SUPPLEMENTARY INFORMATION

"Phenotypic Profiling of 6PPD, 6PPD-quinone and Structurally Diverse Antiozonants in RTgill-W1 Cells Using the Cell Painting Assay"

Felix R. Harris[1,2], Mark D. Jankowski[3], Daniel L. Villeneuve[4], Joshua A. Harrill[1,*]

Affiliations:

[1]U.S. EPA, Office of Research and Development, Center for Computational Toxicology and Exposure, Biomolecular and Computational Toxicology Division, Research Triangle Park, NC 27711, United States

[2]Oak Ridge Institute for Science and Education, Oak Ridge, TN 37830, United States

[3]U.S. EPA, Region 10, Laboratory Services and Applied Science Division, Seattle, WA 98101, United States

[4]U.S. EPA, Office of Research and Development, Center for Computational Toxicology and Exposure, Great Lakes Toxicology Division, Duluth, MN 55804, United States

*Corresponding Author

Contents

Supplemental Methods

Method S1. Sample Processing and Imaging

Cell culture and chemical dosing

RTgill-W1 stocks were thawed and grown for two consecutive passages (approximately 3 population doublings) in media containing 5% heat inactivated fetal bovine serum (HI-FBS) in an incubator (19 °C) shielded from light. Cells were seeded into flasks at no more than +/- 10% of $7.0x10^4$ cells/cm^2 and grown for seven days, with complete media replacement on day 3. After two consecutive passages, cells were seeded into Revvity PhenoPlate™ 384-well plates at 22,500 cells/well (corresponding to approximately 2,117 cells/cm^2) in 40 µL of growth media (Table S3) using a Certus Flex liquid handler and 0.45/0.15 microvalve.

Approximately 24 hours after seeding, the growth media was removed and replaced with 40 µL of minimal media containing no serum (Table S3). After the media change, 200 nL of 200X chemical stocks solubilized in DMSO were dispensed into each well of all plates using a LabCyte Echo 550 acoustic dispenser. The day after dosing (but before further processing was done, ~20 hours after dosing) each plate was imaged using the same imaging protocols described in "Imaging" to determine if any chemical treatments produced visible artifacts in the fluorescent channels used for Cell Painting that would be indicative of chemical precipitation in the assay medium. The images were manually inspected for artifacts, and none were found that were both systemic throughout every well containing that chemical concentration and bright enough to interfere with or overpower either the cell viability or Cell Painting fluoroprobes.

Cell viability assay

Exactly 23.5 hours after dosing, 2 µL of the cell viability labeling solution containing propidium iodide and Hoechst 33342 (Table S4) were added to each well of every cell viability plate using a Certus Flex liquid handler and a 0.1/0.03 microvalve. Plates were returned to the 19 °C incubator for 30 minutes, after which at exactly 24 hours post-dosing cells were fixed by application of 12 µL of 16% paraformaldehyde using a MultiFlo FX Liquid Dispenser. The fixative remained on the cells for 10 minutes. Then each plate was washed using 1X PBS, also using a MultiFlo FX Liquid Dispenser, and sealed with an optical plate seal. Cell viability plates were wrapped in foil and stored at 4 °C until imaging.

Cell Painting assay

Similar to the cell viability assay, at exactly 23.5 hours after dosing, 2 µL of a "live-label" solution containing MitoTracker™ DeepRed (Table S4) was applied to each well of every Cell Painting plate using a Certus Flex liquid handler and a 0.1/0.03 microvalve. Plates were then returned to the 19 °C incubator for thirty minutes. At exactly 24 hours post-dosing cells were fixed via addition of 12 µL of 16% paraformaldehyde using a MultiFlo FX Liquid Dispenser. The fixative remained on the cells for 10 minutes before plates were washed with 1X PBS, also using a MultiFlo FX Liquid Dispenser. The plates were then sealed with aluminum seals, wrapped in foil, and stored at 4 °C until further processing.

All plates were stained either 1 (for biological replicates 3 and 4) or 2 days (for biological replicates 1 and 2) after fixation. Plates were removed from the refrigerator and equilibrated to room temperature for 1 hour. Then the volume of 1X PBS in each well was drained to 40 µL. Cells were permeabilized by addition of 10 µL of 0.5% Triton X-100 solution filtered using a syringe and 0.22 µm filter, using a Certus Flex liquid handler. After 30 minutes of permeabilization plates were washed once again with 1X PBS using a MultiFlo FX Liquid Dispenser, leaving a final volume of 40 µL 1X PBS in each well. The remaining fluorescent probes were combined (Table S4) and 2 µL of this "full stain" was applied to each well using a Certus Flex liquid handler. The full stain contained Hoechst 33342 to visualize DNA, SYTO14 to visualize RNA and nucleoli, Alexa Fluor™ 488 conjugated Concanavalin A to visualize the endoplasmic reticulum (ER), Alexa Fluor™ 568 conjugated phalloidin to visualize the actin cytoskeleton, Alexa Fluor™ 555 conjugated Wheat Germ Agglutinin (WGA) to visualize Golgi apparatus and plasma membrane. The plates were then moved into dark conditions for 30 minutes. Then a final 1X PBS wash was conducted using a MultiFlo FX Liquid Dispenser. The plates were sealed with optical plate seals and stored at 4 °C until imaging.

Imaging

Images of all plates were acquired at least 24 hours after the final labeling step. Plates were allowed to equilibrate to room temperature for at least 1 hour prior to imaging. Cell Painting and cell viability plates were both acquired using a Revvity Opera Phenix Plus high content imaging system and analyzed using Revvity Harmony software, as described previously[1-3].

Cell Painting plates were imaged using a 20X water immersion objective across four individual fluorescent channels that corresponded to: DNA (405 nm excitation/435-480 nm emission); RNA/ER (488 nm excitation/500-550 nm emission); actin cytoskeleton, Golgi bodies and plasma membrane (AGP) (561

nm excitation/570-630 nm emission) and mitochondria (Mito) (640 nm excitation/650-760 nm emission). Confocal imaging was used to acquire 7 z-planes 2.5 µm apart to capture three-dimensional variations in cell colony growth. Z-planes were combined into one maximum projection image in each fluorescent channel for analysis. A previously developed Harmony analysis protocol segmented cells and derived 1294 phenotypic features for each cell that describe cell morphology.

Cell viability plates were imaged in non-confocal mode across two channels: Hoechst 33342 and propidium iodide. Z-offsets for each channel were set so the nuclei of the cells were in focus. Cells were segmented and intranuclear propidium iodide fluorescence was measured using previously developed Harmony protocols and exported into a tab separated value file for further analysis.

Method S2. Data analysis

Cell viability assay

Cell-level values were imported into R, and cells were initially filtered based on nuclei size (between 30 and 1,000 μM^2) and roundness (>0.3) to exclude any artifacts present in the image. Cells were then identified as propidium iodide positive if the intranuclear propidium iodide fluorescence was greater than the 95^{th} percentile of intranuclear propidium iodide fluorescence in vehicle control wells. Cell-level data was then aggregated to well-level data by defining two values. The first value is the percentage of cells which are propidium iodide positive; higher values indicate a loss of cell viability. The second value is the normalized cell count calculated as a percent by dividing the number of cells in the well by the median number of cells in vehicle control wells and multiplying by 100; lower values indicate a loss of cell viability.

The well-level propidium iodide and normalized cell count endpoints were then subject to concentration-response modeling using the R package *tcplfit2* v0.1.7[4]. Normalized cell count was fit to either a constant or hill model with a BMR of 50 to derive a benchmark dose (BMD) at which point the normalized cell count was decreased by 50% compared to vehicle control. The propidium iodide endpoint was fit to either a constant, hill, or gain-loss model with a BMR of 5 * the normalized median absolute deviation (nMAD) of vehicle control wells.

Cell Painting assay

Profile Generation and Dimensional Reduction

Within each well, feature data from each cell was first aggregated by taking the median of each feature, resulting in 1294 raw values per well. These well-level values were then normalized to the median and normalized median absolute deviation (nMAD) of vehicle control wells within a plate, and then z-scaled using the mean and standard deviation of all vehicle control wells in the study. This resulted in 1294 z-scaled feature values per well. These are the features that are then later used for the feature selection and profile correlation analysis.

Synthetic null chemicals were generated as described previously[5] from the lowest two concentrations of chemicals which exhibited no cytotoxicity below the second highest chemical treatment level. Responses from these two dose levels were sampled twice and assigned concentrations in the same range as 6PPD and the other PPD parent compounds (100, 30, 10, 3, 1, 0.3, 0.1, and 0.03 µM). There were enough eligible data points to generate 12 null chemicals, with one response value from each culture plate. The null chemicals were subject to the same concentration-response modeling procedure as the true test chemicals and were used to select *tcplfit2* hitcall thresholds that minimize false positive rates for identifying active chemicals [3].

All 1,294 features were then analyzed using dimensional reduction techniques as described previously[3]. This consisted of either the "Global Mahalanobis" approach that used all 1,294 features to generate a single latent variable for concentration-response modeling or the "Category-level Mahalanobis" approach where features are partitioned according to the channel, compartment, and module from which they were derived and latent variables calculated for each of the 49 feature categories. Wells with at least 500 cells were then used for principal component analyses to calculate rotation matrices for each approach. The Global Mahalanobis approach retained 16 eigenfeatures that covered 95% of variance, and the Category-level Mahalanobis approach retained enough eigenfeatures for each group to cover 95% of variance.

Concentration-response modeling and PAC determination

Concentration-response modeling was performed using the R package *tcplfit2* (v0.1.7)[4] as described previously[1, 3] using the global and category-level Mahalanobis distances. Data from cytotoxic concentrations were discarded from curve fitting by removing all data above the lowest observable effect concentration found from the cell viability assay. Noise levels were determined by using distance values

from the two lowest concentrations of all test chemicals, and a Tukey's outer fence test was conducted. Values that exceeded Tukey's outer fence (i.e., 75th percentile + 3x the interquartile range of potential noise distance values) were removed to ensure no bioactivity was included in the treatments used to model noise in the data and define the *tcplfit2* cutoff value for each endpoint. Curves were fit using one of nine models: constant, hill, gain-loss, first and second order polynomial, power, and exponential 1-4. A curve was only fit to the data if values exceeded a cutoff of 1 SD, and the BMR was set to 1.349 * SD.

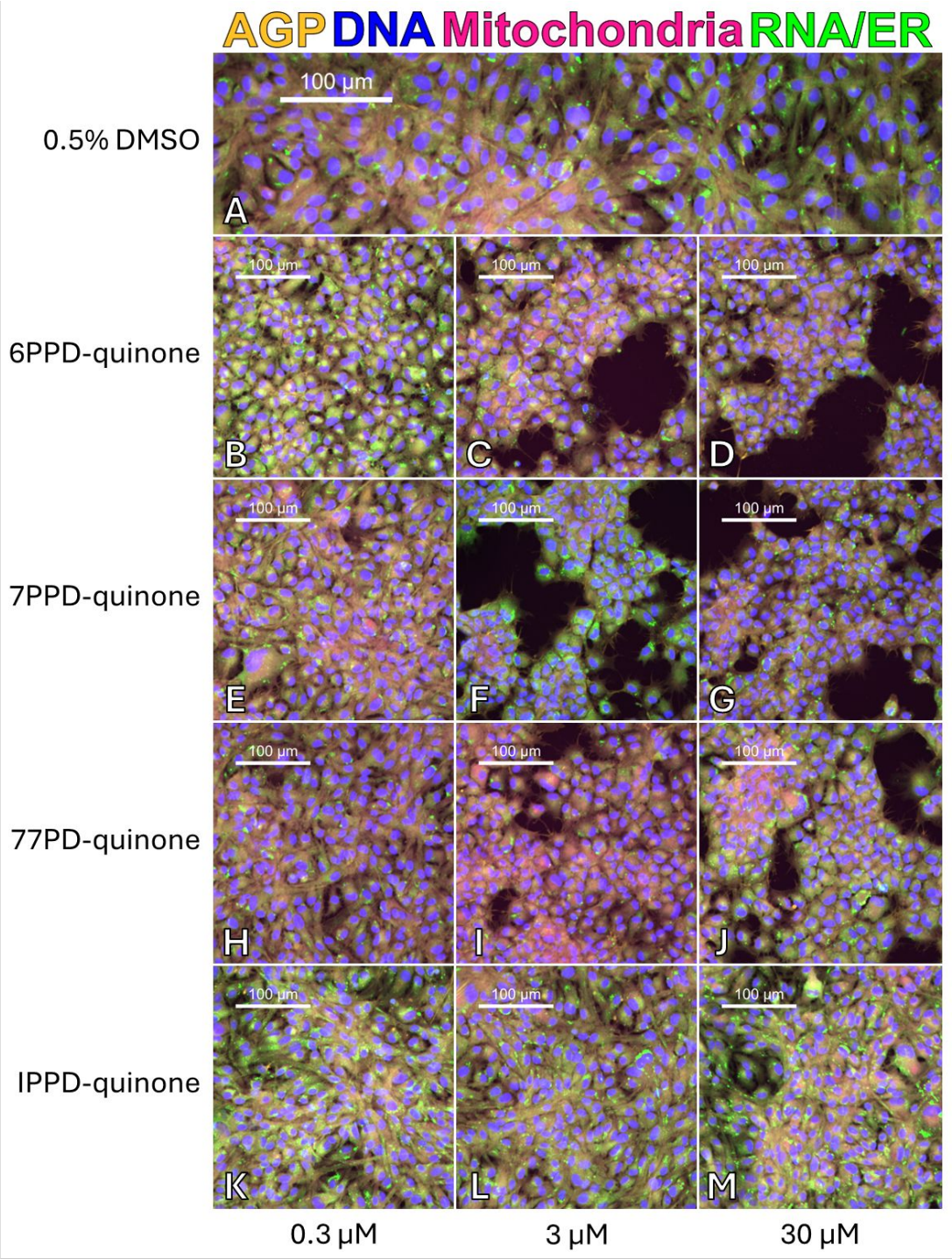
Global Mahalanobis BMCs were valid if the BMD from the best fit curve was lower than the highest tested concentration. Category-level BMCs were valid if the BMD from the best fit curve was lower than the highest tested concentration AND if a continuous hitcall ≥ 0.95 was achieved. The hitcall threshold of 0.95 was determined by choosing a hitcall threshold which resulted in no perceived biological activity in any of the synthetic null chemicals. The global BMC was also defined as the global phenotype altering concentration (PAC), and the lowest category-level BMC was defined as the category-level PAC. The overall PAC is the minimum between the global and category-level PACs.

Feature selection and profile correlation analysis

The phenotypic profile correlation analysis was carried out as described previously[5]. The full set of 1294 features contained some features which were found to be highly correlated or redundant after processing. As a result, a feature selection process was implemented to remove correlation and redundancy from the feature set. Feature selection was a multi-step process which began by: 1) removing any features that had constant values across treatments (0 features eliminated at this step in this study), 2) removing features that were not correlated with themselves across multiple biological replicates to remove any features which are not reproducible between plates (195 features eliminated), and 3) systematically removing features with high correlation to another feature until all remaining pairs of features had Kendall correlation of 0.75 or less with another feature (637 features eliminated). This left 462 features for the profile correlation analysis. The features used for profile correlation are listed in File S2.

Profiles of normalized Cell Painting data were then filtered to include only the 462 features remaining after feature selection, and feature values between -1 and 1 (within 1 nMAD of vehicle control) were set to 0 to reduce noise and focus on distinct biological effects. Concentrations used for profile correlation began at the test concentration directly above the PAC value (or the highest concentration if no PAC was derived) and ranged up to two test concentrations above, given none of these concentrations

were cytotoxic. This excluded styrenated phenol from profile correlation analysis, as the first concentration above the PAC was also experiencing cytotoxicity. Profile similarity was then computed using Kendall correlation, and the maximum correlation value for each chemical pair was kept as the correlation between chemicals.



Supplemental Figures

Figure S1. Representative images across equimolar concentrations of 6PPD-quinone, 7PPD-quinone, 77PD-quinone, and IPPD-quinone. Images are all taken from the same experimental plate. Conditions shown are (A) 0.5% DMSO (vehicle control); 6PPD-quinone at 0.3 µM (B), 3 µM (C), and 30 µM (D); 7PPD-

quinone at 0.3 µM (E), 3 µM (F), 30 µM (G); 77PD-quinone at 0.3 µM (H), 3 µM (I), and 30 µM (J); and IPPD-quinone at 0.3 µM (K), 3 µM (L), and 30 µM (M).

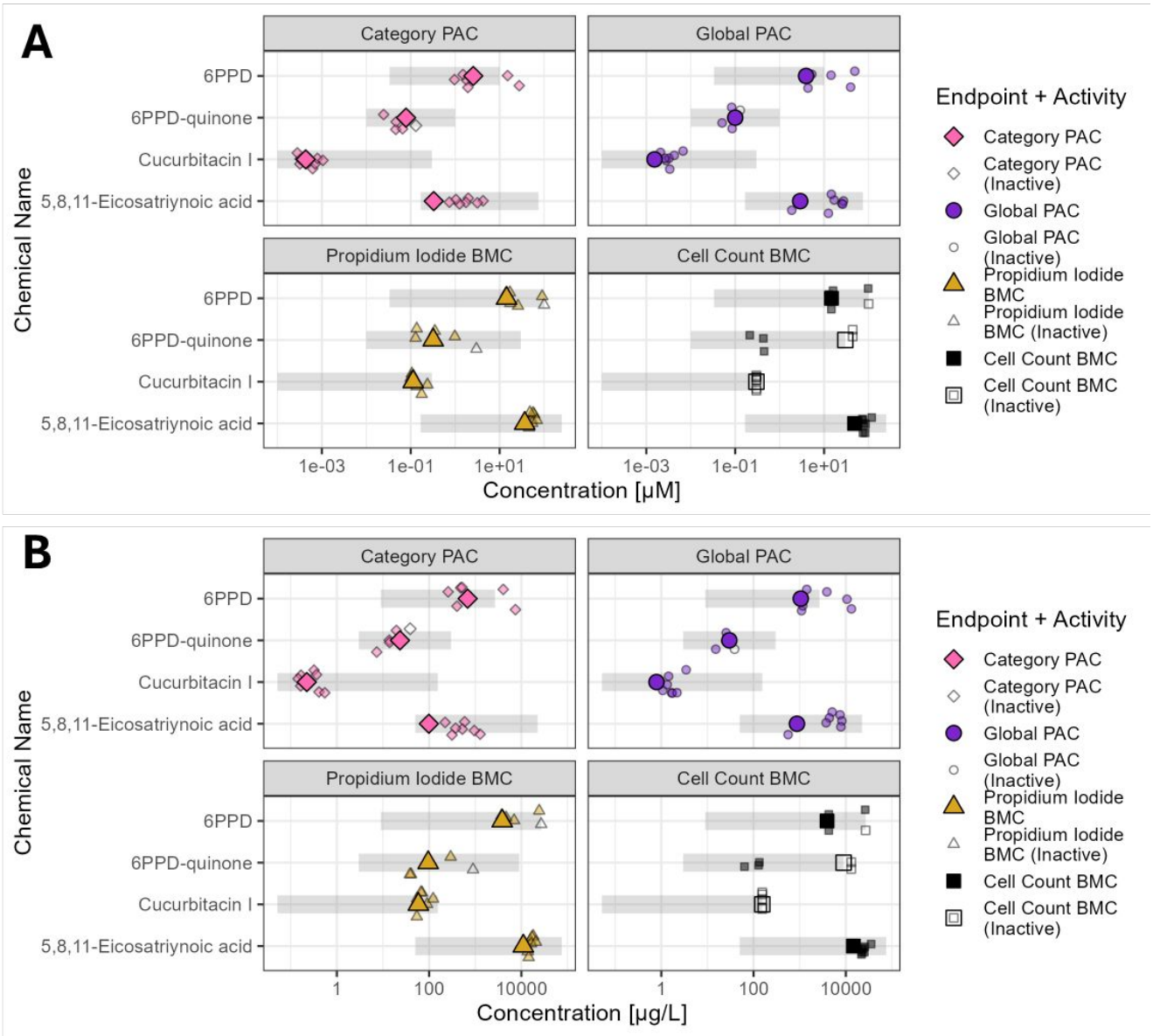


Figure S2. Reproducibility of PAC and cell viability BMC values between this experiment and previous experiments. Data for 6PPD and 6PPD-quinone from the present study and Jankowski et al. (2025) are displayed. Data for phenotypic reference chemicals 5,8,11-Eicosatriynoic acid and Cucurbitacin I from the present study, Jankowski et al. (2025), and Nyffeler et al. (2025) are displayed. Concentrations are shown in µM (A) and µg/L (B). Gray bars indicate dose ranges used for concentration-response modeling. For PAC endpoints the dose ranges are truncated for decreases in cell viability as determined in the current experiment. The complete range of test concentrations are shown for cell viability BMC endpoints. Points for the current experiment are indicated by large symbols. Data from previous studies are represented

with smaller symbols. Solid symbols indicate activity, and inactive symbols are open and displayed at the highest concentration used for curve-fitting.

Supplemental Tables

Table S1. Materials, reagents, and instruments used with sourcing information.

Item	Provider	Catalog #
Cell culture		
Neubauer Hemacytometer Counting Chamber	Hausser Scientific	3200
25cm^2 T25 cell culture flask	Corning	431082
75cm^2 T75 cell culture flask	Corning	430641
225cm^2 T225 cell culture flask	Corning	431082
CRL-2523™ RTgill-W1 Cells	ATCC	CRL-2523
0.4% Trypan Blue Cell Dye	ThermoFisher	T10282
LifeTech™ TrypLE™ Select Enzyme	ThermoFisher	12563011
10X Phosphate Buffered Saline	Sigma-Aldrich	P7059
Gibco™ Leibovitz's L-15 medium (no phenol red) with Galactose (900 mg/L), Sodium Pyruvate (550 mg/L), and L-Glutamine (300 mg/L)	ThermoFisher	21083-027
Gibco™ Leibovitz's L-15 medium with phenol red, Galactose (900 mg/L), Sodium Pyruvate (550 mg/L), and L-Glutamine (300 mg/L)	ThermoFisher	11415064
Heat-inactivated Fetal Bovine Serum	Sigma Aldrich	F4135
HyClone™ Penicillin (10kU/mL) and Streptomycin (10mg/mL) in 0.85% NaCl	Cytiva	SV30010
Dimethyl Sulfoxide (DMSO)	ATCC	4-X
Sodium chloride (NaCl)	Sigma Aldrich	S5886
Potassium chloride (KCl)	Sigma Aldrich	P5405
Magnesium sulfate ($MgSO_4$)	Sigma Aldrich	208094
Magnesium chloride ($MgCl_2$)	Sigma Aldrich	M8266
Calcium chloride ($CaCl_2$)	Sigma Aldrich	C5670
Disodium phosphate (Na_2HPO_4)	Sigma Aldrich	5136
Potassium dihydrogen phosphate (KH_2PO_4)	Sigma Aldrich	P5655
Galactose	Sigma Aldrich	G5388
Sodium Pyruvate (100 mM)	Gibco	11360-070
Fisherbrand™ Mini Low Temperature Refrigerated Incubator, 18 L	Fisher Scientific	15-015-2632
PhenoPlate 384-well, black, optically clear flat-bottom, tissue-culture treated	Revvity	6057300
Instruments & their supplies		
Certus Flex Liquid Dispenser	Trajan Scientific and Medical	certus-flex-liquid-dispenser
Certus Flex SMLD 300GC Valve (0.45mm/0.15mm)	Trajan Scientific and Medical	CERT.21772
Certus Flex SMLD 300GC Valve (0.20mm/0.10mm)	Trajan Scientific and Medical	CERT.21769

Certus Flex SMLD 300GC Valve (0.10mm/0.03mm)	Trajan Scientific and Medical	CERT.21765
CyBio FeliX	AnalytikJena	--
CyBio FeliX 96/250 Tips	AnalytikJena	--
Echo® 550 Liquid Handler	Beckman Coulter	--
Echo® 555 Liquid Handler	Beckman Coulter	--
Echo® Qualified 384-Well Low Dead Volume (LDV) Source Microplates	Beckman Coulter	LP-0200
Echo® Qualified 384-Well Polypropylene (PP) Source Microplates	Beckman Coulter	PPT-0200
MultiFlo™ FX Microplate Dispenser	BioTek®	MFXP1
BioTek® Small Volume (1µL) Cassette	BioTek®	7170012
5 mL Luer-Lock™ Syringe	Fisher Scientific	BD 309646
1.7 mL microcentrifuge tubes	BioExpress	C-3262-1
MicroAmp™ Optical Adhesive Film	ThermoFisher	4311971
Cell Painting Reagents		
Hoechst 33342	Invitrogen	H3570
Concanavalin A	Invitrogen	C11252
AlexaFluor™ 568 Phalloidin	Invitrogen	A12380
SYTO 14	Invitrogen	S7576
MitoTracker Deep Red	Invitrogen	M22426
Wheat Germ Agglutinin, AlexaFluor™ 555 Conjugate	Invitrogen	W32464
16% Paraformaldehyde	Electron Microscopy Sciences	15710S
Triton™ X-100 Detergent	Sigma-Aldrich	T8787
Sodium Bicarbonate	Sigma-Aldrich	S6014
Sodium Azide	Sigma-Aldrich	S2002
Bovine Serum Albumin	Sigma-Aldrich	A2153
Cell viability reagents		
Propidium Iodide	ThermoFisher	P3566

Table S2. Wells excluded from analysis due to processing failures.

Plate Barcode	Assay	Wells	Reason
TC00005021	Cell Painting	A1:A24, B1:B24	Issue dispensing fixative, wells empty
TC00005022	Cell Viability	P1:P24	Issue dispensing stain, cells unstained

Table S3. Chemical information supplied by vendors, including catalog number, molecular weight, lot number and purity.

Chemical Name	CAS RN (DTXSID)	Molecular Weight (g/mol)	Supplier (Item No.)	Lot No.	Purity
6PPD	793-24-8 (DTXSID9025114)	268.40	HPC Standards (687875)	819484	98.09%
6PPD-quinone	2754428-18-5 (DTXSID301034849)	298.40	HPC Standards (687855)	818631	98.65%
5,8,11-Eicosatriynoic acid	13488-22-7 (DTXSID10159018)	300.40	Cayman Chemical (90200)	0666022-3	100%
Cucurbitacin I	2222-07-3 (DTXSID501015546)	514.70	Cayman Chemical (14747)	0504456-9	100%
7PPD	3081-01-4 (DTXSID5027516)	282.40	Alfa Chemistry (ACM3081014)	A24Q1828HL	96%
7PPD-quinone	2894124-00-4 (NODTXSID)	312.40	Cayman Chemical (40610)	0713246-3	97.6%
IPPD	101-72-4 (DTXSID1025485)	226.32	HPC Standards (682900)	822474	98.01%
IPPD-quinone	68054-73-9 (DTXSID401352921)	256.30	HPC Standards (689129)	827038	98.24%
77PD	3081-14-9 (DTXSID2024618)	304.51	HPC Standards (688170)	822486	98.38%
77PD-quinone	None (DTXSID801352923)	334.50	HPC Standards (689879)	813463	96.74%
CPPD	101-87-1 (DTXSID2051508)	266.38	HPC Standards (690149)	823710	91.69%
CPPD-quinone	68054-78-4 (DTXSID101352922)	296.36	HPC Standards (689878)	813102	99.26%
DPPD	74-31-7 (DTXSID9020538)	260.33	HPC Standards (677669)	809065	99.19%
DPPD-quinone	3421-08-7 (DTXSID10279149)	290.32	HPC Standards (689475)	809858	93.30%
CCPD	4175-38-6 (DTXSID8063335)	272.43	Alfa Chemistry (ACM4175386)	A24Q18281AQ2	95%
DTPD	68953-84-4 (DTXSID2021444)	288.39	HPC Standards (690038)	827819	99.58%
o-DTPD-quinone	252950-56-4 (NODTXSID)	318.37	HPC Standards (689877)	813103	99.94%
p-DTPD-quinone	10015173-46-5 (NODTXSID)	318.37	HPC Standards (690037)	817458	97.82%
Irganox 1076	2082-79-3 (DTXSID8027456)	530.86	Sigma-Aldrich (367079)	MKCW0680	99.3%
Irganox 1520	110553-27-0 (DTXSID6044199)	424.75	TCI (M3227)	RA86G-AK	97.3%
Octyl gallate	1034-01-1 (DTXSID4040713)	282.34	Sigma-Aldrich (PHR3511)	LRAC7992	99.6%
Styrenated phenol	61788-44-1 (DTXSID8028052)	406.57	Sigma-Aldrich (A215577)	A215577-QH4	99%

Table S4. RTgill-W1 growth media and minimal media components.

Component	Stock Concentration	Final Concentration
RTgill-W1 Growth Media		
Leibovitz's L-15 medium with no phenol red		--
Heat-inactivated fetal bovine serum		5% V/V
Penicillin/Streptomycin		1% V/V
RTgill-W1 Exposure Minimal Media		
Deionized water		--
Salt Solution A		6% V/V
Sodium chloride	2.28 M	137 mM
Potassium chloride	89.4 mM	5.4 mM
Magnesium sulfate	13.5 mM	0.8 mM
Deionized water	--	--
Salt Solution B		1% V/V
Calcium chloride	126 mM	1.26 mM
Deionized water	--	--
Salt Solution C		3% V/V
Disodium phosphate	44.6 mM	1.34 mM
Potassium dihydrogen phosphate	14.7 mM	0.441 mM
Deionized water	Fisher Scientific	BD 309646
Galactose solution		1% V/V
Galactose	0.5 M	5 mM
Deionized water	--	--

Table S5. Cell Painting and cell viability stain components.

Reagent	Stock Concentration	Labelling Solution Concentration	Final Well Concentration	Amount per 384 well plate (including dead volume)
Cell Painting Live Labeling Solution				1,000 µL
MitoTracker DeepRed	1 mM	10 µM	0.48 µM	10 µL
RTgill-W1 Minimal Media	--	--	--	990 µL
Cell Viability Live Labeling Solution				992 µL
Hoechst 33342	16.2 mM	202.5 µM	9.64 µM	11.53 µL
Propidium Iodide	1495.11 µM	74.81 µM	3.56 µM	46.1 µL
RTgill-W1 Minimal Media	--	--	--	864.4 µL
Cell Painting Full Staining Solution				1,000 µL
Hoechst 33342	16.2 mM	40.5 µM	1.93 µM	2.5 µL
Concanavalin A	1 µg/µL	600 µg/µL	28.6 µg/mL	600 µL
SYTO14	5 mM	18 µM	0.86 µM	3.6 µL
Phalloidin	6.6 µM	247.2 nM	11.8 nM	37.6 µL
Wheat Germ Agglutinin	1 mg/mL	30 µg/mL	1.43 µg/mL	30 µL
1% Bovine Serum Albumin in 1X PBS	--	--	--	326.4 µL

Supplemental Files

File S1. 2D chemical structures for all antiozonant compounds tested in this study.

File S2. List of features used for profile correlation analysis. This spreadsheet contains the columns: “channel”, referring to the relevant fluorescent channel (or shape/position if applicable); “module”, referring to the type of cellular morphology measured; “compartment”, referring to the location of the feature. All other columns are shared with those from File S3.

File S3. Cell Painting and cell viability concentration-response curves for each chemical tested. Each page contains the curve corresponding to the relative cell count endpoint (black squares, upper left), the propidium iodide endpoint (gold triangles, upper right), the most sensitive category-level Mahalanobis distance endpoint (pink diamonds, lower left), and the global Mahalanobis distance endpoint (purple circles, lower right). Cell viability responses are clipped for cytotoxicity as described in Method S2.

File S4. Curve-fitting results for cell viability. A table containing results from concentration-response modeling of relative cell count from the HTPP data using tcplfit2[4]. This table contains the following columns:

Data Column Name	Data Type	Description	Example(s)
assay	string	An acronym used to denote which assay a sample was generated from.	HTPP CV
pg_id	string	In an HTPP study, the chemical set is divided into sub-groups according to the number of chemicals that can be tested on an assay plate (typically 40 to 42 chemicals in 8-point dilution series). Each chemical sub-group is tested in several assay plates originating from independent cultures. Assay plates that receive the same sub-group of chemicals belong to the same “plate group”. Data from all plates in a plate group are combined for concentration-response modeling	AltAntiozonants
stype	string	The sample type. This value denotes different types of samples on the assay plate: e.g., “vehicle control” (DMSO wells), “reference chemical”, “test sample”.	vehicle control, reference chemical, test sample, viability positive control
chem_id	string	Chemical identifier, for this study a shortened version of the chemical name or abbreviation.	Eico 6PPD 77PD-quinone
min_conc	float	The minimum concentration of test chemical included in concentration-response modeling.	0.03 1
max_conc	float	The maximum concentration of test chemical included in concentration-response modeling.	100 30
n_conc	integer	The number of dose levels included in concentration-response modeling.	8 7
ctr_mean	float	The mean of vehicle control values used for concentration-response modeling.	1.953546

ctr_sd	float	The standard deviation of vehicle control values used for concentration-response modeling.	0.7636382
approach	string	A description of the type of endpoint that was subject to concentration-response modeling. The three possible values are global, category and feature.	global category feature
endpoint	string	The name of the endpoint that was subject to concentration-response modeling.	rel_cell_count percent_responder_pi
n_gt_cutoff	float	Standardized output from the tcplfit2 R-package (https://cran.r-project.org/web/packages/tcplfit2/vignettes/tcplfit2-vignette.html).	Examples are provided in the vignette and documentation associated with the tcplfit2 R-package.
cutoff	float		
fit_method	string		
rmse	float		
er	float		
bmr	float		
hitcall	float		
conc	string		
resp	string		
top_over_cutoff	float		
tp	float		
p	float		
ga	float		
bmdl	float		
bmdu	float		
ac50	float		
top	float		
ac5	float		
ac10	float		
ac20	float		
acc	float		
ac1sd	float		
bmd	float		

File S5. Curve-fitting results for Cell Painting. A table containing results from concentration-response modeling of global Mahalanobis, category-level Mahalanobis and feature-level Cell Painting data using tcplfit2[4]. This table contains the following columns:

Data Column Name	Data Type	Description	Example(s)
pg_id	string	In an HTPP study, the chemical set is divided into sub-groups according to the number of chemicals that can be tested on an assay plate (typically 40 to 42 chemicals in 8-point dilution series). Each chemical sub-group is tested in several assay plates originating from independent cultures. Assay plates that receive the same sub-group of chemicals belong to the same “plate group”. Data from all plates in a plate group are combined for concentration-response modeling	AltAntiozonants

stype	string	The sample type. This value denotes different types of samples on the assay plate: e.g., “vehicle control” (DMSO wells), “reference chemical”, “test sample”.	vehicle control, reference chemical, test sample, viability positive control
chem_id	string	Chemical identifier, for this study a shortened version of the chemical name or abbreviation.	Eico 6PPD 77PD-quinone
min_conc	float	The minimum concentration of test chemical included in concentration-response modeling.	0.03 1
max_conc	float	The maximum concentration of test chemical included in concentration-response modeling.	100 30
n_conc	integer	The number of dose levels included in concentration-response modeling.	8 7
ctr_mean	float	The mean of vehicle control values used for concentration-response modeling.	1.953546
ctr_sd	float	The standard deviation of vehicle control values used for concentration-response modeling.	0.7636382
approach	string	A description of the type of endpoint that was subject to concentration-response modeling. The three possible values are global, category and feature.	global category feature
endpoint	string	the name of the endpoint that was subject to concentration-response modeling. Possible values are global, category names, and alphanumeric feature identifiers as described for File S3.	global, AGP_Axial_Cells f_1 f_100 f_3000
n_gt_cutoff	float	Standardized output from the tcplfit2 R-package (https://cran.r-project.org/web/packages/tcplfit2/vignettes/tcplfit2-vignette.html).	Examples are provided in the vignette and documentation associated with the tcplfit2 R-package.
cutoff	float		
fit_method	string		
rmse	float		
er	float		
bmr	float		
hitcall	float		
conc	string		
resp	string		
top_over_cutoff	float		
tp	float		
p	float		
ga	float		
bmdl	float		
bmdu	float		
ac50	float		
top	float		
ac5	float		
ac10	float		
ac20	float		
acc	float		
ac1sd	float		
bmd	float		

File S6. List of all Cell Painting features. Table with the detailed name of each phenotypic feature as output by the high content imaging instrument, a refined feature name suitable for use in the R statistical programming environment, a shorthand alphanumerical designation for each feature (e.g., f_10) and the phenotypic category that the feature belongs to. This table contains the following columns:

Data Column Name	Data Type	Description	Example(s)
feature_name_harmony	string	The original name of the features as they appear in the raw data export files from the imaging instrument.	Cells Non-Border - AGP_Cells_Morph_STAR Axial Length Ratio SER-Bright
feature_name_r	string	A modified name of the features that is compatible with the R statistical computing environment. Does not contain spaces or special characters.	AGP_Cells_Morph_STAR_ Axial_Length_Ratio_SER-Bright
column_type	string	M = metadata or F = Feature	M, F
feature_id	integer	A numerical designation for each feature.	1,2,…,1294
feature_name_mongo	string	A shorthand alphanumerical designation for each feature.	f_1, f_2,…,f_1294
category_name_r	string	The name of the category a feature belongs to. Describes the fluorescent channel, feature type and cellular compartment from which the feature was derived. Used as the basis for category-level Mahalanobis distance modeling.	AGP_Axial_Cells Mito_Profile_Nuclei

File S7. Normalized Cell Painting well-level values. A table containing well-level feature data that has been z-scaled to vehicle control well across all plates in a plate group. Data are z-scaled by plate group. Each row in the table is a unique sample. The first several columns in the table are metadata used to identify and analyze the sample. The last several hundred columns of the table are the z-scaled feature data. This table also contains data from a series of NULL chemicals constructed from the two lowest concentrations of test chemicals in each plate group. These NULL data are subject to concentration-response modeling in the same manner as the authentic test chemicals and is used to evaluate false positive rate when criteria for defining active chemicals is applied to either global Mahalanobis, category-level Mahalanobis or feature level concentration-response modeling results. This table contains the following columns:

Data Column Name	Data Type	Description	Example(s)
replicate_num	integer	A numerical value associated with the culture replicate. There are typically 3-4 culture replicates in each HTPP experiment.	1,2,3,4
culture_id	string	An identifier for the culture replicate associated with an assay plate. Typically begins with a	c2022-02-07

		lowercase “c” followed by a YYYY-MM-DD notation.	
assay	string	An acronym used to denote which assay a sample was generated from.	HTPP CV
pg_id	string	In an HTPP study, the chemical set is divided into sub-groups according to the number of chemicals that can be tested on an assay plate (typically 40 to 42 chemicals in 8-point dilution series). Each chemical sub-group is tested in several assay plates originating from independent cultures. Assay plates that receive the same sub-group of chemicals belong to the same “plate group”. Data from all plates in a plate group are combined for concentration-response modeling.	AltAntiozonants
doseplate_id	string	An identifier for the dose plate that was used to dispense a treatment. Typically, a barcode.	LAB-000281
plate_id	string	An identifier for the assay plate. Typically, a barcode.	TC00001234
stype	string	The sample type. This value denotes different types of samples on the assay plate: e.g., “vehicle control” (DMSO wells), “reference chemical”, “test sample”.	vehicle control, reference chemical, test sample, viability positive control
chem_id	string	Chemical identifier, for this study a shortened version of the chemical name or abbreviation.	Eico 6PPD 77PD-quinone
dose_level	integer	A numerical index corresponding to ascending dose levels within a study. For an eight-point dilution series, 1 represents the lowest tested concentration and 8 represents the highest tested concentration of a chemical. Vehicle control wells are represented by a zero.	0,1,2,3,…,8
conc	float	The chemical test concentration. Should be a positive numerical value. Vehicle control wells are assigned a value of 0.	100, 30, 0
conc_units	string	The units associated with the chemical test concentration. Units commonly used in HTPP studies are µM or µg/mL	µM, µg/mL
trt_name	string	The concatenation of chem_id and dose_level.	6PPD_7
sample_id	string	Unique identifier assigned to a sample. It is the concatenation of the plate_id and well_id.	TC00001234_K11
well_id	string	A well identifier associated with the position of the sample on the assay plate. First character is the plate row (alphabetical). Second two digits are the plate column (numerical).	A01 K11 P24
n_fields	integer	The number of unique fields of view imaged in the well associated with the sample.	3,2,1
n_cells_total	integer	Total number of cells identified across all unique fields of view that were imaged in an assay well.	1041 23 767
n_cells_keep	integer	The total number of cells retained in the analysis follow application of filtering criteria in	1041 22

		during data pipelining. Often n_cells_total equals n_cells_keep; n_cells_total should never exceed n_cells_keep.	763
rel_cell_count	float	The number of cells (n_cells_keep) in an assay well relative to the median number of cells counted in vehicle control wells on the same assay plate, expressed as a percentage.	Values typically range from 50 to 150. In cases of overt cytotoxicity rel_cell_count will be very low.
f_1, f_2,…,f_1300	float	Z-scaled well-level feature data. Values can be positive or negative and are reported to three decimal places.	7.156 -4.362

File S8. Global Mahalanobis distance values. Table containing the global Mahalanobis distances calculated from the HTPP data. Each row in the table is a unique sample. The table includes results for test chemicals, reference chemicals, vehicle control wells, viability positive control wells and null chemicals. This table contains the following columns:

Data Column Name	**Data Type**	**Description**	**Example(s)**
replicate_num	integer	These data columns are the same as described for File S7.	Example of values for these data columns are provided for File S7.
culture_id	string		
pg_id	string		
doseplate_id	string		
stype	string		
chem_id	string		
dose_level	string		
conc	string		
conc_units	string		
sample_id	string		
plate_id	string		
well_id	string		
trt_name	string		
assay	string		
n_fields	integer	The number of unique fields-of-view in the assay well.	3, 2
n_cells_total	integer	The total number of cells identified across all unique fields-of-view in an assay well.	1365 1280
n_cells_keep	integer	The total number of cells retained for analysis, summed across all unique fields-of-view in an assay well.	1365 1276
rel_cell_count	float	Relative cell count calculated as the number of cells retained for analysis (n_cell_keep) divided by the mean number of cells retained for analysis from vehicle control wells x 100.	113.8 91.7 8.8
n_features	integer	The number of phenotypic features measured and reported for the assay well	1300
n_pc	integer	The number of principal components used in calculation of the global Mahalanobis distance values	65 47 8
d	float	The global Mahalanobis distance value for this assay well.	178.466 35.782

			4.33

File S9. Category-level Mahalanobis distance values. Table containing the category-level Mahalanobis distances calculated from the HTPP data. Each unique sample within the table will have 49 rows corresponding to each of the 49 HTPP categories. The table includes results for test chemicals, reference chemicals, vehicle control wells, viability positive control wells and null chemicals. This table contains the following columns:

Data Column Name	Data Type	Description	Example(s)
replicate_num	integer	These data columns are the same as described for File S7.	Example of values for these data columns are provided for File S7.
culture_id	string		
pg_id	string		
doseplate_id	string		
stype	string		
chem_id	string		
dose_level	string		
conc	string		
conc_units	string		
sample_id	string		
plate_id	string		
well_id	string		
trt_name	string		
assay	string		
n_fields	integer	The number of unique fields-of-view in the assay well.	3, 2
n_cells_total	integer	The total number of cells identified across all unique fields-of-view in an assay well.	1365 1280
n_cells_keep	integer	The total number of cells retained for analysis, summed across all unique fields-of-view in an assay well.	1365 1276
rel_cell_count	float	Relative cell count calculated as the number of cells retained for analysis (n_cell_keep) divided by the mean number of cells retained for analysis from vehicle control wells x 100.	113.8 91.7 8.8
category_name_r	string	The name of the category a feature belongs to. Describes the fluorescent channel, feature type and cellular compartment from which the feature was derived. Used as the basis for category-level Mahalanobis distance modeling.	AGP_Axial_Cells Mito_Profile_Nuclei
n_features	integer	The number of features belonging to a category	20, 40, 9
n_pc	integer	The number of principal components used in calculation of the category Mahalanobis distance values	3, 6, 12
d	float	The category Mahalanobis distance value for this assay well.	178.466 35.782 4.33

Supplemental References

(1) Nyffeler, J.; Harris, F. R.; Willis, C.; Byrd, G.; Blackwell, B.; Escher, B. I.; Kasparek, A.; Nichols, J.; Haselman, J. T.; Patlewicz, G.; et al. A combination of high-throughput in vitro and in silico new approach methods for ecotoxicology hazard assessment for fish. *Environ Toxicol Chem* **2025**. DOI: 10.1093/etojnl/vgae083 From NLM Publisher.
(2) Nyffeler, J.; Willis, C.; Lougee, R.; Richard, A.; Paul-Friedman, K.; Harrill, J. A. Bioactivity screening of environmental chemicals using imaging-based high-throughput phenotypic profiling. *Toxicol Appl Pharmacol* **2020**, *389*, 114876. DOI: 10.1016/j.taap.2019.114876 From NLM Medline.
(3) Nyffeler, J.; Haggard, D. E.; Willis, C.; Setzer, R. W.; Judson, R.; Paul-Friedman, K.; Everett, L. J.; Harrill, J. A. Comparison of Approaches for Determining Bioactivity Hits from High-Dimensional Profiling Data. *SLAS Discov* **2021**, *26* (2), 292-308. DOI: 10.1177/2472555220950245 From NLM Medline.
(4) *tcplfit2: A Concentration-Response Modeling Utility*; 2024. https://CRAN.R-project.org/package=tcplfit2 (accessed 2024/10/08).
(5) Nyffeler, J.; Willis, C.; Harris, F. R.; Foster, M. J.; Chambers, B.; Culbreth, M.; Brockway, R. E.; Davidson-Fritz, S.; Dawson, D.; Shah, I.; et al. Application of Cell Painting for chemical hazard evaluation in support of screening-level chemical assessments. *Toxicol Appl Pharmacol* **2023**, *468*, 116513. DOI: 10.1016/j.taap.2023.116513 From NLM Medline.