; A. T. Szafran et al., 2017a; Welboren et al., 2009). The newly developed inducible IER-GFP cell line has greater recruitment of coregulator proteins which is associated with greater changes in visible array size and receptor expression in response to E2 treatment compared to the original GFP-ER cell line. Despite no new biology being incorporated into the iER-GFP cell line, the larger dynamic range in observed phenotypes may improve the accuracy classifying agonist and antagonist estrogenic chemicals using machine-learning models compared to efforts based on the original GFP-ER:PRL-HeLa cell line (Mukherjee et al., 2020).

A key advantage to a full ERα model developed using a diverse set of in vitro assays is resilience to the false-positive and false-negative results that are often inherent to the technology being used by any one assay (Judson et al., 2015, 2017). Although the increased assay sensitivity observed with the iGFP-ER and iER-GFP cell models decreased the frequency of false negatives in the panel of reference chemicals, it came at the cost of an increased false-positive rate resulting in specificity of 0.58 in both cell lines. It is important to note that the lower specificity was largely driven by the detection of weak estrogenic activity from non-estrogenic steroids (progesterone, corticosterone, and spironolactone). These steroids lack the characteristic phenolic A-ring of estrogens, but share a common A-ring structure with an oxo group at position 3. At the high concentrations tested (upper limit 100 μM) and in the presence of the multimerized PRL promoter array, this common structure may allow the weak chromatin binding of ERα observed. This is supported by both progesterone and corticosterone having generated (anti-)estrogenic activity in yeast-based estrogenic assays despite no in vivo estrogenic activity being reported (Kolle et al., 2010).

We also unexpectedly detected significant ERα chromatin binding induced by phenobarbital sodium. Although this result was replicated with multiple copies of the reference chemicals obtained from the EPA, these copies originated from the same laboratory stock solution, and we could not secure an independent supplier. Although phenobarbital is known to influence E2 metabolism through the CAR receptor, no evidence exists of a direct interaction between phenobarbital and ERα itself. However, phenobarbital activates CAR through an EGFR-mediated cell signaling pathway through a threonine motif that is conserved in CAR and ERα (Fashe et al., 2018). We have previously shown the ability of EGF to induce ERα chromatin binding in the PRL-HeLa cell line, suggesting a possible mechanism for the observed phenobarbital activity
Further studies would be needed to understand if phenobarbital induces ERα chromatin binding through such a mechanism, or if the current observations are a false-positive result because of contamination or assay interference. Determining the sensitivity and specificity in the iGFP-ER and iER-GFP models is important in validating the assays and highlights the need for a well-defined set of interference chemicals, those likely to cause false-positives, to defining the domain of applicability of assay. Importantly, this knowledge and the multi-endpoint nature of the chromatin binding assays will allow the design of a focused battery of in vitro assays that will efficiently probe for chemical effects on multiple aspects of the ERα signaling pathway(s).

Although the ability to screen individual chemicals using the PRL-HeLa cell line has been shown here and elsewhere, most environmental exposure to toxicants occurs as complex mixtures and need to be studied as such (Hsieh et al., 2021; Watt et al., 2016). To examine how the PRL-HeLa chromatin binding assay may be used in this setting, a panel of complex environmental mixtures were screened for their ability to induce ERα chromatin binding. Many of these mixtures have been previously examined using a panel of induced pluripotent stem cell (iPSC) models to understand functional and cytotoxic effects (Hsieh et al., 2021). Focusing on low dose effects defined by EC_{10} and point-of-departure (POD) values, many mixtures were found to be active, and effects best predicted by models based on a common pathway where component effects are summed by their relative concentration (concentration addition, CA). When examined here using shorter exposure times (2 h) and a focused mechanistic cell model (iER-GFP), two mixtures (AC50 Mix, Abdo Mix) met the 0.1 AUC activity threshold with a response phenotype that matched weak or very weak agonist references. An additional 8 mixtures had very low activity (AUC 0.01 to 0.1). Like the prior iPSC cytotoxicity assays, a concentration addition model (GCA) produced more accurate results as concentrations increased as both ES and IA models overestimated mixture effects because of their limited ability to account for partial agonists. As the dilution increased, all tested models (GCA, IA, ES) approximated each other, agreeing with the accuracy that others have found in nuclear receptor driven in vitro assays (Silva et al., 2002; Watt et al., 2016).

A key limitation of the use of GCA equation is the requirement that it must be inverted to predict mixture effects. That is achieved by limiting the Hill coefficient to 1 during dose curve fitting. Although generally considered valid for nuclear receptors such as PPARγ which heterodimerize and therefore bind ligands at a single binding site, it is unclear if the same limitation can be applied to ERα which forms stable homodimers with two ligand binding sites (Watt et al., 2016). For active individual compounds, we found we could apply the Hills slope coefficient limitation and still achieve R^2 values >0.98, suggesting that, in the context of the highly multimerized iER-GFP chromatin binding assay, each ligand binding site can be considered to act independently. Furthermore, for the AC50 Mix mixture, we found that prediction accuracy at higher concentrations improved by limiting models to the two most active individual components. This suggests that, as saturating concentrations are reached, the most efficacious components dominate the mixture effect possibly because of lower dissociation rates and longer receptor-ligand interactions (Oostenbrink et al., 2000; Tummino and Copeland, 2008). These factors are less likely to play a role at lower concentrations, further explaining the convergence of all model predictions at higher dilutions.

The continued evolution of the PRL-HeLa chromatin binding assay from transient transfection of GFP-ERα into a library of stable cell lines expressing GFP-ER, GFP-ERβ, inducible chimeric progesterone receptor (iGFP-PRER), inducible chimeric androgen receptor (iGFP-ARER), and now iGFP-ER and iER-GFP offers a platform of tools to screen EDCs that can provide early insight into potential mechanisms of action. Although the original PRL-HeLa stably expressing GFP-ERα demonstrated a lower sensitivity, the updated inducible GFP-ERα and ERα-GFP cell lines achieve sensitivity ~equivalent to endogenously expressed ERα while correlating well with models based on 18 in vitro assays. This warrants consideration for inclusion in screening programs as more complex high throughput in vitro assays will play an important role in the early assessment of potential EDCs, complex environmental mixtures, therapeutic modulators of ER, and how all might interact in human exposure.

Limitations of the study
Although the inducible PRL-HeLa cell models have advantages over traditional methods for detecting estrogenic activity, challenges and study limitations remain beyond those already discussed above. First, reliable detection of chromatin binding by iER-GFP or iGFP-ER is dependent on the quality of sample preparation, requires high resolution automated imaging platforms, and, as with most high content assays, is
dependent on custom image analysis and data analysis algorithms. However, these resources are increas-
ingly common in the modern laboratory setting. Second, our cluster analysis of response phenotypes as
either agonist- or antagonist-like were based on published in vivo data, not directly analyzed in the cell
models using RNA FISH. Direct RNA FISH analysis of transcriptional activity in the iER-GFP cell model after
exposure to EPA reference chemicals has recently been published and is in agreement with the results pre-
sented (Bolt et al., 2021). Finally, to avoid further complexities of downstream effects from ligand exposure
and fit the assay within a high throughput context, short term experiments were used, which prevents ob-
servations of responses requiring more time. Cell proliferation assays for the detection of estrogenic EDC
activity are well established, simple, and sensitive (Wang et al., 2012). Because cell proliferation assays
typically last for days/weeks, they potentially allow the detection of active metabolites or other important
secondary changes that may not be detected in an HT assay focusing on early (~1-2h) mechanistic steps in
receptor activation.

STAR METHODS
Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell lines
  - Reference chemical libraries
- METHOD DETAILS
  - Cell treatment with chemicals
  - Staining and immunofluorescence labeling
  - Sample imaging
- QUANTIFICATION AND DATA ANALYSIS
  - Image analysis and quantification
  - Data analysis
  - Prediction of chemical binding to ERα
  - Mixture prediction models
  - Statistical analysis

SUPPLEMENTAL INFORMATION
Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.105200.

ACKNOWLEDGMENTS
M.A.M., F.S., and A.T.S. are funded by NIEHS (Project 4, M.A.M. PI, and Superfund Research Program grant,
P42ES027704, PI, I. Rusyn), M.A.M. and F. S. are supported via the CPRIT-funded GCC Center for Advanced
Microscopy and Image Informatics (RP170719; PI, M.A.M.). M.A.M. received additional support from the
NIEHS-funded Center for Precision Environmental Health (P30ES030285, PI, C. Walker). Imaging is also
supported by the Integrated Microscopy Core at Baylor College of Medicine with funding from NIH
(DK56338, and CA125123), CPRIT (RP150578), the Dan L. Duncan Comprehensive Cancer Center, and
the John S. Dunn Gulf Coast Consortium for Chemical Genomics. Additional thanks for Dr. Moczygemba
and the Flow Cytometry and Cell Sorting Facility (FCCSF) located at the Texas A&M University Institute of
Bioscience and Technology. The authors are also indebted to Keith Houck, PhD (EPA), for coordinating ac-
cess to reference chemicals used in this study. Finally, we would like to thank members of the Texas A&M
Superfund Research Program, especially Zunwei Chen, Ivan Rusyn, and Weihueh Chiu for providing infor-
mation and access to reference chemicals and mixtures to Baylor College of Medicine.

AUTHOR CONTRIBUTIONS
Conceptualization, A.T.S, F.S., and M.A.M.; Investigation, A.T.S. and M.G.M.; Methodology, A.T.S. and
M.G.M.; Writing – original draft, A.T.S.; Revisions, A.T.S.; Visualization, A.T.S.; Supervision, F.S. and
M.A.M.; Funding acquisition, M.A.M.
DETECTION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: May 10, 2022
Revised: July 26, 2022
Accepted: September 21, 2022
Published: October 21, 2022

REFERENCES

Ashcroft, F.J., Newberg, J.Y., Jones, E.D., Mikic, I., and Mancini, M.A. (2011). High content imaging-based assay to classify estrogen receptor-α ligands based on defined mechanistic outcomes. Gene 477, 42–52. https://doi.org/10.1016/j.gene.2011.01.009.

Berno, V., Amazit, L., Hinojos, C., Zhong, J., Mancini, M.G., Sharp, Z.D., and Mancini, M.A. (2008). Activation of estrogen receptor-alpha by E2 or EGF induces temporally distinct patterns of large-scale chromatin modification and mRNA transcription. PLoS One 3, e2286. https://doi.org/10.1371/journal.pone.0002286.

Bolt, M.J., Singh, P., Obkirchner, C.E., Powell, R.T., Mancini, M.G., Szafrań, A.T., Stossi, F., and Mancini, M.A. (2021). Endocrine disrupting chemicals differentially alter intranuclear dynamics and transcriptional activation of estrogen receptor-α. iScience 24, 103227. https://doi.org/10.1016/j.isci.2021.103227.

Browne, P., Judson, R.S., Casey, W.M., Kleinsteuber, N.C., and Thomas, R.S. (2015). Screening chemicals for estrogen receptor bioactivity using a computational model. Environ. Sci. Technol. 49, 8804–8814. https://doi.org/10.1021/acs.est.5b02291.

Chen, Z., Liu, Y., Wright, F.A., Chiu, W.A., and Rusyn, I. (2020). Rapid hazard characterization of environmental chemicals using a compendium of human cell lines from different organs. ALTEX 37, 623–638. https://doi.org/10.14573/altex.2002291.

Dahman-Wright, K., Cavailles, V., Fuqua, S.A., Jordan, V.C., Katzenellenbogen, J.A., Korach, K.S., Maggi, A., Muramatsu, M., Parker, M.G., and Gustafsson, J.-A. (2006). International union of pharmacology. LXIV. Estrogen receptors. Pharmacol. Rev. 58, 773–781. https://doi.org/10.1124/pr.58.4.8.

Ezechial, M., Janochova, J., Filipová, A., Křesnová, Z., and Cajthaml, T. (2016). Widely used pharmaceuticals present in the environment revealed as in vitro antagonists for human estrogen and androgen receptors. Chemosphere 152, 284–291. https://doi.org/10.1016/j.chemosphere.2016.02.067.

Fashe, M., Hashiguchi, T., Yi, M., Moore, R., and Negishi, M. (2018). Phenobarbital-induced phosphorylation converts nuclear receptor RXRα from a repressor to an activator of the estrogen sulfotransferase gene Sult1e1 in mouse livers. FEBS Lett. 592, 3769–3780. https://doi.org/10.1002/1873-4689.13199.

Gore, A.C., Chappell, V.A., Fenton, S.E., Flaws, J.A., Nadal, A., Putluri, V., and Zöllner, R.T. (2015). Executive summary to EDC-2: the endocrine society’s second scientific statement on endocrine-disrupting chemicals. Endocr. Rev. 36, 593–602. https://doi.org/10.1210/er.2015-1093.

Hewitt, S.C., and Korach, K.S. (2018). Estrogen receptors: Risk characterization and probabilistic concentration-response modeling of complex environmental mixtures using new approach methodologies (NAMs) data from organotypic in vivo human stem cell assays. Environ. Health Perspect. 129, 17004. https://doi.org/10.1289/EHP760.

Judson, R.S., Houck, K.A., Watt, E.D., and Thomas, R.S. (2017). On selecting a minimal set of in vitro assays to reliably determine estrogen agonist activity. Regul. Toxicol. Pharmacol. 91, 39–49. https://doi.org/10.1016/j.yrtph.2017.09.022.

Judson, R.S., Magpantay, F.M., Chickarmane, V., Haskell, C., Tania, N., Taylor, J., Xia, M., Huang, R., Rotroff, D.M., Filer, D.L., et al. (2015). Integrated model of chemical perturbations of a biological pathway using 18 in vitro high throughput screening assays for the estrogen receptor. Toxicol. Sci. 148, 137–154. https://doi.org/10.1093/toxsci/kfv168.

Kolle, S.N., Tania, N., Taylor, J., Xia, M., Huang, R., Rotroff, D.M., Filer, D.L., et al. (2015). Integrated model of chemical perturbations of a biological pathway using 18 in vitro high throughput screening assays for the estrogen receptor. Toxicol. Sci. 148, 137–154. https://doi.org/10.1093/toxsci/kfv168.

Mukherjee, R., Beykal, B., Szafrań, A.T., Onel, M., Stossi, F., Mancini, M.G., Lloyd, D., Wright, F.A., Zhou, L., Mancini, M.A., and Pistikopoulos, E.N. (2020). Classification of estrogenic compounds by coupling high content analysis and machine learning algorithms. PLoS Comput. Biol. 16, e1008191. https://doi.org/10.1371/journal.pcbi.1008191.

Oostenbrink, B.C., Pitera, J.W., van Lipzig MM, van L.M., Meerman, J.H., and van Gunsteren WF, van G.W. (2000). Simulations of the estrogen receptor ligand-binding domain: affinity of natural ligands and xenoestrogens. J. Med. Chem. 43, 4594–4605. https://doi.org/10.1021/jm001045d.

Ramasprasad, A.S.E., Smith, M.T., McCoy, D., Hubbard, A.E., la Merrill, M.A., and Durkin, K. (2022). Predicting the binding of small molecules to nuclear receptors using machine learning. Brief. Bioinform. 23, bbac114. https://doi.org/10.1093/bib/bbac114.

Safe, S.H. (1995). Environmental and dietary estrogens and human health: is there a problem? Environ. Health Perspect. 103, 346–351. https://doi.org/10.1289/ehp.9510034.

Sharpe, Z.D., Mancini, M.A., Hinojos, C.A., Dai, F., Berno, V., Szafrań, A.T., Smith, K.P., Lele, T.P., Ingber, D.E., and Mancini, M.G. (2006). Estrogen-receptor-exchange and chromatin dynamics are ligand- and domain-dependent. J. Cell Sci. 119, 4365. https://doi.org/10.1242/jcs.03264.

Silva, E., Rajapakse, N., and Kortenkamp, A. (2002). Something from “nothing” - eight weak estrogenic chemicals combined at concentrations below NOECs produce significant mixture effects. Environ. Sci. Technol. 36, 1751–1756. https://doi.org/10.1021/es0101227.

Stossi, F., Bolt, M.J., Ashcroft, F.J., Lamerdin, J.E., Melnick, J.S., Powell, R.T., Dandeker, R.D., Mancini, M.G., Walker, C.L., Westwick, J.K., and Mancini, M.A. (2014). Defining estrogenic mechanisms of bisphenol A analogs through high throughput microscopy-based contextual assays. Chem. Biol. 21, 743–753. https://doi.org/10.1016/j.chembiol.2014.03.013.

Stossi, F., Dandeker, R.D., Bolt, M.J., Newberg, J.Y., Mancini, M.G., Kaushik, A.K., Putluri, V., Sreekumar, A., and Mancini, M.A. (2016). High throughput microscopy identifies bisphenol A, a bisphenol A analog, as a novel AR.
down-regulator. Oncotarget 7, 16962–16974. https://doi.org/10.18632/oncotarget.7655.

Stossi, F., Dandekar, R.D., Mancini, M.G., Gu, G., Fuqua, S.A.W., Nardone, A., De Angelis, C., Fu, X., Schiff, R., Bedford, M.T., et al. (2020). Estrogen-induced transcription at individual alleles is independent of receptor level and active conformation but can be modulated by coactivators activity. Nucleic Acids Res. 48, 1800–1810. https://doi.org/10.1093/nar/gka2172.

Szafran, A.T., Bolt, M.J., Obkirchner, C.E., Mancini, M.G., Helsen, C., Claessens, F., Stossi, F., and Mancini, M.A. (2020). A mechanistic high-content analysis assay using a chimeric androgen receptor that rapidly characterizes androgenic chemicals. SLAS Discov. 25, 695–708. https://doi.org/10.1177/2472555220922917.

Szafran, A.T., and Mancini, M.A. (2014). The myImageAnalysis project: a web-based application for high-content screening. Assay Drug Dev. Technol. 12, 87–99. https://doi.org/10.1089/adt.2013.532.

Szafran, A.T., Stossi, F., Mancini, M.G., Walker, C.L., and Mancini, M.A. (2017a). Characterizing properties of non-estrogenic substituted bisphenol analogs using high throughput microscopy and image analysis. PLoS One 12, e0180141. https://doi.org/10.1371/journal.pone.0180141.

Szafran, A.T., Stephan, C., Bolt, M., Mancini, M.G., Marcelli, M., and Mancini, M.A. (2017b). High-content screening identifies src family kinases as potential regulators of AR-V7 expression and androgen-independent. Prostate 77, 82–93. https://doi.org/10.1002/pros.23251.

Tummino, P.J., and Copeland, R.A. (2008). Residence time of receptor-ligand complexes and its effect on biological function. Biochemistry 47, 8465. https://doi.org/10.1021/bi801226m.

"Endocrine disruptor screening program (EDSP) overview." US EPA Official Website, US EPA, updated 2018a, [WWW Document] https://www.epa.gov/endocrine-disruption/endocrine-disruptor-screening-program-eds-overview. Accessed 1 May 2019.

"Toxicology Testing in the 21st Century (Tox21)." US EPA Official Website, US EPA, updated 2018b, [WWW Document] https://www.epa.gov/chemical-research/toxicology-testing-21st-century-tox21. Accessed 1 May 2019.

Wang, S., Aarts, J.M.J.G., Evers, N.M., Peijnenburg, A.A.C.M., Rietjens, I.M.C.M., and Bovee, T.F.H. (2012). Proliferation assays for estrogenicity testing with high predictive value for the in vivo uterotrophic effect. J. Steroid Biochem. Mol. Biol. 128, 98–106. https://doi.org/10.1016/j.jsbmb.2011.11.009.

Watt, J., Webster, T.F., and Schlezinger, J.J. (2016). Generalized concentration addition modeling predicts mixture effects of environmental PPARγ agonists. Toxicol. Sci. 153, 18–27. https://doi.org/10.1093/toxsci/kfw100.

Welboren, W.-J., van Driel, M.A., Janssen-Megens, E.M., van Heeringen, S.J., Sweep, F.C., Span, P.N., and Stunnenberg, H.G. (2009). ChIP-Seq of ERα and RNA polymerase II defines genes differentially responding to ligands. EMBO J. 28, 1418–1428. https://doi.org/10.1038/emboj.2009.88.

Zoeller, R.T., Brown, T.R., Doan, L.L., Gore, A.C., Skakkebaek, N.E., Soto, A.M., Woodruff, T.J., and Vom Saal, F.S. (2012). Endocrine-disrupting chemicals and public health protection: a statement of principles from the Endocrine Society. Endocrinology 153, 4097–4110. https://doi.org/10.1210/en.2012-1422.
## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| **Antibodies**      |        |            |
| Anti-Estrogen Receptor α Antibody, clone 60C, rabbit IgG monoclonal | Millipore-Sigma | 04-820, RRID: AB_1587018 |
| Anti-SRC-1 Antibody, clone 8, mouse IgG monoclonal | BD Biosciences | 612378, RRID: AB_399738 |
| Anti-SRC-3 Antibody, Rabbit polyclonal | Abcam | ab2831, RRID: AB_2151072 |
| Anti-BRG1 Antibody, clone EPNCIR111A, rabbit IgG monoclonal | Abcam | ab110641, RRID: AB_10861578 |
| Anti-RNA Pol II (phospho-Ser5), clone 4H8, mouse IgG monoclonal | Abcam | ab5408, RRID: AB_304868 |
| Anti-RNA Pol II (phospho-Ser2), rabbit IgG polyclonal | Abcam | ab5095, RRID: AB_304749 |
| Anti-Histone H3 (acetyl K9, AcH3) Antibody, clone Y28, rabbit recombinant | Abcam | ab32129, RRID: AB_732920 |
| Anti-Histone H3 (di methyl K4, DiMeH3) Antibody, clone Y47, rabbit recombinant | Abcam | ab32356, RRID: AB_732924 |
| **Chemicals, peptides, and recombinant proteins** | | |
| 17β-estradiol (E2) | Sigma | E8875 |
| 4-hydroxytamoxifen (4HT) | Sigma | H6278 |
| Bazedoxifene | Sigma | P20018 |
| AZD9496 | Sellek Chemicals | S8372 |
| H3B-5942 | Sellek Chemicals | H1-112611 |
| ICI 182,790 | Tocris Bioscience | 1047/1 |
| ZK 164015 | Tocris Bioscience | 2183/10 |
| Toremifene citrate | Toronto Research Chemicals | T547500 |
| EPA 45 ER Reference Library | Dr. Keith Houck (US EPA, Research Triangle Park, NC) | N/A |
| ATSDR Reference Mixtures and Chemicals | Dr. Ivan Rusyn (Texas A&M, College Station, TX) | N/A |
| **Experimental models: Cell lines** | | |
| GFP-ER:PRL-HeLa cell line | Msancini Lab | N/A |
| PRL-HeLa cell line | Mancini Lab | N/A |
| GFP-ER:PRL-HeLa cell line | This paper | N/A |
| iGFP-ER:PRL-HeLa cell line | This paper | N/A |
| MCF7 cell line | ATCC | HTB-22 |
| T47D cell line | ATCC | HTB-133 |
| **Recombinant DNA** | | |
| pEGFP-C1-ER alpha | Mancini lab/Adgene | Plasmid #28230 |
| pEGFP-N1-ER alpha | Mancini lab | N/A |
| plINDUCER20 | Dr. Trey Westbrook (Baylor College of Medicine, Houston, TX)/Adgene | Plasmid #44012 |
| **Software and algorithms** | | |
| Pipeline Pilot Analysis software | Biovia | [https://www.3ds.com/products-services/biovia/products/data-science/pipeline-pilot/](https://www.3ds.com/products-services/biovia/products/data-science/pipeline-pilot/) |
RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources and reagents should be directed to, and will be fulfilled by the corresponding author and lead contact, Michael A. Mancini (mancini@bcm.edu).

Materials availability
All unique/stable reagents generated in this study are available from the lead contact with a completed Materials Transfer Agreement.

Data and code availability
- Image data reported in this paper will be shared by the lead contact upon request.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines

**iGFP-ER and iER-GFP PRL-HeLa cell lines**
PRL-HeLa cells expressing tetracycline-inducible full length GFP-tagged ERα was generated by cloning the EGFP-C1-ER or the EGFP-N1-ER coding region into pINDUCER20, containing a tetracycline-inducible promoter, a gift from Dr. Trey Westbrook (Baylor College of Medicine, Houston, TX). Despite the large packaging size of the vectors, high titer Lentivirus was generated by System Biosciences LLC (Palo Alto, CA). In a parallel process, PRL-HeLa cells were transduced with either inducible GFP-ER (iGFP-ER) or inducible ER-GFP (iER-GFP) viral particles and cells with a stable integration were enriched using geneticin (G418) drug selection, flow cytometry, and single cell cloning. The final populations of iGFP-ER:PRL-HeLa and iER-GFP:PRL-HeLa cells are >95% GFP positive following doxycycline induction. Both cell lines are maintained in phenol red-free Dulbecco’s modified Eagle’s medium (DMEM-HG) supplemented with 5% FBS, 200 μg/ml hygromycin, and 400 μg/ml G418.

For experiments, iGFP-ER and iER-GFP PRL-HeLa cells were seeded in phenol red-free DMEM with 5% charcoal-stripped/dialyzed FBS without selection agents or inhibitors on 384-well plates (384-IQ, Aurora Biotechnologies) at a target density of 3,000 cells/well. Cells were left to adhere overnight. Cell lines, cells were treated with doxycycline (0.8 μg/ml) for 20 hours before media was exchanged prior to treatment

**GFP-ER:PRL-HeLa cell line**
GFP-ER:PRL-HeLa cells were maintained in phenol red-free DMEM containing 5% FBS, 0.8 μg/ml blastidin, 200 μg/ml hygromycin, and 10 nM 4-hydroxytamoxifen as previously described (Ashcroft et al., 2011). GFP-ER cells were seeded in phenol red-free DMEM with 5% charcoal-stripped/dialyzed FBS without selection agents or inhibitors on 384-well plates (384-IQ, Aurora Biotechnologies) at a target density of 3,000 cells/well. Cells were left to adhere overnight.

**T47D and MCF7 cell lines**
T47D and MCF7 are genotype verified female breast cancer cell lines obtained from the AmericanType Culture Collection (ATCC) and were maintained according to recommended conditions in DMEM media with 10% FBS. For experiments, cells were seeded in phenol red-free DMEM with 5% charcoal-stripped/dialyzed...
FBS without selection agents or inhibitors on 384-well plates (384-IQ, Aurora Biotechnologies) at a target density of 3,000 cells/well. Cells were left to adhere overnight prior to use.

Reference chemical libraries
A set of 45 estrogenic EDC reference chemicals (Table S1) with well documented in vivo activity were provided in DMSO via agreement with Dr. Keith Houck (US EPA, Research Triangle Park, NC).

Mixtures and individual chemicals used in the complex mixture analysis were prepared as previously described (Chen et al., 2020; Hsieh et al., 2021). Mixtures (Table S2) were designed using a set of criteria that included 1) diverse set of chemicals belonging to multiple classes of environmental toxicants and 2) concentrations based on known in vitro data found in the ToxCast database, in vivo data used to determine regulatory oral non-cancerous reference doses (RfDs), or general population exposure levels derived from the ExpoCast database. Individual chemicals for comparison to active mixtures were from the Agency for Toxic Substances and Disease Registry (ATSDR 2020) priority list of hazardous substances. These chemicals are frequency detected at U.S. National Priority List sites (i.e., Superfund sites). Chemicals were from Sigma-Aldrich, except heptachlor, heptachlor epoxide, 2,4,5-trichlorophenol, parathion, benzidine and o,p’-DDT, which were from ChemService and provided in DMSO by the lab of Dr. Ivan Rusyn (Texas A&M, College Station, TX) as part of the Superfund Research Program. All mixtures and individual chemicals were then tested using six 10-fold serial dilutions.

METHOD DETAILS
Cell treatment with chemicals
During initial cell line characterization, reference compounds were manually prepared at concentrations ranging from 1 pM to 100 nM and added to cells for 2 hours. The EPA and ATSDR reference chemicals and mixtures were provided in pre-formatted multi-well plates in DMSO at a known concentration. In vitro screening in cells was performed at concentrations ranging from 1 nM to 100 μM (EPA, ATSDR) and 200-fold to 200,000,000-fold dilution (mixtures) and a treatment time of 2 hours. Immediately prior to compound treatment, the stock compound solutions were transferred to 96-well working dilution plates in phenol red-free DMEM with 5% SD-FBS. Working dilutions of compounds were then added in quadruplicate to assay plates containing prepared cells. For all experiments, negative control wells containing 0.5% DMSO and positive control wells containing E2 ranging from 0.01 nM to 1000 nM were included in the plate layout. All dilutions and cell treatments using reference chemicals and mixtures were performed using the Beckman Biomek NX liquid handling platform.

Staining and immunofluorescence labeling
After incubation period was completed, cells were fixed using 4% EM-grade formaldehyde in PEM buffer (80 mM potassium PIPES [pH 6.8], 5 mM EGTA, and 2 mM MgCl2) and quenched with 0.1 M ammonium chloride for 10 min. Cell membranes were disrupted by incubating samples with a 0.5% Triton X-100 solution for 10 minutes. Nuclei were stained using DAPI (1 μg/ml) for 10 minutes. For samples in which antibody labeling was used, incubation with 0.5% Triton-X was extended to 30 minutes and cells were incubated in blotto (5% milk in Tris-buffered saline/Tween 20) for 30 minutes prior to addition of primary antibody solution (ERα, SRC-1, SRC-3, BRG1, phospho-Ser5 RNA polymerase II, phospho-Ser2 RNA polymerase II, 0.5 μg/mL, ACh3, DiMeH3). Primary antibody solution was incubated overnight at 4C prior to incubation with secondary antibody (Alexa 647 conjugated Anti-Ms IgG, Molecular Probes) for 1 hour at room temperature. Nuclei were stained using DAPI (1 μg/ml) for 10 minutes. All samples were stored in PBS buffer containing calcium, magnesium, and sodium azide at 4C prior to imaging.

Sample imaging
Automated imaging was carried out using a Vala Sciences IC-200 (San Diego, CA) image cytometer. Image acquisition was performed with a 20x/0.75NA objective and a scientific CMOS camera. DAPI, GFP, and A647 (antibody-labeled samples) channels were imaged using Z-stacks captured at 1 μm intervals (for a total of 5 μm) at 1 x 1 binning. Z-stacked were projected into a single 2-D image per channel using Vala Sciences proprietary synthetic focus algorithm using medium filtering. To capture a minimum of 500 cells per sample, a total of 4 to 5 fields were imaged.
QUANTIFICATION AND DATA ANALYSIS

Image analysis and quantification

Image analyses were performed using custom measurement and reporting routines incorporated into myImageAnalysis/Pipeline Pilot software (Szafran and Mancini, 2014). Cell, nucleus, and array (if present) were identified and metrics describing shape and intensity features quantified. Aggregated cells, mitotic cells, and apoptotic cells were removed using filters based on nuclear size, nuclear shape, and nuclear intensity. The fraction of cells with a detected/visible array was calculated per sample using Equation (1):

\[
\text{Fraction of Cells with Array (FCA)} = \frac{N_{\text{Array}}}{N_{\text{Total}}}
\]

(Equation 1)

where \(N_{\text{Array}}\) is the number of cells with a detected array and \(N_{\text{Total}}\) is the total number of cells analyzed. The relative GFP intensity at the array compared to the nucleoplasm was calculated per cell using Equation (2):

\[
\text{Intensity Ratio (IR)} = \frac{\text{Intensity}_{\text{Array}}}{\text{Intensity}_{\text{NP}}}
\]

(Equation 2)

where \(\text{Intensity}_{\text{Array}}\) is the average GFP pixel intensity in the array mask and \(\text{Intensity}_{\text{NP}}\) is the average GFP pixel intensity in the nucleoplasm mask. Chromatin binding was calculated per cell using Equation (3):

\[
\text{Chromatin Binding (CB)} = (IR - 1) \times A_{\text{Array}}
\]

(Equation 3)

where \(IR\) is the intensity ratio and \(A_{\text{Array}}\) is the area in pixels contained in the array mask.

Data analysis

The chemical responses were analyzed as previously described using a standardized analysis pipeline generated using Pipeline Pilot (Biovia) (Szafran et al., 2017b, 2020). This pipeline performed all necessary baseline correction, normalization, curve-fitting, and hit-calling. All screening data were normalized to a fractional range with DMSO and 100 nM E2 treated controls set at 0 and 1. All dose response data was fit to either a constant model or a Hill function using least squares curve fitting. The activity observed for each chemical on measured endpoints is summarized by integrating the response over the tested concentration range (Area Under Curve, AUC). AUC values are normalized to E2 and labeled as significant if greater than 0.1. This threshold was selected based on its use in previous EPA computational models (Judson et al., 2015).

Prediction of chemical binding to ER\(\alpha\)

Prediction of chemical binding to ER\(\alpha\) was performed using the NR-ToxPred online tool (http://nr-toxpred.chem.berkeley.edu/predict) (Ramaprasad et al., 2022). Chemical Pubmed CID numbers were used to retrieve simplified molecular-input line-entry system (SMILES) structures used as input into NR-ToxPred. Analysis was performed using Morgan fingerprints, the superlearner algorithm, and multiclass analysis. Predicted agonism- and antagonism-like binding were compared and the highest value used to describe each chemical.

Mixture prediction models

Prediction models used in mixture analysis were calculated as previously described (Hsieh et al., 2021; Watt et al., 2016). Since a generalized concentration addition model requires an invertible dose-response function, the effect on an individual chemical on iER-GFP chromatin binding was fit to the Hill Equation (4):

\[
f([A]) = \min + \frac{(\max - \min)([A])}{K_a + [A]}
\]

(Equation 4)

where \(K_a\) is the equilibrium constant for chemical A (equivalent to the EC50 when the Hill coefficient was restricted to 1), min and max are the curve minimum and maximum effect level, and [A] is the concentration of chemical A. Effect summation (ES) assumes each chemical in the mixture acts independently and sums the individual responses, was calculated for a mixture of \(n\) chemicals using Equation (5):

\[
E_{\text{ES}} = \sum_{i=1}^{n} f([X_i])
\]

(Equation 5)
where \( [X] \) is the concentration of the \( i \)th chemical in the mixture. Independent action (IA) assumes chemicals with differing efficacies act through different mechanisms and have a maximal effect of 1, was calculated using Equation (6):

\[
E_{IA} = \alpha_{max} \left( 1 - \prod_{i=1}^{n} \left( 1 - \frac{f_i([X]_i)}{\alpha_{max}} \right) \right),
\]

(Equation 6)

where the maximum efficacy (\( \alpha_{max} \)) is based on 17\( \beta \)-estradiol. Generalized concentration addition assumes chemicals act through a common mechanism and tolerates chemicals with differing efficacies and the presence of partial agonists, was calculated using Equation (7):

\[
E_{GCA} = \frac{\sum_{i=1}^{n} \frac{[X]_i}{K_i} \frac{1}{1 + \sum_{i=1}^{n} \frac{[X]_i}{K_i}}}{1 + \sum_{i=1}^{n} \frac{[X]_i}{K_i}}.
\]

(Equation 7)

where \( \alpha \) is the maximal efficacy, \( K \) is the EC50, and \( [X]_i \) is the concentration of the \( i \)th chemical of the mixture.

**Statistical analysis**

Data presented were acquired from a minimum of 3 independent experiments. Statistical significance was determined using one-way ANOVA with Tukey’s post hoc test for comparisons across multiple samples. Tests were performed using OriginPro (version 2020b) software, and \( p < 0.05 \) was considered statistically significant. All graphs were generated using OriginPro (version 2020b) software.