

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/259472871>

Benchmarking Organic Micropollutants in Wastewater, Recycled Water and Drinking Water with In Vitro Bioassays

ARTICLE in ENVIRONMENTAL SCIENCE & TECHNOLOGY · DECEMBER 2013

Impact Factor: 5.33 · DOI: 10.1021/es403899t · Source: PubMed

CITATIONS

42

READS

331

40 AUTHORS, INCLUDING:



Peter A Bain

The Commonwealth Scientific and Industrial ...

26 PUBLICATIONS 220 CITATIONS

[SEE PROFILE](#)



Klára Hilscherová

Masaryk University

127 PUBLICATIONS 2,692 CITATIONS

[SEE PROFILE](#)



Barbora Jarosova

Masaryk University

9 PUBLICATIONS 197 CITATIONS

[SEE PROFILE](#)



Erik Prochazka

Griffith University

9 PUBLICATIONS 65 CITATIONS

[SEE PROFILE](#)

Benchmarking Organic Micropollutants in Wastewater, Recycled Water and Drinking Water with In Vitro Bioassays

Beate I. Escher,^{*,†} Mayumi Allinson,^{‡,§} Rolf Altenburger,^{||} Peter A. Bain,[⊥] Patrick Balaguer,[#] Wibke Busch,^{||} Jordan Crago,[▽] Nancy D. Denslow,[○] Elke Dopp,[◆] Klara Hilscherova,[¶] Andrew R. Humpage,[☆] Anu Kumar,[†] Marina Grimaldi,[#] B. Sumith Jayasinghe,[○] Barbora Jarosova,^{||} Ai Jia,[■] Sergei Makarov,[□] Keith A. Maruya,[◊] Alex Medvedev,[□] Alvine C. Mehinto,[◊] Jamie E. Mendez,[⊗] Anita Poulsen,^{†,●} Erik Prochazka,[■] Jessica Richard,[◆] Andrea Schifferli,[⊗] Daniel Schlenk,[▽] Stefan Scholz,^{||} Fujio Shiraishi,[§] Shane Snyder,[■] Guanyong Su,[∞] Janet Y. M. Tang,[†] Bart van der Burg,[○] Sander C. van der Linden,[●] Inge Werner,[⊗] Sandy D. Westerheide,[⊗] Chris K. C. Wong,[■] Min Yang,[ⓐ] Bonnie H. Y. Yeung,[■] Xiaowei Zhang,[∞] and Frederic D. L. Leusch[■]

[†]The University of Queensland, National Research Centre for Environmental Toxicology (Entox), 39 Kessels Rd, Brisbane, QLD 4108, Australia

[‡]Centre for Aquatic Pollution Identification and Management (CAPIM), School of Chemistry, The University of Melbourne, 1-100 Grattan Street, Parkville, VIC 3010 Australia

[§]National Institute for Environmental Studies, 16-2 Onogawa, Tsukuba 305-8506 Japan

^{||}UFZ - Helmholtz Centre for Environmental Research, Department of Bioanalytical Ecotoxicology, Permoserstr. 15, 04318 Leipzig, Germany

[⊥]CSIRO Land and Water, Private Bag No. 2, Glen Osmond, SA 5064, Australia

[#]Cancer Research Institute Montpellier, CRLC Val, d'Aurelle, Parc Euromédecine, 34298 Montpellier, Cedex 5, France

[▽]Aquatic Ecotoxicology, Department of Environmental Sciences, University of California, Riverside, Riverside, California 92521, United States

[○]University of Florida, Department of Physiological Sciences, PO Box 110885, Gainesville, Florida 32611, United States

[◆]IWW Water Centre, Department of Toxicology, Moritzstrasse 26, 45476 Mülheim/Ruhr, Germany

[¶]Research Centre for Toxic Compounds in the Environment (RECETOX), Masaryk University, Kamenice 753/5, 62500, Brno, Czech Republic

[☆]Australian Water Quality Centre, 250 Victoria Square, Adelaide SA 5001, Australia

[■]University of Arizona, 1133 E. James E. Rogers Way, Harshbarger 108, Tucson, Arizona 85721-0011, United States

[□]ATTAGENE, PO Box 12054, Research Triangle Park, North Carolina 27709, United States

[◊]Southern California Coastal Water Research Project Authority (SCCWRP), 3535 Harbor Blvd., Suite 110, Costa Mesa, California 92626-1437, United States

[⊗]Department of Cell Biology, Microbiology and Molecular Biology, University of South Florida, 4202 E Fowler Ave, ISA 2015, Tampa, Florida 33620, United States

[■]Smart Water Research Centre, Griffith University, Edmund Rice Dr, Griffith University Gold Coast Campus, Southport, QLD 4222, Australia

[⊗]Swiss Centre for Applied Ecotoxicology, Eawag-EPFL, Überlandstr. 133, 8600 Dübendorf, Switzerland

[∞]State Key Laboratory of Pollution Control and Resources Reuse, School of the Environment, Nanjing University, Nanjing 210046, PR China

[●]BioDetection Systems, Science Park 406, 1098 XH, Amsterdam, The Netherlands

[■]Department of Biology, Croucher Institute for Environmental Sciences, Level 10, Science Tower, Ho Sin Hang Campus, Hong Kong Baptist University, 224 Waterloo Rd. Kowloon Tong, Kowloon, Hong Kong

[ⓐ]State Key Laboratory of Environmental Aquatic Chemistry, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, PO Box 2871, Beijing 100085, PR China

Supporting Information

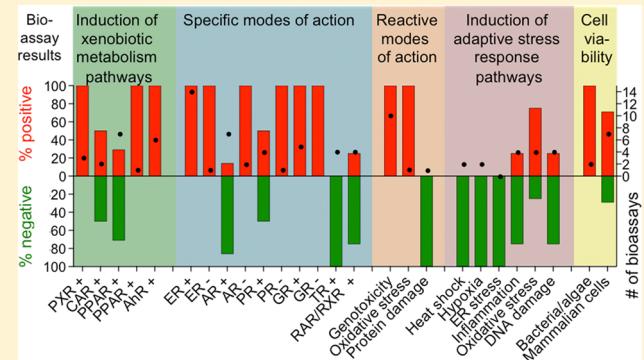
Received: September 2, 2013

Revised: December 9, 2013

Accepted: December 10, 2013

Published: December 10, 2013

ABSTRACT: Thousands of organic micropollutants and their transformation products occur in water. Although often present at low concentrations, individual compounds contribute to mixture effects. Cell-based bioassays that target health-relevant biological endpoints may therefore complement chemical analysis for water quality assessment. The objective of this study was to evaluate cell-based bioassays for their suitability to benchmark water quality and to assess efficacy of water treatment processes. The selected bioassays cover relevant steps in the toxicity pathways including induction of xenobiotic metabolism, specific and reactive modes of toxic action, activation of adaptive stress response pathways and system responses. Twenty laboratories applied 103 unique *in vitro* bioassays to a common set of 10 water samples collected in Australia, including wastewater treatment plant effluent, two types of recycled water (reverse osmosis and ozonation/activated carbon filtration), stormwater, surface water, and drinking water. Sixty-five bioassays (63%) showed positive results in at least one sample, typically in wastewater treatment plant effluent, and only five (5%) were positive in the control (ultrapure water). Each water type had a characteristic bioanalytical profile with particular groups of toxicity pathways either consistently responsive or not responsive across test systems. The most responsive health-relevant endpoints were related to xenobiotic metabolism (pregnane X and aryl hydrocarbon receptors), hormone-mediated modes of action (mainly related to the estrogen, glucocorticoid, and antiandrogen activities), reactive modes of action (genotoxicity) and adaptive stress response pathway (oxidative stress response). This study has demonstrated that selected cell-based bioassays are suitable to benchmark water quality and it is recommended to use a purpose-tailored panel of bioassays for routine monitoring.



INTRODUCTION

The Tox21, a joint program of the National Institutes of Health, U.S. Environmental Protection Agency (EPA), and U.S. Food and Drug Administration¹ and the U.S. EPA ToxCast Program² aim to advance molecular toxicology, systems biology, and computational toxicology to overcome shortcomings of traditional *in vivo* toxicity testing of chemicals. Parallel initiatives exist in Europe, for example, the EU project ChemScreen.³ Jointly these programs bring a paradigm shift in toxicity testing as *in vitro* methods help elucidate mechanisms of toxicity, prioritize chemicals for further testing and develop predictive models in order to refine, reduce, or replace future *in vivo* testing. These programs rely heavily on high-throughput screening (HTS) using cell-based and cell-free *in vitro* bioassays of large numbers of chemicals to elucidate their toxicity pathways. While Tox21 focuses on the bioanalytical profiling of individual chemicals, these tools should also be applicable to environmental samples of unknown and complex composition, and this study brings together for the first time selected bioassays from Tox21 with established bioassays for water quality assessment.⁴

Cell-based bioassays have been developed to target all steps of the toxicity pathway (Figure 1).⁵ While the induction of xenobiotic metabolism may not lead to cytotoxicity, it is an indicator of the presence of pollutants. Some bioassays provide measures of mechanisms of toxicity by visualizing the interaction of stressors, for example, chemicals, with specific biological targets, for example, binding to endocrine receptors or reaction with DNA. The exposed cells may respond through induction of adaptive stress response pathways, a feature that can be used for the assessment of environmental pollutants, although it is not an adverse effect per se. Cell viability, growth and/or proliferation are indicators of adverse effects on a cellular level. If the cell represents a specialized tissue, this may give an indication of tissue-specific impairment.

Cellular responses will not always imply higher-level effects but are a prerequisite for whole organism and population responses.⁶ The direct detection of initiating key events in a

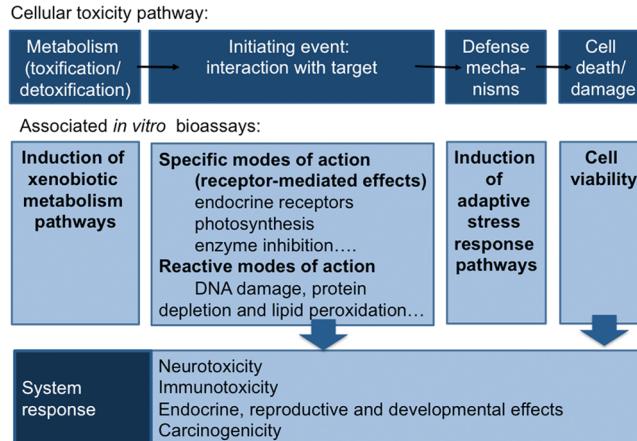


Figure 1. Classification of *in vitro* bioassays according to cellular toxicity pathways (adapted from refs 4, 5, and 117).

bioassay only provides a measure of a potential adverse effect because repair and defense mechanisms may ultimately prevent toxicity. From a precautionary perspective, however, the potential to cause an adverse effect is a crucial assessment endpoint. Cell-based bioassays are not suitable to replace regulatory *in vivo* tests but provide hazard information for screening and prioritization of chemicals.⁷

A large variety of cell-based bioassays have been applied for water quality assessment in the past⁴ but many of the end points evaluated under HTS toxicology programs⁸ are yet to be included. Many studies rely on a small set of bioassays and each study uses different types of water samples, sample preparation methods, bioassays, and data evaluation methods. Only a few cell-based bioassays have standardized protocols, such as the OECD or ISO guidelines. The goals of this study were to evaluate as many bioassays as practically achievable using one set of water samples with one sample preparation method and to recommend a screening test battery for water quality testing.

We selected cell-based bioassays using three criteria. First, we selected bioassays that have previously been used for water quality assessment. Second, a comprehensive literature review allowed us to identify cell-based bioassays that responded to environmentally relevant organic micropollutants but had not been used for water quality assessment prior to this study. Third, for final selection of endpoints, we screened 25 nuclear receptors (NR) and 48 transcription factor (TF) response elements in HepG2 human liver carcinoma cell lines⁹ to ensure inclusion of endpoints relevant for the particular samples tested.

We evaluated 10 samples, including nine ambient water samples ranging from effluent, recycled water to drinking water, plus one procedural blank. The samples were extracted and concentrated with an optimized solid phase extraction (SPE) method and sent to a total of 20 worldwide laboratories applying 103 bioassays for bioanalytical testing. All experimental data were evaluated using a common method specifically developed for this study to harmonize the different approaches to data evaluation. The results were used not only to validate and compare the different bioassays for application with water samples but also to benchmark water quality. The data may also serve to compare the efficacy of different water treatment processes for removal of organic micropollutants.

It was not the goal of this study to directly compare bioassay protocols and performance of bioassays but rather to obtain an overview of the biological endpoints responsive to typical water contaminants. One goal was to cover the major health-relevant toxicity pathways introduced in Figure 1 and to evaluate which pathways were relevant for water quality testing. The study was further expected to identify the most robust and responsive bioassays, thus, not only mammalian but also bacterial assays were included. The outcome includes recommendations on the makeup of a screening test battery and on indicator bioassays that appear to be particularly relevant for further investigation in water quality monitoring programs.

MATERIALS AND METHODS

Samples. Ten grab samples of water were collected in December 2011 and January 2012 (Supporting Information (SI), Section S1 and Table S1). Sample Eff1 is a secondary treated sewage effluent (activated sludge treatment) that serves as influent to a Water Reclamation Plant (WRP) that produces high quality recycled water for indirect potable reuse. Three samples were taken at different stages of treatment: after microfiltration (MF), reverse osmosis (RO) and advanced oxidation (AO) using H₂O₂/UV. Bioanalytical assessments had been previously undertaken at this WRP.^{10,11}

The second investigated WRP treats secondary sewage effluent (Eff2) by ozonation followed by biologically activated carbon filtration (O₃/BAC) to produce recycled water for irrigation and industrial usage. The fate of micropollutants in this plant has been previously characterized in more detail.^{12–14}

River water (RW) and drinking water (DW) samples were collected at the inlet and outlet of a metropolitan drinking water treatment plant applying chlorination and chloramination, which was also previously assessed with bioanalytical tools.^{10,15}

The stormwater sample (SW) was collected from a stormwater drain in Brisbane, Australia, that receives runoff from a residential catchment.¹⁶ The laboratory blank consisted of ultrapure water (Milli-Q water) run through the same SPE process as the samples. Fourteen liters were collected for each of Eff1, MF, Eff2, and SW, while 28 L were collected for the remaining samples.

Sample Preparation and Distribution to the Participating Laboratories. The SPE was performed according to Macova et al.¹⁰ using the sorbent materials Oasis HLB (Waters) followed by Supelclean coconut charcoal (Sigma-Aldrich), a combination that was confirmed previously to extract a broad range of organic micropollutants.¹⁷ Details and information on sample preparation and the logistics of sample distribution are summarized in the SI, Section S1. All distributed samples were labeled with codes for blind sample processing.

Sample Characterization. 293 organic micropollutants were previously characterized in these samples.¹⁸ Dissolved organic carbon concentrations are reported in the SI, Table S1.

Bioassays. The majority of selected bioassays was based on mammalian, bacterial or yeast cells. A zebrafish embryo test was included because fish embryos are considered as nonprotected life stages and as ethically accepted alternative to the testing of (adult) animals.¹⁹ Only one bioassay employed a naked enzyme (acetylcholinesterase inhibition assay).

All applied bioassays and their associated experimental methods are listed in Table 1 and categorized according to the toxicity pathways outlined in Figure 1. For bioassays where the protocol was modified or applied for the first time, Section S2 in the SI and Table S2 give additional information on the experimental procedures.

Concentration-Effect Assessment. A critical aspect when working with diverse biological endpoints is a consistent data evaluation process. It was the goal of the present study to harmonize the data evaluation as much as possible, which is challenging given the different types of endpoints measured but a prerequisite for quantitative comparison between different bioassays.

The concentrations of samples were expressed in units of relative enrichment factor REF. The REF is the product of the enrichment factor of the SPE process and the dilution of the extract in the bioassay (for derivation of equations, see SI, Section S3). A REF > 1 means that the sample is enriched in the bioassay (e.g., a REF of 10 means the sample was concentrated 10-fold in the bioassay), a REF < 1 means it was diluted in the bioassay, and a REF of 1 is equivalent to the organic micropollutants in the ambient (undiluted and unconcentrated) sample stripped from inorganics, metals and most colloidal organic matter by SPE.

For each bioassay, the measured responses were plotted against the sample concentration expressed as REF (Figure 2). For assays with a known maximum response, responses were converted to percent of maximum effect. For all endpoints that relate to cell viability and cell population growth, the controls can be expressed as 0% effect, while 100% relates to “no growth” or “all cells dead”. For reporter gene assays that measure the binding to a receptor or the transactivation of a receptor, 0% refers to the basal activity of the receptor, while 100% is defined using an appropriate reference compound that can saturate the receptor without causing cytotoxicity.

The ideal case would be a full concentration-effect curve covering 0% to 100% of effect (Figure 2A), which typically has a sigmoidal shape that can be described with a log-logistic equation. Any log-logistic concentration-effect curve will be linear with respect to (nonlogarithmic) concentrations at low effect level (up to 20–30% effect). As the water samples investigated in the present study often showed only very low effect levels, the linear form of the concentration-effect curves was used for derivation of the effect concentration causing 10% of maximum effect (EC₁₀) of the samples

Table 1. Bioassays Used and Their Toxicity Pathway Classifications^a

#	laboratory	bioassay	literature reference for method development	experimental approach (literature reference/more information in SI)
Xenobiotic Metabolism				
1	ATG	PXR-cisFACTORIAL	46	9
2	ATG	PXR-transFACTORIAL	46	9
3	IRCM	HGSLN PXR	47	48
4	ATG	CAR-transFACTORIAL	46	9
5	CAPIM	CAR-yeast	49	see SI, Section S2-A.
6	ATG	PPAR α -transFACTORIAL	46	9
7	ATG	PPAR γ -transFACTORIAL	46	9
8	IRCM	HELN-PPAR γ	50	28
9	BDS	CALUX-PPAR α	51	52
10	BDS, CSIRO	CALUX-PPAR γ 2	53	52
11	HK	MCF7-PPAR	54	54
12	GU	PPAR γ -GeneBLAzer	27	55
13	GU	Anti-PPAR γ -GeneBLAzer	27	see SI, Section S2-B.
14	CAPIM	AhR-yeast	56	57 and SI, Section S2-B.
15	UQ, RECETOX	CAFLUX	58	RECETOX: 59 UQ: 60
16	RECETOX	H4IIELuc	61	59
17	HK	MCF7DRE	54	54
18	ATG	AhR-cisFACTORIAL	46	9
19	UFZ	DART cyp1a induction	62	63
Specific Modes of Action				
20	UQ, SWISS	algae photosynthesis inhibition	31	64
21	UQ	acetylcholinesterase inhibition	65	30
Specific MOA: ER				
22	GU, CSIRO, BDS, IWW	ER-CALUX	66	33
23	UQ	E-SCREEN	67	10
24	SWISS, CSIRO, UA	YES	68	69
25	CAPIM	hER yeast	49	70
26	CAPIM	medER yeast	49	70
27	IRCM	HELN-ER α	71	72
28	IRCM	HELN-ER β	71	73
29	ATG	ERE-cisFACTORIAL	46	9
30	RECETOX	hER α -HeLa-9903	74	74
31	HK	MCF7-ERE	54	54
32	ATG	ER α -transFACTORIAL	46	9
33	NJU	Steroidogenesis (estrogens)	75	76
34	UFZ	DART cyp19a1b	77	63
35	UF, USF, UCR, SCCWRP	ER α -GeneBLAzer	27	55 see SI, Section S2-C.
36	CSIRO, GU	Anti ER-CALUX	66	17
Specific MOA: AR				
37	GU, BDS, CSIRO	AR-CALUX	66,78	33
38	IRCM	HELN-AR	79	80
39	HK	MCF7-ARE	54	54
40	UA, CSIRO	YAS	81	82
41	UF, USF, UCR, SCCWRP	AR-GeneBLAzer	27	55 see SI, Section S2-C.
42	ATG	AR-transFACTORIAL	46	9
43	RECETOX	MDA-kb2	83	84
44	RECETOX	Anti-MDA-kb2	83	84
45	CSIRO, GU	Anti-AR-CALUX	66,78	17
Specific MOA: GR				
46	GU, BDS, CSIRO	GR-CALUX	66	33
47	UA	GR Switchgear		see SI, Section S2-D.
48	ATG	GR-transFACTORIAL	46	9
49	RECETOX	GR-MDA-kb2	83	84
50	GU, UF, USF, SCCWRP	GR-GeneBLAzer	27	55
51	GU	Anti-GR-GeneBLAzer	27	
52	GU	Anti-GR-CALUX	66	33
Specific MOA: PR				
53	UF, USF, UCR, SCCWRP	PR-GeneBLAzer	27	55 see SI, Section S2-B.
54	GU, BDS, CSIRO	PR-CALUX	66	33

Table 1. continued

#	laboratory	bioassay	literature reference for method development	experimental approach (literature reference/more information in SI)
Specific MOA: PR				
55	GU	Anti-PR-CALUX	66	
56	NJU	steroidogenesis (progesterone)	75	76
57	NJU	steroidogenesis (17α OH-progesterone)	75	76
Specific MOA: TR				
58	BDS, GU	TR-CALUX	66	17
59	UQ	T-SCREEN	85	
60	ATG	THR α 1-transFACTORIAL	46	9
61	IRCM	HELN-TR		86
Specific MOA: Reproductive and Developmental Effects				
62	HK	MCF7-RARE	54	54
63	UQ	P19/A15	87	see SI, Section S2-E.
64	ATG	ROR β -transFACTORIAL	46	9
65	CAPIM	hRAR-Yeast Assay	49	88
Reactive MOA				
66	UQ, RCEES	umuC TA1535/pSK1002	89	UQ: 90 RCEES: 89
67	UQ	umuC TA1535/pSK1002 +S9	89	90
68	RCEES	umuC NM5004	91	89
69	RECETOX	SOS chromotest	92	93
70	UA, IWW	Ames TA98	94	IWW: 95
71	UA, IWW	Ames TA98+ S9	94	IWW: 95,96
72	UA	Ames TAmix	94	
73	UA	Ames TAmix +S9	94	
74	UQ, IWW	Ames TA100	94	UQ: 97 IWW: 95
75	AWQC	micronucleus assay	98	99
76	CSIRO	ROS formation RTG2	100	101
77	UQ	protein damage <i>E.coli</i>	102	38
Adaptive Stress Response				
78	ATG	HSE-cisFACTORIAL	46	9
79	UFZ	hspb11 induction DART	63	63
80	ATG	HIF-1a-cisFACTORIAL	46	9
81	UA	Hypoxia-Switchgear		see SI, Section S2-F.
82	ATG	NF-kB-cisFACTORIAL	46	9
83	UQ	NF-kB-Geneblazer	27	55 see SI, Section S2-C. 33
84	BDS	NF-kB-CALUX	103	
85	GU	Jurkat E6.1 IkB	none	see SI, Section S2-G.
86	UQ	AREc32	104	39
87	UA	Nrf2-keap	105	see SI, Section S2-H.
88	ATG	Nrf2/ARE-cisFACTORIAL	46	9
89	BDS	Nrf2-CALUX	106	52
90	ATG	p53-cisFACTORIAL	46	9
91	BDS	p53-CALUX	106	52
92	BDS	p53-CALUX +S9	106	52
93	UF	p53-GeneBLAzer	27	55 see SI, Section S2-C.
Cytotoxicity and Indicators of System Response				
94	UQ	AREc32 cell viability	39	39
95	GU	Caco 2 NRU	107	17 see SI, Section S2-I.
96	CSIRO	RTG2MTT	108	
97	UFZ	DART 48h lethality	109	110
98	UFZ	DART 120h sublethal	111	110
99	GU	SK-N-SH cytotoxicity	112	see SI, Section S2-J.
100	GU	THP1 cytokine	113	17 see SI, Section S2-K.
101	UQ	algae growth inhibition	114	64
102	UQ, SWISS	<i>Vibrio fischeri</i> (Microtox)	115	SWISS: 69 UQ: 18
103	RCEES	<i>Photobacterium phosphoreum</i>	116	

^aAssay numbers (#) are equivalent to the numbers in Figure 3. Literature references are provided for the method development and for how the assay was performed. Modifications of the assay methods are summarized in the Supporting Information, SI, Table S2.

(Figure 2C). A detailed derivation of EC₁₀ is provided in the SI, Section S3-A.

For assays where the maximum response was unknown or unachievable, the responses were normalized to growth medium

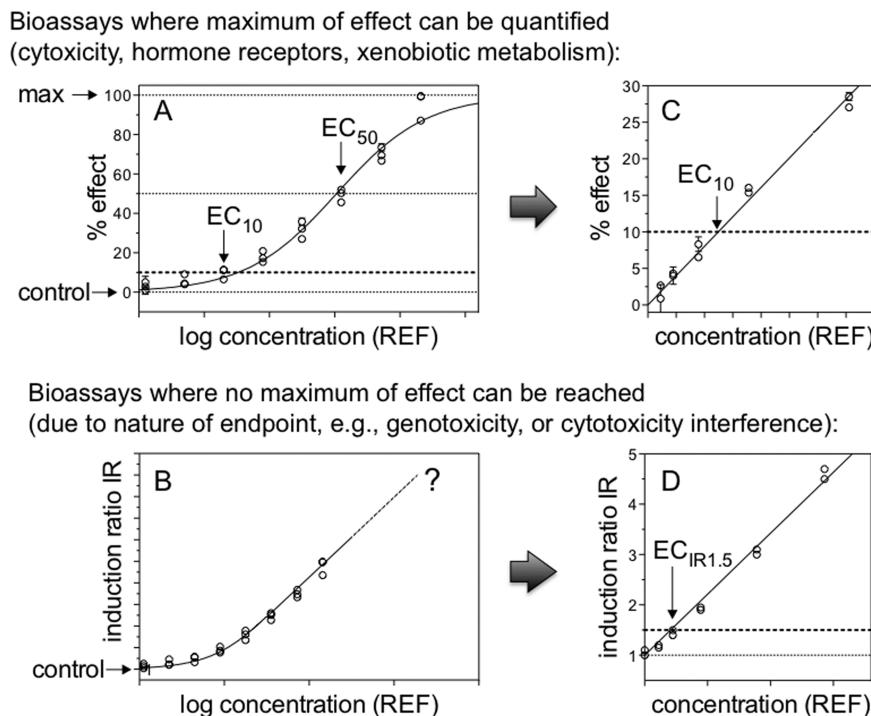


Figure 2. Overview of the concentration-effect models applied to derive benchmark effect concentrations (EC).

controls (including the same percentage of solvent as in the corresponding samples) and assessed via the induction ratio (IR; that is, fold induction relative to control). The maximum response was unknown if no appropriate reference compound existed, or if cytotoxicity quenched the reading of the reporter activity. The problem also arose if the endpoint is inducible or, by nature, there is no clear upper limit, for example, for DNA damage or adaptive stress response. In practice, linear regression through the control effect level (IR = 1) was used for derivation of the effect concentration that causes an IR of 1.5, or EC_{IR1.5} (Figure 2D). A detailed derivation of EC_{IR1.5} is provided in the SI, Section S3-B together with a discussion of the comparability of EC₁₀ versus EC_{IR1.5} (SI Table S3).

A bioassay can be run in antagonistic mode if the receptors are occupied with a constant concentration of a known and potent agonist. Varying concentrations of sample are added and if the signal of the control is suppressed the sample can be considered to exhibit an antagonistic effect. The effect concentration causing a suppression ratio SR of 0.2, EC_{SR0.2}, was used to describe all antagonistic effects and correspondingly an EC_{CD0.2} was defined for endpoints that are based on chaperone dissociation (e.g., I_KB dissociation from NF_κB). A detailed derivation is provided in the SI, Section S3-C and Figure S1.

Data Presentation. A heatmap presenting all measured EC values was generated using the R Software package gplots (www.r-project.org/). Hierarchical clustering was performed using the “complete linkage” method to find similar clusters of water samples.

RESULTS

Repeatability of Bioassays. A number of bioassays were performed simultaneously in multiple laboratories. As discussed in detail in the SI, Section S4 and Table S4, the results were consistent between laboratories and therefore the results of the same bioassays were averaged.

Initial Screening of Nuclear Receptors and Transcription Factors. The FACTORIAL bioassays⁹ were used for an initial profiling of the water samples after enrichment by SPE to a REF of 4. As discussed in the SI, Section S5 and Figure S2, an IR of 1.5 is regarded as the threshold for positive effects. In the SI, Figure S3, the activity profiles of all samples are depicted.

The blank did not induce any of the tested endpoints indicating that the sample extraction and enrichment process did not negatively influence the test outcome. The Eff1 and Eff2 samples caused an activation of five of 25 nuclear receptors (NR) and 5 of 48 transcription factors (TF) tested (SI Figure S3A and C). The active NRs were the pregnane X receptor (PXR), the peroxisome proliferator-activated receptor (PPAR γ), the estrogen receptor (ER α) and marginally the glucocorticoid receptor (GR) and the liver X receptor (LXR). The active TFs were related to the aryl hydrocarbon receptor (AhR), PXR, the oxidative stress response (nrf2/ARE), the estrogen response element (ERE), and the RAR-related orphan receptor (RORE).

The screening provides strong support to expand the test battery to include additional endpoints to those routinely employed.⁴ Particularly the PXR and AhR, which are related to xenobiotic metabolism, warrant more attention as these had the highest activity in the prescreening assay. The LXR is relevant as its activation induces the PPAR²⁰. The PPAR pathway is related to obesity²¹ and has gained much attention in recent years and therefore various PPAR-related endpoints have been included in the test battery if not specifically LXR. The oxidative stress response pathway appeared to be of high relevance and has rarely been investigated with water samples prior to this study.

In response to the findings of this screening, the active three NRs and five TFs and two others associated to relevant pathways (CAR, PPAR α , AR, GR, THR α 1, ROR β , HSE, HIF1a, NF_κB, p53) were included in the detailed dose-response analysis.

Responsiveness of the Bioassays. Figure 3, Table 2, and Table S5 in the SI give an overview of all results of the 101 different bioassays tested plus the two bacterial cytotoxicity assays.

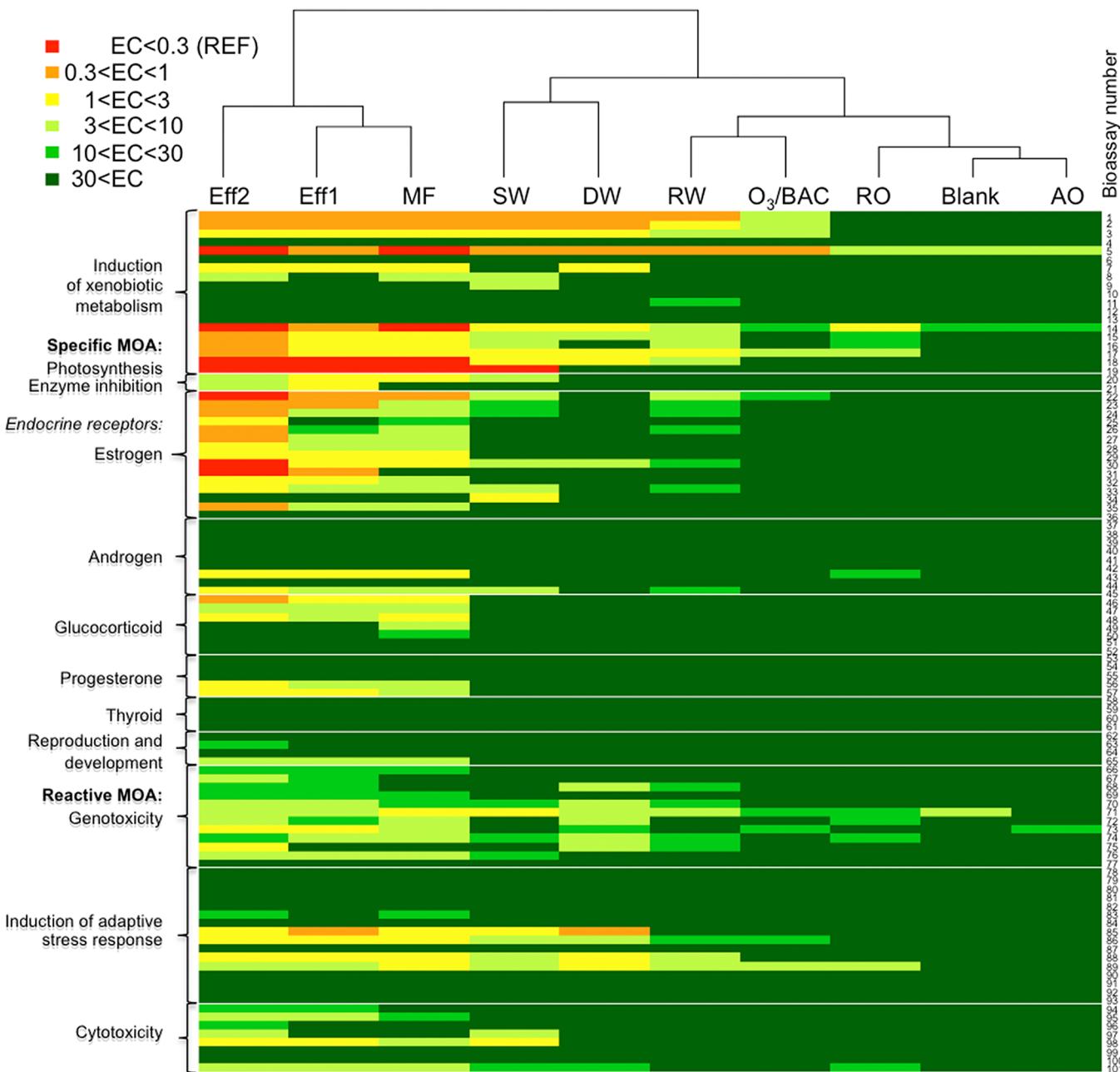


Figure 3. Summary of results in 101 bioassays (excluding inactive FACTORIAL and the bacterial cytotoxicity assays). Plotted are the effect concentrations (EC_{10} , $EC_{IRI,5}$, or $EC_{SR0,2}$) in units of REF (relative enrichment factors). The colors encode for the magnitude of the EC. Green stands for high effect concentrations (low potency) and transitions to red for low effect concentrations (high potency). Dark green are EC values that were >30 REF (which means that the sample that is enriched 30 times still does not show an effect), green from 10 to 30 REF, light green from 3 to 10 REF. A sample that has its EC at concentrations of the native sample up to three times enriched is denoted in yellow. Samples that have to be diluted for the EC are orange for up to 3 times diluted (REF 1 to 0.3) and red for over 3 times diluted. Numbers on the right refer to bioassay numbers in Table 1.

A positive response is defined here as $EC_{IRI,5}$ or EC_{10} lower than the highest tested REF, thus no extrapolations were performed.

The first two questions that we have to answer to judge the suitability of bioassays for water quality assessment are (A) “do “polluted” samples induce a response?” and (B) “is the response acceptably low in control samples?” (A) Sixty-five bioassays showed a response in at least one of the water samples. For Eff2, which can be considered as a moderately “polluted” sample, the number of positive results was 60. (B) No solvent blank caused any effect and the procedural blank with ultrapure water only produced effect in five bioassays (4.9%).

Procedural Blank. Even small impurities leaching out of the material or present in the solvent would likely contribute to the nonspecific effect of the blank. Here we applied two different SPE sorbent materials (HLB and coconut charcoal), which were eluted separately and required twice the amount of solvent. This consideration rationalizes the low but positive results of the blanks in the two bioluminescence inhibition assays with marine luminescent bacteria, which were higher than previously seen when only one type of solid-phase material were applied.¹⁰

Furthermore, the yeast-based assays AhR-yeast and CAR-yeast showed responses in the blanks but only at much higher REF

Table 2. Summary of Responsive and Non-Responsive Bioassays (Total = Number of Bioassays, In Parentheses Number of Replicates), + = Number of Positive Responses, - = Number of Negative Responses)

toxicity pathway	MOA	inducing chemicals/ positive controls	total	+	-	positive response	negative response
xenobiotic metabolism	pregnane X receptor (PXR) constitutive androstane receptor (CAR)	steroids/ phenobarbital, various pharmaceuticals	3 (1)	3 1	0 1	PXR-cisFACTORIAL, PXR-transFACTORIAL, HGSLN PXR CAR-yeast	-
	peroxisome proliferator-activated receptor (PPAR) aryl hydrocarbon receptor (AhR)	phthalates, fibre pharmaceuticals PAHs, PCDDs, coplanar PCBs	2 (1)	1	5	PPAR γ -transFACTORIAL, HELN-PPARY	CAR-transFACTORIAL PPAR α -transFACTORIAL, CALUX- PPAR α , CALUX-PPARY, PPARY GenBLAzer, MCF7-PPAR Anti-PPARY GeneBLAzer
	pPAR suppression	-	1	0	1	-	-
specific MOA	acetylcholinesterase (AChE) photosystem II	insecticides herbicides	1 (1)	0	1	-	AChE enzyme inhibition
specific receptor-mediated MOA	estrogen receptor (ER) androgen receptor (AR)	human hormones and industrial chemicals (xenoestrogens), 17 β -estradiol 4-Hydroxy-tamoxifen (Dihydro)-testosterone	14 (9)	14	0	ER-CALUX, E-SCREEN, YES, HELN ERA, HELN ERB, ERE-cisFACTORIAL, hER κ -HeLa-9903, MCF7-ER, ER α -transFACTORIAL, hER κ -HeLa-9903, MCF7-ER, (aromatase), ER-GeneBLAzer, hER yeast, medER yeast	-
	AR suppression	-	1 (1)	1	0	Anti-ER-CALUX	AR-CALUX, HELN-AR, MCF7-ARE, Yeast Androgen Screen (YAS), AR-GeneBLAzer, AR-transFACTORIAL
	PR suppression	-	7 (6)	1	6	MDA-kb2 (but coexpression with GR)	-
	GR suppression	-	2 (1)	2	0	Anti-AR-CALUX, anti MDA-kb2	PR-CALUX, PR-GeneBLAzer
	thyroid receptor (TR)	Flutamide Levonorgestrel Mifepristone Dexamethasone	4 (5)	2	2	Steroidogenesis, induction of progesterone and of 17 α OH-progesterone	-
RAR/RXR (Reproductive and development effects)	Mifepristone 3,3'-S-Triido-thyronine Retinoic acid	Mifepristone 3,3'-S-Triido-thyronine Retinoic acid	5 (6)	1	0	Anti-PR-CALUX	PR-CALUX, PR-GeneBLAzer
reactive modes of action	genotoxicity oxidative stress	4-Nitroquinoline-N-oxide PAH, electrophilic chemicals, <i>t-butyl</i> <i>hydroquinone</i> Sea-Nine	11 (4)	11	0	GR-CALUX, GR Switchgear, GR-transFACTORIAL, GR-MDA-kb2 (AR suppressed), GR-GeneBLAzer	-
	protein damage	oxygen depletion (can be caused by metals)	2	0	2	Anti-GR-CALUX, anti GR-GeneBLAzer	-
adaptive stress response pathway	heat shock response hypoxia	tunicamycin, caplain	1	0	1	TR-CALUX, T-SCREEN, THRaI-trans- FACTORIAL, HELN-TR MCF7-RARE, P19/A15, ROR β -trans- FACTORIAL,	-
	endoplasmic reticulum stress inflammation	high salt, glycol metals, PCBs, smoke, particles	4	1	3	umuC \pm S9, SOS chromotest, Ames \pm S9, micronucleus assay oxidative stress in RTG2 cells	protein damage <i>E.coli</i> GSH \pm
	oxidative stress	reactive oxygen species, <i>t-butyl</i> <i>hydroquinone</i>	4	3	1	HSE-cisFACTORIAL, hspb11 induction in DART after 120h HIF-1 α -cisFACTORIAL, Hypoxia- Switchgear	-
						NF-kB-CALUX, NF-1kB-GeneBLAzer, NF-kB-cisFACTORIAL	-
						Nrf2-keap	-

toxicity pathway	MOA	inducing chemicals/ positive controls	positive response			negative response
			total	+	-	
cytotoxicity indicative of system response	DNA damage	electrophilic chemicals, UV radiation, <i>nutlin-3</i>	4 (1)	0	4	p53-cisFACTORYL, p53-CALUX-59, p53-CALUX +s9, p53-GeneBLAzer
bacterial cytotoxicity		all	2 (1)	2	0	<i>Vibrio fischeri</i> (Microtox), <i>Photobacterium phosphoreum</i>
algal growth		all (+herbicides)	1 (1)	1	0	algae growth inhibition
cytotoxicity (mammalian cells)		all	4	4	0	AREC32 cell viability, Caco 2 NRU, RTG2MTT, DART 48h lethality
development		all	1	1	0	DART 120h sublethal
neurotoxicity		insecticides	1	0	1	-
immunotoxicity		immunosuppressive chemicals	1	0	1	SK-N-SH cytotoxicity
						THP1 cytokine

Table 2. continued

than the samples. A positive blank value was observed in one of the various Ames assays and is most likely due to measurement uncertainty as this value was derived from only one data point.

Bacterial Cytotoxicity Screening Assays. The bacterial cytotoxicity assays with *Vibrio fischeri* and *Photobacterium phosphoreum* were used as quick bioanalytical assessment tools, as these tests are rapid with only 15–30 min of exposure. Bioluminescence tests are nonspecific assays, as all stressors can impair the energy production and thus decrease bioluminescence. They provide a high responsiveness, that is, often indicate effects for the “diluted” sample at REFs of less than 1. However, the high sensitivity if compared to other cellular assays may result from a higher bioavailability, that is, absence of serum proteins typically used in *in vitro* assay. This increased bioavailability could result in a detection of trace amounts of coextracted dissolved organic carbon (DOC). In particular the low-molecular fraction of assimilable organic carbon can add to the observed effects as previously demonstrated.²² However, the high sensitivity may also result from specific interactions with bacterial physiology. Therefore, and because the control sample exhibited some effects as well, the luminescence assays have been excluded from the heatmap (see below).

Heatmap. The summary of 101 EC values (excluding the two bacterial cytotoxicity assays) in each of the 10 samples is presented in form of a heatmap (Figure 3, data summarized in SI, Table S5). The similarity of bioanalytical fingerprints between different water samples was characterized by hierarchical clustering.

Evidently, quantitative comparison is difficult because EC were expressed as EC₁₀ or EC_{IR1.5} and these two values are only directly comparable if the maximum IR is around 6 (see discussion in SI, Section S3-B). Therefore hierarchical clustering was only performed on samples and not on bioassays.

The closest similarity existed between the blank and the highly treated AO sample, while the RO sample clustered with this group on the next level of hierarchy (Figure 3). Surface water and ozonated recycled water clustered together. Both of these groups (cluster RW + O₃/BAC and cluster RO + Blank + AO) clustered closely on the next level of similarity. Of the more polluted samples, Eff1 and MF were highly similar. This is not unexpected, as microfiltration, the only treatment step separating the two samples, is ineffective at removing micropollutants. Slightly higher effects were, however, observed in the MF sample likely due to disinfection by chloramination of the membrane to avoid biofouling.¹⁰ On the next level of hierarchical clustering, the two WWTP effluents Eff1 + MF and Eff2 showed high similarity.

The largest separation was observed between the cluster of Eff1 + Eff2 + MF and all other samples (Figure 3), clearly demonstrating that cell-based bioassays can distinguish between wastewater and reclaimed water samples.

The bioanalytical fingerprints can also help distinguish between different water types: WWTP effluents not only showed the highest effects but also distinct responses related to known environmental pollutants, including pesticides, industrial chemicals, pharmaceuticals, and personal care products, for example, the activation of the aryl hydrocarbon receptor by PAHs or methylmercaptoaniline,²³ the activation of the estrogen receptor by natural hormones and xenoestrogens, the activation of the glucocorticoid receptor by dexamethasone and organotin compounds,²⁴ or photosynthesis inhibition by herbicides. The specific effects, caused by chemicals that bind to receptors, were decreased substantially in the WRPs.

Stormwater had a slightly different pattern to WWTP effluents but was also dominated by pesticides, as represented for example by herbicidal activity that was absent in other samples of its cluster.

In other studies on stormwater one could even identify sewer cross-contamination by bioanalytical profiling.¹⁶

In contrast, disinfection by chlorination lead to disinfection byproducts, which showed distinct bioassay response patterns with clearly increased genotoxicity and oxidative stress responses in DW as compared to its source water (RW), while specific-receptor mediated effects were low in RW and almost fully disappeared in DW (Figure 3). This is consistent with previous bioanalytical profiling of the drinking water treatment process.¹⁵

DISCUSSION

Multiplexed Assays As a Screening Tool. The FACTORIAL assay was applied here for the first time to water samples and yielded interesting fingerprints of effects that were consistent with the results of the other bioassays. However, more work is required to implement reference chemicals and include a more rigorous dose-response assessment. The effect fingerprints were qualitatively consistent with responses seen in the initial ToxCast I screening of 320 pesticides, where 73% of the pesticides were responsive in PXR, 52% in Nrf2/ARE and 46% in PPAR γ .⁹

Benchmarking Water Quality. A detailed discussion of the responses of each bioassay is provided in the SI, Section S6. Here follows a summary of the responsive and nonresponsive endpoints in relation to the associated step in the toxicity pathway (Table 2). It should be noted that, although responsiveness is related to assay sensitivity, even the most sensitive assay will not respond in the absence of chemicals capable of activating the bioassay endpoint. In other words, the absence of a measurable effect does not necessarily prove an assay insufficiently sensitive. Relative assay sensitivity can be assessed by comparing the effect concentrations and limits of detection of reference chemicals between assays. While such comparison was not possible for all assays tested in the present study, the results obtained can give some indication regarding the suitability of bioassays for monitoring purposes.

Induction of Xenobiotic Metabolism Pathways. Induction of metabolic pathways is not per se an indicator of toxicity but gives an indication of exposure to bioactive chemicals. Metabolism can detoxify or bioactivate chemicals. Omiecinski et al.²⁵ stressed the relevance and the toxicological implications of a number of xenobiotic metabolism pathways and associated NR, including the PXR, PPAR α , β and γ , AhR, and CAR.

Three and six bioassays were evaluated for the PXR and AhR, respectively, and all showed positive responses in less treated samples and negative responses in recycled water and the blank (Table 2, SI, Section S6-A, Figure S5).

CAR plays a role in both phase I and II metabolism and plays a protective role against toxicity induced by bile acids as well as regulation of physiological functions. The target chemicals for CAR are less clearly defined than for AhR and while a few pesticides (e.g., methoxychlor, carbaryl propazine, 6-deisopropylatrazine) induced the CAR in the CAR-trans-FACTORIAL assay in previous work,⁹ no response was detected in the water samples in the CAR-trans-FACTORIAL assay up to an REF of 4. In contrast, the CAR-yeast showed a response in all samples (EC_{IRL5} from 0.1 to 9.4) (Table 2, SI, Section S6-A, Figure S5).

For PPAR, only two of seven bioassays (PPAR γ -trans-FACTORIAL and HELN-PPAR γ) gave signals in the four most polluted samples (Table 2, SI, Section S6-A, Figure S5). PPAR is strongly linked to the regulation of glucose and lipid metabolism as well as inflammation, and is less important for xenobiotic metabolism.²⁶ In a high throughput study of 3000 environmentally relevant chemicals, roughly 1% of the tested chemicals were PPAR γ agonists and 8% were PPAR γ

antagonists.²⁷ Organotins²⁸ and polyhalogenated bisphenol A²⁹ were found to induce PPAR α and γ . The higher activity of PPAR γ over PPAR α for water samples is consistent with the finding that 146 of 309 ToxCast phase I chemicals were active in PPAR γ -trans-FACTORIAL, while the other isoforms were less responsive.⁹

Specific Modes of Toxic Action. Most specific modes of action involve binding to receptors or inhibition of enzymes. In the past, direct enzyme inhibition assays have been popular tools for water quality testing. Recent work on the influence of dissolved organic matter (DOM) on the acetylcholinesterase assay has demonstrated that DOM nonspecifically impacts the assay at relatively low concentrations.³⁰ The implication of these findings is that for most tests with naked enzymes, water samples cannot be concentrated above a REF of 2. In the present study, only the two wastewater samples produced a valid response in this assay. Despite the high relevance of this biological endpoint for many insecticides, it thus proves unsuitable to investigate recycled water samples.

Photosynthesis Inhibition. An important group of environmental contaminants are herbicides that inhibit photosynthesis. While they are specifically designed to target photosynthesis inhibition, herbicides can nevertheless be toxic to humans and are regulated in recycled and drinking water guidelines. The most sensitive species to detect herbicides are algae, for which the inhibition of photosystem II by triazines and phenylurea herbicides can be specifically measured by pulse-amplitude modulated fluorometry.³¹ This endpoint was very responsive in the water samples Eff1, Eff2, MF, SW that were suspected to contain herbicides (SI, Table S5).

Estrogen Receptor. The most relevant receptor-mediated effects are related to endocrine disruption (SI, Section S6-B, Figure S6). Estrogenic effects are by far the most prominent and environmentally relevant endocrine effects for aquatic species but they are overshadowed by other endocrine endpoints when it comes to human health. Fourteen different bioassays indicative of estrogenic effects were evaluated and all were active in four to five samples (Table 2). The absolute responsiveness was highest for ER-CALUX and MCF7-ERE but the effect pattern across the different samples was similar for all bioassays (SI, Table S5). No antiestrogenic activity could be detected in any of the samples, which is typical for samples that contain estrogenic chemicals.¹⁷

Androgen Receptor. Of seven bioassays (bioassay nos. 37–43, Table 1), only the MDA-kb2 produced positive results in the wastewater samples. Both GR and AR are expressed in this cell line and they share the same DNA response element, so it is unclear if the activity in this assay is purely AR-mediated, although incubation with flutamide indicates that the contribution of AR to the overall effect is higher than of GR.

Both bioassays for antiandrogenicity (anti-AR-CALUX and anti-MDA-kb2) were positive in some samples although only at very high REF. The WWTP effluents, which typically have highest antiandrogenic effects, were not responsive in anti-MDA-kb2, presumably due to the interfering agonistic response of effluent, similar to what has been observed with YAS and anti-YAS in an earlier study.³²

Progesterone Receptor. The two transactivation assays for the PR, PR-CALUX, and PR-GeneBLAzer, did not exceed the 10% effect threshold in all samples. However, progesterogenic activity has been detected previously in aquatic samples.^{17,33,34} The anti-PR (anti-PR-CALUX) assay was also negative with samples tested to a REF of 2. The increased levels of progesterone and 17 α -hydroxyprogesterone in the H295R bioassay

for steroidogenesis were most likely due to an inhibitory effect on CYP21A.

Glucocorticoid Receptor. ER, AR, and PR are important for the development and functioning of the reproductive system. The GR is more abundant and found in all cell types. Given that the GR has important functions in glucose metabolism and the immune feedback mechanisms, it has been linked to a wide spectrum of diseases, including cardiovascular, inflammatory and immune disease, diabetes, and obesity, and is therefore of high potential relevance. Five bioassays targeting activation of the GR were included in this study, all of which were active in one or more samples (SI, Figure S7). The observed GR activity is in agreement with previous studies on similar water types.^{17,33}

Thyroid Receptor. No assay indicative of modulation of the thyroid hormone system showed response to any of the water samples (SI, Table S5), which comes as no surprise because the most commonly observed thyroid agonists and goitrogens are oxyanions such as the perchlorate and nitrate.³⁵ These inorganic ions do not act via TR binding and are not extracted with SPE.

RAR/RXR. The retinoic acid signaling pathway is crucial for reproduction and development as well as for cell homeostasis and immune function.³⁶ Two receptors are key to this pathway, the retinoic acid receptor (RAR) and the retinoic X receptor (RXR). RXR is a heterodimer partner not only for the RAR but also for other nuclear receptors including PPAR, PXR, CAR, and TR.³⁷

We tested four bioassays that are connected to the retinoic acid-signaling pathway but only the two-hybrid assay, where RAR γ is inserted into yeast with lacZ as reporter gene, showed activity in three samples. The ROR β -transFACTORIAL did not show any response with the water samples tested, although this endpoint tested positive in 30% of the ToxCast I chemicals.⁹ Clearly, the role of RXR for water quality assessment should be further explored in the future.

Reactive Toxicity. Testing for reactive toxicity focused on genotoxicity and mutagenicity (Table 2, SI Figure S8). Only one bioassay, the micronucleus assay, detects DNA damage directly; the Ames test relies on reverse mutations and the umuC assay on detection of DNA repair. Three samples were active in the micronucleus assay, Eff2, RW and DW (SI, Figure S8). The SOS chromotest and umuC assays gave consistent results and were responsive at lower REFs but the Ames assay gave more variable responses and even false-positive responses (presumably due to the high inherent degree of endogenous gene mutation in bacteria).

Tests for genotoxicity can be run in the presence and absence of a rat liver metabolic enzyme mix (S9 fraction) to differentiate between chemicals that require metabolic activation and those that are detoxified by metabolism. In the umuC and the Ames assay, there was no discernible difference between response with and without S9.

The *E. coli* assay for protein damage relies on growth inhibition differences between a strain that is glutathione-deficient (GSH-) and the corresponding parent strain (GSH+).^{38,39} These assays were found to be unsuitable for samples with high organic matter content.³⁸ In the present study no effects could be detected although there appeared to be a qualitative difference in growth inhibition between the GSH- and the GSH+ strains.

Only one assay attempted to quantify reactive oxygen species formation in RTG2 cells and results were positive and

consistent with the activation of oxidative stress response pathway discussed below.

Induction of Adaptive Stress Response Pathways. Both the heat shock response and the hypoxia induction were negative in all assays tested (SI, Table S5). No bioassay for endoplasmic reticulum stress could be identified and therefore this potentially relevant endpoint had to be omitted.

Response to inflammation was tested by enzyme-linked immunosorbent assay (ELISA) in the human T-lymphoblast cell line Jurkat E6.1 by quantifying IkB, which is a chaperone for NF- κ B that keeps NF- κ B inactive and prevents it from entering the nucleus. Five samples tested positive in this assay (SI, Table S5). In contrast, the NF- κ B-CALUX, NF- κ B-GeneBLAzer and the NF- κ B-cisFACTORIAL did not respond to any of the samples. These latter assays are relatively new, have not yet been applied for water quality assessment and possibly require further validation work to improve their detection limits.

Three of four bioassays indicative of the oxidative stress response were active in six to eight samples, highlighting the potential importance of this stress response pathway. The AREc32, Nrf2/ARE-cisFACTORIAL and the Nrf2-CALUX were all able to detect effects at low sample enrichment. The data also showed a wide dynamic range between different samples, which makes them ideal water quality indicators, although their relevance to health effects is less evident than for other bioassays.

The p53 protein plays an important role as a tumor suppression factor but all evaluated assays did not show any effect, both, in presence and absence of S9.

General Cytotoxicity and Models for System Response. The overarching effect overlying each of the cellular toxicity pathways is cytotoxicity (Figure 1). As cytotoxicity normally manifests at higher concentrations than induction of response pathways, this endpoint is best implemented as a quality control measure for all induction assays to verify that cell vitality is not adversely affected. We did not complete full dose-response curves for cytotoxicity in all mammalian reporter gene assays apart from AREc32 for which cytotoxicity was similar to the targeted cytotoxicity assays in a human colon cancer cell line Caco 2 NRU. The fish cell line RTG2 was of relatively low responsiveness. Acute toxicity in the zebrafish embryo (DART 48h lethality) was only observed in two samples Eff2 and SW but at high enrichment factors (REF 5–6).

Cytotoxicity assays may also give information about system toxicity if appropriate cell lines are used, although a recent report by the ACuteTox project suggests little difference in response using different cell lines in vitro.⁴⁰ Here we considered the sublethal endpoint in the zebrafish embryo toxicity test after 120 h of incubation as an indicator of developmental and potentially long-term apical effects. This effect was clearly more responsive than the 48 h acute lethality endpoint in the zebrafish embryo.

The SK-N-SH neuroblastoma cell line⁴¹ is sensitive to chemicals that block the sodium channels and similar cell lines have been used previously in an assay to evaluate paralytic shellfish poisons caused by neurotoxic freshwater cyanobacteria.⁴² We used a simpler version of this assay in this study, evaluating cytotoxicity as a coarse measure of cellular neurotoxicity. This endpoint was not active for our water samples tested to a REF of 2.

Expression of various cytokines in the human acute monocytic leukemia cell line THP-1 gives an indication of potential immunotoxicity.¹⁷ Although a previous study reported

detectable inhibition of IL1 β secretion in chlorinated waters,¹⁷ we did not observe detectable effects in our study up to a REF of 2.

Benchmarking Treatment Efficacy. The bioassay results can be used to assess and monitor treatment processes (see SI, Section S7 for more detail). Best suited for this purpose are bioassays that show a clear decrease in response with increasing treatment and will not fall below the detection limit after treatment. We refer to these assays as “indicator bioassays” from hereon. Reverse osmosis (RO) is known as highly efficient in removing trace constituents, and only 13 indicator bioassays remained above detection limit after RO (but below LOD after AO, SI, Figure S10). After ozonation and BAC treatment, (another) 13 indicator bioassays remained above detection limit (SI, Figures S11 and S12). Between the two water reclamation plants, 18 suitable indicator bioassays were identified, including those indicative of AhR, PXR, CAR, ER, algal toxicity, genotoxicity and oxidative stress (SI, Figures S11 and S12).

In contrast, chlorination and chloramination increased the response in 15 of the 101 bioassays in drinking water samples (SI, Figure S13). The increased effect was most pronounced in the induction of xenobiotic metabolism and the reactive modes of action and oxidative stress response, which is consistent with the formation of chlorinated disinfection byproducts that cause genotoxicity and oxidative stress.^{15,43} This comparison demonstrates that there is no single battery of bioassays that can be applied universally, but rather that a panel of assays should be tailored to fit the needs of each application.

A Routine Test Battery of Indicator Bioassays. Because a single bioassay is not capable of assessing water quality comprehensively, a set of relevant biological endpoints that are sensitive to micropollutants typically encountered in water samples can be used collectively as indicators of water quality. A battery of bioassays should include integrative endpoints such as cytotoxicity as well as endpoints specific to relevant steps in the cellular toxicity pathway. As a minimum, indicator bioassays should cover examples found responsive to and representative of “induction of xenobiotic metabolism”, “endocrine disruption” and “adaptive stress responses”. Relevant endpoints are proposed in the following:

1. Induction of xenobiotic metabolism. Our results confirmed that activation of the aryl hydrocarbon receptor, already one of the most widely applied endpoints in water quality assessment, is a relevant indicator of the presence of chemicals and should be included in any routine test battery. The pregnane X receptor showed high responsiveness to water samples and responds to a wide range of chemicals and should be further explored for routine application.
2. Endocrine disruption. Specific receptor-mediated modes of action including estrogenic and androgenic effects showed the most promise for routine water quality screening applications. Recent work using GR-CALUX applied to various environmental chemicals and water samples³³ also support our findings that GR activity is present and could be detected in secondary treated effluent with the current battery of GR bioassays. In addition, given the co-occurrence of progestins and synthetic estrogens in hormone replacement therapy, PR activity remains of interest, despite the negative findings

here. Lastly, it is vital to test for antagonistic as well as agonistic effects.

3. Reactive modes of action. Genotoxicity as measured by well-established bioassays such as umuC or SOS chromotest served the purpose well. Bioassays derived from mammalian cells would be more relevant for human health and thus preferable to bacterial assays. As the p53 assays did not show the hoped-for responsiveness it is recommended to further evaluate alternative bioassays.
4. Adaptive stress response pathways. Oxidative stress response appears to be a highly sensitive and yet selective indicator of environmental pollution that responds to a wide range of chemicals as well as to transformation products and disinfection byproducts.⁴⁴ This is consistent with previous chemical testing in Nrf2/ ARE-cisFACTORIAL, where almost 50% of the ToxCast chemicals were active.⁹ Thus this mode of action is recommended to be included in any routine test battery, especially if transformation reactions are expected.
5. Cytotoxicity and systemic response. The bacterial cytotoxicity assays (*V. fischeri* and *P. phosphoreum*) are very fast and sensitive screening assays but their high sensitivity and effects caused by controls indicated that the responses may not be of human health relevance. In contrast, cytotoxicity assays with mammalian cells are comparatively less sensitive and clearly the bioassays for toxicity pathways are more relevant. With limitations, specific cell lines may be used as indicators of organ/systemic response. Nevertheless, further work has to be invested in the selection of appropriate tests systems and protocols for cell-based bioassays for organ/systemic responses as these are much less developed than nonspecific cytotoxicity assays and bioassays targeting cellular toxicity pathways. The enzymatic AChE inhibition assay to test for one aspect of neurotoxicity failed completely but there is the potential to implement neurotoxicity endpoints in the zebrafish embryo toxicity test.⁴⁵ A whole organism *in vitro* assay, such as the zebrafish embryo assay may help to link specific responses from the cellular assays to systemic responses by the observed phenotypes.

We also recommend that more attention be paid to the basal activities of cell lines in use. As metabolism is the most crucial modifier of toxicity, detoxifying many chemicals but activating others, the metabolic capacity of bioassay cell lines needs to be considered when selecting or designing a bioassay. Many available cell lines have low metabolic activity and for these it is advisable to run each experiment in parallel in the presence of an exogenous metabolic mixture, for example, liver S9 fraction.

In summary, an ideal battery of bioassays for water quality assessment and testing should contain sensitive bioassays that cover a wide range of cellular toxicity pathways (Figure 1). For induction of xenobiotic metabolism pathways, we recommend AhR and PXR. For specific modes of action, the receptor-mediated hormonal effects related to the estrogenic, glucocorticoid and antiandrogenic pathways appear to be most relevant as most are responsive to water samples. The oxidative stress response clearly stands out as a highly responsive defense mechanism. Cell viability (“cytotoxicity”) assays should be further developed with a focus on those representative of systemic responses.

■ ASSOCIATED CONTENT

§ Supporting Information

Additional information on the water samples, the bioassay methods, the data evaluation, initial screening, bioassay results (including a large table with all detailed results), and a text on monitoring the treatment efficacy at WRPs is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*(B.I.E.) Phone +61 7 3274 9180; fax: +61 7 3274 9003; e-mail: b.escher@uq.edu.au.

Present Address

University of California Davis, Department of Environmental Toxicology, One Shields Avenue, Davis, California 95616–8588, United States.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare the following competing financial interest(s): Attagene and BDS are companies that market the bioassays applied by them in the present study.

■ ACKNOWLEDGMENTS

We thank Unitywater and Seqwater for access to their treatment plants. We acknowledge J.P. Giesy of University of Saskatchewan, Canada, for sharing the H4IIE-luc cells with RECETOX and M. Denison from University of California, Davis, U.S. for sharing the AhR-CAFLUX used at UQ and RECETOX. AREc32 cells were kindly provided by C.R. Wolf from University of Dundee, UK, to U.Q. We thank Michael Bartkow for helpful discussions, Peta Neale, Julien Reungoat, and Jatinder Sidhu for sampling assistance and Daniela Baumberger, Mriga Dutt, Eva Glenn, Ling Jin, and Shane McCarty for experimental assistance. This work was supported, mainly, by the WaterReuse Research Foundation (WRF 10-07), and, in part, by the California Water Resources Control Board (Agreement No. 10-096-250) and the European Union, project Demeau, grant agreement number 308339.

■ ABBREVIATIONS

AO	advanced oxidation
AR	androgen receptor
ASR	adaptive stress response
ATG	Attagene
AWQC	Australian Water Quality Centre
BDS	BioDetection Systems
BEQ	bioanalytical equivalent concentration
CAPIM	Centre for Aquatic Pollution Identification and Management
CAR	constitutive androstane receptor
CSIRO	Commonwealth Scientific and Industrial Research Organisation
CT	cytotoxicity
CYP	cytochrome P450 monooxygenase
DART	embryo toxicity test with the zebrafish <i>Danio rerio</i>
DW	drinking water
EC	effect concentration
EEQ	estradiol equivalent concentration
Eff	effluent

ER	estrogen receptor
GU	Griffith University
GR	glucocorticoid receptor
HK	Hong Kong Baptist University
HTS	high-throughput screening
IR	induction ratio
IRCM	Cancer Research Institute of Montpellier
ISO	International Organization for Standardization
IWW	Institute for Water Research in North-Rhine Westfalia, Germany
MF	microfiltration
MOA	mode of action
MTT	(3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NJU	Nanjing University
NIH	National Institutes of Health
NRU	neutral red uptake
OECD	Organisation for Economic Co-operation and Development
PPAR	peroxisome proliferator-activated receptor
PR	progesterone receptor
PXR	pregnane-X-receptor
RAR	retinoic acid receptor
RCEES	Research Center for Eco-Environmental Sciences
RECETOX	Research Centre for Toxic Compounds in the Environment
REF	relative enrichment factor
RFU	relative fluorescence units
RLU	relative light units
RO	reverse osmosis
ROS	reactive oxygen species
RT-PCR	real-time polymerase chain reaction
RW	river water
SCCWRP	Southern California Coastal Water Research Project
SPE	solid phase extraction
SW	stormwater
SWISS	Swiss Centre for Applied Ecotoxicology
TEQ	toxic equivalent concentration
UA	University of Arizona
UCR	University of California Riverside
UF	University of Florida
UFZ	Helmholtz Centre for Environmental Research
UQ	The University of Queensland
USF	University of South Florida
WRP	water reclamation plant
WWTP	wastewater treatment plant
XM	xenobiotic metabolism

■ REFERENCES

- (1) NRC. *Toxicity Testing in the 21st Century: A Vision and a Strategy. Report in Brief*; National Research Council (NRC): Washington, D.C., 2007.
- (2) Kaylock, R.; Chandler, K.; Houck, K.; Hunter, S.; Judson, R.; Kleinstreuer, N.; Knudsen, T.; Martin, M.; Padilla, S.; Reif, D.; Richard, A.; Rotroff, D.; Sipes, N.; Dix, D. Update on EPA's ToxCast Program: Providing high throughput decision support tools for chemical risk management. *Chem. Res. Toxicol.* **2012**, *25* (7), 1287–1302.
- (3) van der Burg, B.; Kroese, E. D.; Piersma, A. H. Towards a pragmatic alternative testing strategy for the detection of reproductive toxicants. *Reprod. Toxicol.* **2011**, *31* (4), 558–561.
- (4) Escher, B.; Leusch, F. *Bioanalytical Tools in Water Quality Assessment*; IWA Publishing: London, UK, 2012.

- (5) Collins, F.; Gray, G. N.; Bucher, J. R. Transforming environmental health protection. *Science* **2008**, *319*, 906–907.
- (6) Gohlke, J. M.; Portier, C. J. The forest for the trees: A systems approach to human health research. *Environ. Health Perspect.* **2007**, *115*, 1261–1263.
- (7) Judson, R.; Kavlock, R.; Martin, M.; Reif, D.; Houck, K.; Knudsen, T.; Richard, A.; Tice, R. R.; Whelan, M.; Xia, M.; Huang, R.; Austin, C.; Daston, G.; Hartung, T.; Fowle, J. R., III; Wooge, W.; Tong, W.; Dix, D. Perspectives on validation of high-throughput assays supporting 21st century toxicity testing. *ALTEX* **2013**, *30* (1), 51–66.
- (8) Tice, R. R.; Austin, C. P.; Kavlock, R.; Bucher, J. R. Improving the human hazard characterization of chemicals: A Tox21 update. *Environ. Health Perspect.* **2013**, *121*, 756–765.
- (9) Martin, M. T.; Dix, D. J.; Judson, R. S.; Kavlock, R. J.; Reif, D. M.; Richard, A. M.; Rotroff, D. M.; Romanov, S.; Medvedev, A.; Poltoratskaya, N.; Gambarian, M.; Moeser, M.; Makarov, S. S.; Houck, K. A. Impact of environmental chemicals on key transcription regulators and correlation to toxicity endpoints within EPA's ToxCast program. *Chem. Res. Toxicol.* **2010**, *23* (3), 578–590.
- (10) Macova, M.; Toze, S.; Hodgers, L.; Mueller, J. F.; Bartkow, M. E.; Escher, B. I. Bioanalytical tools for the evaluation of organic micropollutants during sewage treatment, water recycling and drinking water generation. *Water Res.* **2011**, *45* (14), 4238–4247.
- (11) Escher, B. I.; Lawrence, M.; Macova, M.; Mueller, J. F.; Poussade, Y.; Robillot, C.; Roux, A.; Gernjak, W. Evaluation of contaminant removal of reverse osmosis and advanced oxidation in full-scale operation by combining passive sampling with chemical analysis and bioanalytical tools. *Environ. Sci. Technol.* **2011**, *45* (12), 5387–5394.
- (12) Reungoat, J.; Escher, B. I.; Macova, M.; Argaud, F. X.; Gernjak, W.; Keller, J. Ozonation and biological activated carbon filtration of wastewater treatment plant effluents. *Water Res.* **2012**, *46* (3), 863–872.
- (13) Reungoat, J.; Escher, B. I.; Macova, M.; Keller, J. Biofiltration of wastewater treatment plant effluent: Effective removal of pharmaceuticals and personal care products and reduction of toxicity. *Water Res.* **2011**, *45* (9), 2751–2762.
- (14) Reungoat, J.; Macova, M.; Escher, B. I.; Carswell, S.; Mueller, J. F.; Keller, J. Removal of micropollutants and reduction of biological activity in a full scale reclamation plant using ozonation and activated carbon filtration. *Water Res.* **2010**, *44* (2), 625–637.
- (15) Neale, P. A.; Antony, A.; Bartkow, M.; Farre, M.; Heitz, A.; Kristiana, I.; Tang, J. Y. M.; Escher, B. I. Bioanalytical assessment of the formation of disinfection by-products in a drinking water treatment plant. *Environ. Sci. Technol.* **2012**, *46* (18), 10317–10325.
- (16) Tang, J. Y. M.; Glenn, E.; Aryal, R.; Gernjak, W.; Escher, B. I. Toxicity characterization of urban stormwater with bioanalytical tools. *Water Res.* **2013**, *47* (15), 5594–5606.
- (17) Leusch, F. D. L.; Khan, S. J.; Laingam, S.; Prochazka, E.; Trinh, T.; Froscio, S.; Chapman, H.; Humpage, A. Assessment of the application of bioanalytical tools as surrogate measure of chemical contaminants in recycled water. *Water Res.* **2014**, *49* (1), 300–315.
- (18) Tang, J. Y. M.; McCarty, S.; Glenn, E.; Neale, P. A.; Warne, M. S.; Escher, B. I. Mixture effects of organic micropollutants present in water: Towards the development of effect-based water quality trigger values for baseline toxicity. *Water Res.* **2013**, *47* (10), 3300–3314.
- (19) Strähle, U.; Scholz, S.; Geisler, R.; Greiner, P.; Hollert, H.; Rastegar, S.; Schumacher, A.; Selderslaghs, I.; Weiss, C.; Witters, H.; Brauneck, T. Zebrafish embryos as an alternative to animal experiments—A commentary on the definition of the onset of protected life stages in animal welfare regulations. *Reprod. Toxicol.* **2012**, *33* (2), 128–132.
- (20) Seo, J. B.; Moon, H. M.; Kim, W. S.; Lee, Y. S.; Jeong, H. W.; Yoo, E. J.; Ham, J. Y.; Kang, H. J.; Park, M. G.; Steffensen, K. R.; Stulnig, T. M.; Gustafsson, J. A.; Dai Park, S.; Kim, J. B. Activated liver X receptors stimulate adipocyte differentiation through induction of peroxisome proliferator-activated receptor gamma expression. *Mol. Cell. Biol.* **2004**, *24* (8), 3430–3444.
- (21) Janesick, A.; Blumberg, B. Minireview: PPAR gamma as the target of obesogens. *J. Steroid Biochem. Mol. Biol.* **2011**, *127* (1–2), 4–8.
- (22) Neale, P. A.; Escher, B. I. Does dissolved organic carbon cause artefacts in cell-based bioassays? *Chemosphere* **2013**, submitted 25 July 2013.
- (23) Denison, M. S.; Nagy, S. R. Activation of the aryl hydrocarbon receptor by structurally diverse exogenous and endogenous chemicals. *Annu. Rev. Pharmacol. Toxicol.* **2003**, *43*, 309–334.
- (24) Odermatt, A.; Gumy, C.; Atanasov, A. G.; Dzyakanchuk, A. A. Disruption of glucocorticoid action by environmental chemicals: Potential mechanisms and relevance. *J. Steroid Biochem. Mol. Biol.* **2006**, *102* (1–5), 222–231.
- (25) Omiecinski, C. J.; Heuvel, J. P. V.; Perdew, G. H.; Peters, J. M. Xenobiotic metabolism, disposition, and regulation by receptors: From biochemical phenomenon to predictors of major toxicities. *Toxicol. Sci.* **2011**, *120*, S49–S75.
- (26) Scarsi, M.; Podvinec, M.; Roth, A.; Hug, H.; Kersten, S.; Albrecht, H.; Schwede, T.; Meyer, U. A.; Rucker, C. Sulfonylureas and glinides exhibit peroxisome proliferator-activated receptor gamma activity: A combined virtual screening and biological assay approach. *Mol. Pharmacol.* **2007**, *71* (2), 398–406.
- (27) Huang, R. L.; Xia, M. H.; Cho, M. H.; Sakamuru, S.; Shinn, P.; Houck, K. A.; Dix, D. J.; Judson, R. S.; Witt, K. L.; Kavlock, R. J.; Tice, R. R.; Austin, C. P. Chemical genomics profiling of environmental chemical modulation of human nuclear receptors. *Environ. Health Perspect.* **2011**, *119* (8), 1142–1148.
- (28) le Maire, A.; Grimaldi, M.; Roecklin, D.; Dagnino, S.; Vivat-Hannah, V.; Balaguer, P.; Bourguet, W. Activation of RXR-PPAR heterodimers by organotin environmental endocrine disruptors. *EMBO Rep.* **2009**, *10* (4), 367–373.
- (29) Riu, A.; Grimaldi, M.; le Maire, A.; Bey, G.; Phillips, K.; Boulahtouf, A.; Perdu, E.; Zalko, D.; Bourguet, W.; Balaguer, P. Peroxisome proliferator-activated receptor gamma is a target for halogenated analogs of bisphenol A. *Environ. Health Perspect.* **2011**, *119* (9), 1227–1232.
- (30) Neale, P. A.; Escher, B. I. Co-extracted dissolved organic carbon has a suppressive effect on the acetylcholinesterase inhibition assay. *Environ. Toxicol. Chem.* **2013**, *32* (7), 1526–1534.
- (31) Muller, R.; Schreiber, U.; Escher, B. I.; Quayle, P.; Nash, S. M. B.; Mueller, J. F. Rapid exposure assessment of PSII herbicides in surface water using a novel chlorophyll a fluorescence imaging assay. *Sci. Total Environ.* **2008**, *401* (1–3), 51–59.
- (32) Stalter, D.; Magdeburg, A.; Wagner, M.; Oehlmann, J. Ozonation and activated carbon treatment of sewage effluents: Removal of endocrine activity and cytotoxicity. *Water Res.* **2011**, *45* (3), 1015–1024.
- (33) Van der Linden, S. C.; Heringa, M. B.; Man, H. Y.; Sonneveld, E.; Puijker, L. M.; Brouwer, A.; Van der Burg, B. Detection of multiple hormonal activities in wastewater effluents and surface water, using a panel of steroid receptor CALUX bioassays. *Environ. Sci. Technol.* **2008**, *42* (15), 5814–5820.
- (34) Schriks, M.; van der Linden, S. C.; Stoks, P. G. M.; van der Burg, B.; Puijker, L.; de Voogt, P.; Heringa, M. B. Occurrence of glucocorticogenic activity in various surface waters in The Netherlands. *Chemosphere* **2013**, *93* (2), 450–454.
- (35) Bruce, G. M.; Corey, L. M.; Mandel, J. H.; Pleus, R. C. Urinary nitrate, thiocyanate, and perchlorate and serum thyroid endpoints based on NHANES 2001 to 2002. *J. Occup. Environ. Med.* **2013**, *55* (1), 64–70.
- (36) Novák, J.; Benísek, M.; Hilscherová, K. Disruption of retinoid transport, metabolism and signaling by environmental pollutants. *Environ. Int.* **2008**, *34* (6), 898–913.
- (37) Lefebvre, P.; Benomar, Y.; Staels, B. Retinoid X receptors: Common heterodimerization partners with distinct functions. *Trends Endocrinol. Metab.* **2010**, *21* (11), 676–683.
- (38) Tang, J. Y. M.; Glenn, E.; Thoen, H.; Escher, B. I. *In vitro* bioassay for reactive toxicity towards proteins implemented for water quality monitoring. *J. Environ. Monit.* **2012**, *14*, 1073–1081.

- (39) Escher, B. I.; Dutt, M.; Maylin, E.; Tang, J. Y. M.; Toze, S.; Wolf, C. R.; Lang, M. Water quality assessment using the AREc32 reporter gene assay indicative of the oxidative stress response pathway. *J. Environ. Monit.* **2012**, *14*, 2877–2885.
- (40) Clothier, R.; Gomez-Lechon, M. J.; Kinsner-Ovaskainen, A.; Kopp-Schneider, A.; O'Connor, J. E.; Prieto, P.; Stanzel, S. Comparative analysis of eight cytotoxicity assays evaluated within the ACuteTox Project. *Toxicol. Vitro* **2013**, *27* (4), 1347–1356.
- (41) Ba, F.; Pang, P. K. T.; Benishin, C. G. The establishment of a reliable cytotoxic system with SK-N-SH neuroblastoma cell culture. *J. Neurosci. Methods* **2003**, *123* (1), 11–22.
- (42) Humpage, A. R.; Ledreux, A.; Fanok, S.; Bernard, C.; Briand, J. F.; Eaglesham, G.; Papageorgiou, J.; Nicholson, B.; Steffensen, D. Application of the neuroblastoma assay for paralytic shellfish poisons to neurotoxic freshwater cyanobacteria: Interlaboratory calibration and comparison with other methods of analysis. *Environ. Toxicol. Chem.* **2007**, *26* (7), 1512–1519.
- (43) Richardson, S. D.; Plewa, M. J.; Wagner, E. D.; Schoeny, R.; DeMarini, D. M. Occurrence, genotoxicity, and carcinogenicity of regulated and emerging disinfection by-products in drinking water: A review and roadmap for research. *Mutat. Res.* **2007**, *636* (1–3), 178–242.
- (44) Escher, B. I.; van Daele, C.; Dutt, M.; Tang, J. Y. M.; Altenburger, R. Most oxidative stress response in water samples comes from unknown chemicals: The need for effect-based water quality trigger values. *Environ. Sci. Technol.* **2013**, *47* (13), 7002–7011.
- (45) Kokel, D.; Bryan, J.; Laggner, C.; White, R.; Cheung, C. Y. J.; Mateus, R.; Healey, D.; Kim, S.; Werdich, A. A.; Haggarty, S. J.; MacRae, C. A.; Shoichet, B.; Peterson, R. T. Rapid behavior-based identification of neuroactive small molecules in the zebrafish. *Nat. Chem. Biol.* **2010**, *6* (3), 231–237.
- (46) Romanov, S.; Medvedev, A.; Gambarian, M.; Poltoratskaya, N.; Moeser, M.; Medvedeva, L.; Diatchenko, L.; Makarov, S. Homogeneous reporter system enables quantitative functional assessment of multiple transcription factors. *Nat. Methods* **2008**, *5* (3), 253–260.
- (47) Lemaire, G.; Mnif, W.; Pascussi, J. M.; Pillon, A.; Rabenoelina, F.; Fenet, H.; Gomez, E.; Casellas, C.; Nicolas, J. C.; Cavailles, V.; Duchesne, M. J.; Balaguer, P. Identification of new human pregnane X receptor ligands among pesticides using a stable reporter cell system. *Toxicol. Sci.* **2006**, *91* (2), 501–509.
- (48) Creusot, N.; Kinani, S.; Balaguer, P.; Tapie, N.; LeMenach, K.; Maillot-Marechal, E.; Porcher, J. M.; Budzinski, H.; Ait-Aissa, S. Evaluation of an hPXR reporter gene assay for the detection of aquatic emerging pollutants: Screening of chemicals and application to water samples. *Anal. Bioanal. Chem.* **2010**, *396* (2), 569–583.
- (49) Shiraishi, F.; Shirashi, H.; Nishikawa, J.; Nishihara, T.; Morita, M. Development of a simple operational estrogenicity assay system using the yeast two-hybrid system. *J. Environ. Chem.* **2000**, *10* (1), 57–64.
- (50) Seimandi, M.; Lemaire, G.; Pillon, A.; Perrin, A.; Carlavan, I.; Voegel, J. J.; Vignon, F.; Nicolas, J. C.; Balaguer, P. Differential responses of PPAR alpha, PPAR delta, and PPAR gamma reporter cell lines to selective PPAR synthetic ligands. *Anal. Biochem.* **2005**, *344* (1), 8–15.
- (51) Gijsbers, L.; van Eekelen, H. D. L. M.; de Haan, L. H. J.; Swier, J. M.; Heijink, N. L.; Kloet, S. K.; Man, H. Y.; Bovy, A. G.; Keijer, J.; Aarts, J. M. M. J. G.; van der Burg, B.; Rietjens, I. M. C. M. Induction of peroxisome proliferator-activated receptor gamma (PPAR gamma)-mediated gene expression by tomato (*Solanum lycopersicum* L.) extracts. *J. Agric. Food Chem.* **2013**, *61* (14), 3419–3427.
- (52) Piersma, A. H.; Bosgra, S.; van Duursen, M. B. M.; Hermsen, S. A. B.; Jonker, L. R. A.; Kroese, E. D.; van der Linden, S. C.; Man, H.; Roelofs, M. J. E.; Schulpen, S. H. W.; Schwarz, M.; Uibel, F.; van Vuigt-Lussenburg, B. M. A.; Westerhout, J.; Wolterbeek, A. P. M.; van der Burg, B. Evaluation of an alternative in vitro test battery for detecting reproductive toxicants. *Reprod. Toxicol.* **2013**, *38*, 53–64.
- (53) Gijsbers, L.; Man, H. Y.; Kloet, S. K.; de Haan, L. H.; Keijer, J.; Rietjens, I. M.; van der Burg, B.; Aarts, J. M. Stable reporter cell lines for peroxisome proliferator-activated receptor gamma (PPARgamma)-mediated modulation of gene expression. *Anal. Biochem.* **2011**, *414* (1), 77–83.
- (54) Law, A. Y. S.; Wei, X.; Zhang, X. W.; Mak, N. K.; Cheung, K. C.; Wong, M. H.; Giesy, J. P.; Wong, C. K. C. Biological analysis of endocrine-disrupting chemicals in animal meats from the Pearl River Delta, China. *J. Exposure Sci. Environ. Epidemiol.* **2012**, *22* (1), 93–100.
- (55) EPA, U.S. *Prioritization of the Endocrine Disruptor Screening Program Universe of Chemicals for an Estrogen Receptor Adverse Outcome Pathway Using Computational Toxicology Tools*; U.S. Environmental Protection Agency: Washington, DC, 2012.
- (56) Miller, C. A. A human aryl hydrocarbon receptor signaling pathway constructed in yeast displays additive responses to ligand mixtures. *Toxicol. Appl. Pharmacol.* **1999**, *160* (3), 297–303.
- (57) Kamata, R.; Shiraishi, F.; Nakajima, D.; Takigami, H.; Shiraishi, H. Mono-hydroxylated polychlorinated biphenyls are potent aryl hydrocarbon receptor ligands in recombinant yeast cells. *Toxicol. Vitro* **2009**, *23* (4), 736–743.
- (58) Nagy, S. R.; Sanborn, J. R.; Hammock, B. D.; Denison, M. S. Development of a green fluorescent protein-based cell bioassay for the rapid and inexpensive detection and characterization of Ah receptor agonists. *Toxicol. Sci.* **2002**, *65* (2), 200–210.
- (59) Jarosova, B.; Blaha, L.; Vrana, B.; Randak, T.; Grabic, R.; Giesy, J. P.; Hilscherova, K. Changes in concentrations of hydrophilic organic contaminants and of endocrine-disrupting potential downstream of small communities located adjacent to headwaters. *Environ. Int.* **2012**, *45*, 22–31.
- (60) Jin, L.; van Mourik, L.; Gaus, C.; Escher, B. Applicability of passive sampling to bioanalytical screening of bioaccumulative chemicals in marine wildlife. *Environ. Sci. Technol.* **2013**, *47* (14), 7982–7988.
- (61) Sanderson, J. T.; Aarts, J.; Brouwer, A.; Froese, K. L.; Denison, M. S.; Giesy, J. P. Comparison of Ah receptor-mediated luciferase and ethoxyresorufin-O-deethylase induction in H4IIIE cells: Implications for their use as bioanalytical tools for the detection of polyhalogenated aromatic hydrocarbons. *Toxicol. Appl. Pharmacol.* **1996**, *137* (2), 316–325.
- (62) Weil, M.; Scholz, S.; Zimmer, M.; Sacher, F.; Duis, K. Gene expression analysis in zebrafish embryos: A potential approach to predict effect concentrations in the fish early life stage test. *Environ. Toxicol. Chem.* **2009**, *28* (9), 1970–1978.
- (63) Klüver, N.; Yang, L. X.; Busch, W.; Scheffler, K.; Renner, P.; Strähle, U.; Scholz, S. Transcriptional response of zebrafish embryos exposed to neurotoxic compounds reveals a muscle activity dependent hspb11 expression. *PLoS One* **2011**, *6*, e212.
- (64) Escher, B. I.; Bramaz, N.; Mueller, J. F.; Quayle, P.; Rutishauser, S.; Vermeirssen, E. L. M. Toxic equivalent concentrations (TEQs) for baseline toxicity and specific modes of action as a tool to improve interpretation of ecotoxicity testing of environmental samples. *J. Environ. Monitor* **2008**, *10* (5), 612–621.
- (65) Ellman, G. L.; Courtney, K. D.; Andres, V., Jr.; Featherstone, R. M. A new and rapid colorimetric determination of acetylcholinesterase activity. *Biochem. Pharmacol.* **1961**, *7* (2), 88–90 IN1, 91–95.
- (66) Sonneveld, E.; Jansen, H. J.; Ritsema, J. A. C.; Brouwer, A.; van der Burg, B. Development of androgen- and estrogen-responsive bioassays, members of a panel of human cell line-based highly selective steroid-responsive bioassays. *Toxicol. Sci.* **2005**, *83* (1), 136–148.
- (67) Soto, A. M.; Sonnenschein, C.; Chung, K. L.; Fernandez, M. F.; Olea, N.; Serrano, F. O. The E-SCREEN assay as a tool to identify estrogens: An update on estrogenic environmental pollutants. *Environ. Health Perspect.* **1995**, *103* (Suppl. 7), 113–122.
- (68) Routledge, E. J.; Sumpter, J. P. Estrogenic activity of surfactants and some of their degradation products assessed using a recombinant yeast screen. *Environ. Toxicol. Chem.* **1996**, *15* (3), 241–248.
- (69) Escher, B. I.; Bramaz, N.; Quayle, P.; Rutishauser, S.; Vermeirssen, E. L. M. Monitoring of the ecotoxicological hazard potential by polar organic micropollutants in sewage treatment plants and surface waters using a mode-of-action based test battery. *J. Environ. Monit.* **2008**, *10* (5), 622–631.

- (70) Allinson, M.; Shiraishi, F.; Allinson, G. A comparison of recombinant receptor-reporter gene bioassays and a total estrogen enzyme linked immunosorbent assay for the rapid screening of estrogenic activity in natural and waste waters. *Bull. Environ. Contam. Toxicol.* **2011**, *86* (5), 461–464.
- (71) Balaguer, P.; Francois, F.; Comunale, F.; Fenet, H.; Boussioux, A. M.; Pons, M.; Nicolas, J. C.; Casellas, C. Reporter cell lines to study the estrogenic effects of xenoestrogens. *Sci. Total Environ.* **1999**, *233* (1–3), 47–56.
- (72) Maggioni, S.; Balaguer, P.; Chiozzotto, C.; Benfenati, E. Screening of endocrine-disrupting phenols, herbicides, steroid estrogens, and estrogenicity in drinking water from the waterworks of 35 Italian cities and from PET-bottled mineral water. *Environ. Sci. Pollut. Res.* **2013**, *20* (3), 1649–1660.
- (73) Lemaire, G.; Mnif, W.; Mauvais, P.; Balaguer, P.; Rahmani, R. Activation of alpha- and beta-estrogen receptors by persistent pesticides in reporter cell lines. *Life Sci.* **2006**, *79* (12), 1160–1169.
- (74) OECD Test guideline 455. *Stably Transfected Human Estrogen Receptor-A Transcriptional Activation Assay for Detection of Estrogenic Agonist-Activity of Chemicals*, TG 305; Environmental Directorate; Organisation for Economic Co-operation and Development: Paris, France, 1984, 2009.
- (75) Zhang, X. W.; Chang, H.; Wiseman, S.; He, Y. H.; Higley, E.; Jones, P.; Wong, C. K. C.; Al-Khedhairy, A.; Giesy, J. P.; Hecker, M. Bisphenol A disrupts steroidogenesis in human H295R cells. *Toxicol. Sci.* **2011**, *121* (2), 320–327.
- (76) He, Y. H.; Wiseman, S. B.; Hecker, M.; Zhang, X. W.; Wang, N.; Perez, L. A.; Jones, P. D.; El-Din, M. G.; Martin, J. W.; Giesy, J. P. Effect of ozonation on the estrogenicity and androgenicity of oil sands process-affected water. *Environ. Sci. Technol.* **2011**, *45* (15), 6268–6274.
- (77) Brion, F.; Le Page, Y.; Piccini, B.; Cardoso, O.; Tong, S. K.; Chung, B. C.; Kah, O. Screening estrogenic activities of chemicals or mixtures *in vivo* using transgenic (cyp19a1b-GFP) zebrafish embryos. *PLoS One* **2012**, *7*, (S).
- (78) van der Burg, B.; Winter, R.; Man, H. Y.; Vangenechten, C.; Berckmans, P.; Weimer, M.; Witters, H.; van der Linden, S. Optimization and prevalidation of the *in vitro* AR CALUX method to test androgenic and antiandrogenic activity of compounds. *Reprod. Toxicol.* **2010**, *30* (1), 18–24.
- (79) Molina-Molina, J. M.; Hillenweck, A.; Jouanin, I.; Zalko, D.; Cravedi, J. P.; Fernandez, M. F.; Pillon, A.; Nicolas, J. C.; Olea, N.; Balaguer, P. Steroid receptor profiling of vinclozolin and its primary metabolites. *Toxicol. Appl. Pharmacol.* **2006**, *216* (1), 44–54.
- (80) Delfosse, V.; Grimaldi, M.; Pons, J. L.; Boulahouf, A.; le Maire, A.; Cavailles, V.; Labesse, G.; Bourguet, W.; Balaguer, P. Structural and mechanistic insights into bisphenols action provide guidelines for risk assessment and discovery of bisphenol A substitutes. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109* (37), 14930–14935.
- (81) Sohoni, P.; Sumpter, J. Several environmental oestrogens are also anti-androgens. *J. Endocrinol.* **1998**, *158* (3), 327–339.
- (82) Thomas, K. V.; Hurst, M. R.; Matthiessen, P.; McHugh, M.; Smith, A.; Waldock, M. J. An assessment of *in vitro* androgenic activity and the identification of environmental androgens in United Kingdom estuaries. *Environ. Toxicol. Chem.* **2002**, *21* (7), 1456–1461.
- (83) Wilson, V. S.; Bobeine, K.; Lambright, C. R.; Gray, L. E. A novel cell line, MDA-kb2, that stably expresses an androgen- and glucocorticoid-responsive reporter for the detection of hormone receptor agonists and antagonists. *Toxicol. Sci.* **2002**, *66* (1), 69–81.
- (84) Pavlikova, N.; Blahova, L.; Klan, P.; Bathula, S. R.; Sklenar, V.; Giesy, J. P.; Blaha, L. Enantioselective effects of alpha-hexachlorocyclohexane (HCH) isomers on androgen receptor activity *in vitro*. *Chemosphere* **2012**, *86* (1), 65–69.
- (85) Gutleb, A. C.; Meerts, I. A. T. M.; Bergsma, J. H.; Schriks, M.; Murk, A. J. T-Screen as a tool to identify thyroid hormone receptor active compounds. *Environ. Toxicol. Pharmacol.* **2005**, *19* (2), 231–238.
- (86) Fini, J. B.; Riu, A.; Debrauwer, L.; Hillenweck, A.; Le Mevel, S.; Chevolleau, S.; Boulahouf, A.; Palmier, K.; Balaguer, P.; Cravedi, J. P.; Demeneix, B. A.; Zalko, D. Parallel biotransformation of tetrabromobisphenol a in *Xenopus laevis* and mammals: *Xenopus* as a model for endocrine perturbation studies. *Toxicol. Sci.* **2012**, *125* (2), 359–367.
- (87) Novak, J.; Benisek, M.; Pachernik, J.; Janosek, J.; Sidlova, T.; Kiviranta, H.; Verta, M.; Giesy, J. P.; Blaha, L.; Hilscherova, K. Interference of contaminated sediment extracts and environmental pollutants with retinoid signaling. *Environ. Toxicol. Chem.* **2007**, *26* (8), 1591–1599.
- (88) Kamata, R.; Shiraishi, F.; Nishikawa, J.; Yonemoto, J.; Shiraishi, H. Screening and detection of the *in vitro* agonistic activity of xenobiotics on the retinoic acid receptor. *Toxicol. In Vitro* **2008**, *22* (4), 1050–1061.
- (89) Oda, Y.; Nakamura, S.-I.; Oki, I.; Kato, T.; Shinagawa, H. Evaluation of the new system (umu-test) for the detection of environmental mutagens and carcinogens. *Mutat. Res.* **1985**, *147* (5), 219–229.
- (90) ISO13828, *Water Quality—Determination of the Genotoxicity of Water and Waste Water Using the Umu-Test*; International Organization for Standardization (ISO): Geneva, Switzerland, 1999.
- (91) Oda, Y.; Yamazaki, H.; Thier, R.; Ketterer, B.; Guengerich, F. P.; Shimada, T. A new *Salmonella typhimurium* NM5004 strain expressing rat glutathione S-transferase S-5: Use in detection of genotoxicity of dihaloalkanes using an SOS/umu test system. *Carcinogenesis* **1996**, *17* (2), 297–302.
- (92) Quillardet, P.; Huisman, O.; Dari, R.; Hofnung, M. SOS Chromotest, a direct assay of induction of an SOS function in *Escherichia coli* K-12 to measure genotoxicity. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79* (19), 5971–5975.
- (93) Skarek, M.; Janosek, J.; Cupr, P.; Kohoutek, J.; Novotna-Rychetska, A.; Holoubek, I. Evaluation of genotoxic and non-genotoxic effects of organic air pollution using *in vitro* bioassays. *Environ. Int.* **2007**, *33* (7), 859–866.
- (94) Ames, B. N.; McCann, J.; Yamasaki, E. Methods for detecting carcinogens and mutagens with the salmonella/mammalian-microsome mutagenicity test. *Mutat. Res.* **1975**, *31* (6), 347–363.
- (95) Umbuzeiro, G. D.; Rech, C. M.; Correia, S.; Bergamasco, A. M.; Cardenette, G. H. L.; Flückiger-Isler, S.; Kamber, M. Comparison of the *Salmonella*/microsome microsuspension assay with the new microplate fluctuation protocol for testing the mutagenicity of environmental samples. *Environ. Mol. Mutagen.* **2010**, *51* (1), 31–38.
- (96) Richard, J. *Toxicological in Vitro Evaluation of Organic Micropollutants before and after Oxidative Waste Water Treatment*; Universität Duisburg-Essen: Essen, Germany, 2012; <http://duepublico.uni-duisburg-essen.de/servlets/DocumentServlet?id=31124>.
- (97) Reifferscheid, G.; Maes, H. M.; Allner, B.; Badurova, J.; Belkin, S.; Bluhm, K.; Brauer, F.; Bressling, J.; Domeneghetti, S.; Elad, T.; Flückiger-Isler, S.; Grummt, H. J.; Gürtler, R.; Hecht, A.; Heringa, M. B.; Hollert, H.; Huber, S.; Kramer, M.; Magdeburg, A.; Ratte, H. T.; Sauerborn-Klobucar, R.; Sokolowski, A.; Soldan, P.; Smital, T.; Stalter, D.; Venier, P.; Ziemann, C.; Zipperle, J.; Buchinger, S. International round-robin study on the Ames fluctuation test. *Environ. Mol. Mutagen.* **2012**, *53* (3), 185–197.
- (98) Fenech, M.; Chang, W. P.; Kirsch-Volders, M.; Holland, N.; Bonassi, S.; Zeiger, E. HUMN project: Detailed description of the scoring criteria for the cytokinesis-block micronucleus assay using isolated human lymphocyte cultures. *Mutat. Res.* **2003**, *534* (1–2), 65–75.
- (99) Laingam, S.; Froscio, S. M.; Humpage, A. R. Flow-cytometric analysis of *in vitro* micronucleus formation: Comparative studies with WIL2-NS human lymphoblastoid and L5178Y mouse lymphoma cell lines. *Mutat. Res.* **2008**, *656* (1–2), 19–26.
- (100) Rosenkranz, A. R.; Schmalldienst, S.; Stuhlmeier, K. M.; Chen, W. J.; Knapp, W.; Zlabinger, G. J. A microplate assay for the detection of oxidative products using 2',7'-dichlorofluorescin-diacetate. *J. Immunol. Meth.* **1992**, *156* (1), 39–45.
- (101) Bain, P. A.; Schuller, K. A. Use of fluorogenic probes to differentiate between hydrophilic and lipophilic antioxidant activity in a fish cell line. *J. Agric. Food Chem.* **2012**, *60* (3), 699–705.

- (102) Harder, A.; Escher, B. I.; Landini, P.; Tobler, N. B.; Schwarzenbach, R. P. Evaluation of bioanalytical tools for toxicity assessment and mode of toxic action classification of reactive chemicals. *Environ. Sci. Technol.* **2003**, *37* (21), 4962–4970.
- (103) Piersma, A. H. Alternative methods for developmental toxicity testing. *Basic Clin. Pharmacol. Toxicol.* **2006**, *98* (5), 427–431.
- (104) Wang, X. J.; Hayes, J. D.; Wolf, C. R. Generation of a stable antioxidant response element-driven reporter gene cell line and its use to show redox-dependent activation of Nrf2 by cancer chemotherapeutic agents. *Cancer Res.* **2006**, *66* (22), 10983–10994.
- (105) Villeneuve, N. F.; Du, Y.; Wang, X. J.; Sun, Z.; Zhang, D. D. High-throughput screening of chemopreventive compounds targeting Nrf2. In *3rd IEEE International Conference on Nano/Micro Engineered and Molecular Systems*; IEEE: New York, NY, 2008; Vols 1–3, pp 1138–1141.
- (106) Van der Linden, S.; von Bergh, A.; Van Vugt-Lussenburg, B.; Jonker, L.; Brouwer, A.; Teunis, M.; Krul, C.; van der Burg, B.; Development of a panel of high throughput reporter gene assays to detect genotoxicity and oxidative stress. *Mutat. Res.* **2013**, in press, DOI: 10.1016/j.mrgentox.2013.09.009.
- (107) Konsoula, R.; Barile, F. A. Correlation of *in vitro* cytotoxicity with paracellular permeability in Caco-2 cells. *Toxicol. Vitro* **2005**, *19* (5), 675–684.
- (108) Mosmann, T. Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays. *J. Immunol. Meth.* **1983**, *65* (1–2), 55–63.
- (109) Küster, E.; Altenburger, R. Comparison of cholin- and carboxylesterase enzyme inhibition and visible effects in the zebra fish embryo bioassay under short-term paraoxon-methyl exposure. *Biomarkers* **2006**, *11* (4), 341–354.
- (110) Gündel, U.; Kalkhof, S.; Zitzkat, D.; von Bergen, M.; Altenburger, R.; Küster, E. Concentration-response concept in ecotoxicoproteomics: Effects of different phenanthrene concentrations to the zebrafish (*Danio rerio*) embryo proteome. *Ecotox. Environ. Saf.* **2012**, *76*, 11–22.
- (111) Knöbel, M.; Busser, F. J. M.; Rico-Rico, A.; Kramer, N. I.; Hermens, J. L. M.; Hafner, C.; Tanneberger, K.; Schirmer, K.; Scholz, S. Predicting adult fish acute lethality with the zebrafish embryo: Relevance of test duration, endpoints, compound properties, and exposure concentration analysis. *Environ. Sci. Technol.* **2012**, *46* (17), 9690–9700.
- (112) Manger, R. L.; Leja, L. S.; Lee, S. Y.; Hungerford, J. M.; Hokama, Y.; Dickey, R. W.; Granade, H. R.; Lewis, R.; Yasumoto, T.; Wekell, M. M. Detection of paralytic shellfish poison by rapid cell bioassay: Antagonism of voltage-gated sodium channel active toxins *in vitro*. *J. AOAC Int.* **1995**, *78* (2), 521–527.
- (113) Baqui, A.; Meiller, T. F.; Chon, J. J.; Turng, B. F.; Falkler, W. A. Granulocyte-macrophage colony-stimulating factor amplification of interleukin-1 beta and tumor necrosis factor alpha production in THP-1 human monocytic cells stimulated with lipopolysaccharide of oral microorganisms. *Clin. Diagn. Lab. Immunol.* **1998**, *5* (3), 341–347.
- (114) OECD. *Test Guideline No. 201. Alga, Growth Inhibition Test*; Environmental Directorate, Organisation for Economic Co-operation and Development: Paris, France, 1984.
- (115) ISO11348-3. *Water Quality—Determination of the Inhibitory Effect of Water Samples on the Light Emission of Vibrio Fischeri (Luminescent Bacteria Test)*; International Organization for Standardization (ISO): Geneva, Switzerland, 1998.
- (116) GB/T 15441-1995 *Water Quality—determination of the Acute Toxicity—Luminescent Bacteria Test*; National Standard of China, 1995; http://kjs.mep.gov.cn/hjbhbz/bzwb/shjhb/sjcgfffbz/199508/t19950801_67352.htm.
- (117) Ankley, G. T.; Bennett, R. S.; Erickson, R. J.; Hoff, D. J.; Hornung, M. W.; Johnson, R. D.; Mount, D. R.; Nichols, J. W.; Russom, C. L.; Schmieder, P. K.; Serrano, J. A.; Tietge, J. E.; Villeneuve, D. L. Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environ. Toxicol. Chem.* **2010**, *29* (3), 730–741.

Supporting Information

Benchmarking organic micropollutants in wastewater, recycled water and drinking water with *in vitro* bioassays

Beate I. Escher, Mayumi Allinson, Rolf Altenburger, Peter A. Bain, Patrick Balaguer, Wibke Busch, Jordan Crago, Nancy D. Denslow, Elke Dopp, Klara Hilscherova, Andrew R. Humpage, Anu Kumar, Marina Grimaldi, B. Sumith Jayasinghe, Barbora Jarosova, Ai Jia, Sergei Makarov, Keith A. Maruya, Alex Medvedev, Alvine C. Mehinto, Jamie E. Mendez, Anita Poulsen, Erik Prochazka, Jessica Richard, Andrea Schifferli, Daniel Schlenk, Stefan Scholz, Fujio Shiraishi, Shane Snyder, Guanyong Su, Janet Y.M. Tang, Bart van der Burg, Sander C. van der Linden, Inge Werner, Sandy D. Westerheide, Chris K.C. Wong, Min Yang, Bonnie H.Y. Yeung, Xiaowei Zhang, and Frederic D.L. Leusch

Table of contents

Section S1. Additional information on collected water samples

Table S1. Description of samples and concentration of dissolved organic carbon (DOC, mgC L⁻¹)

Section S2. Additional information on the bioassay methods

Table S2. Modifications in the bioassay methods in comparison to the literature references.

Section S3. Additional information on data evaluation

Table S3. Comparison of 10% effect level with IR1.5 effect level.

Figure S1. Derivation of EC_{SR0.2}.

Section S4. How robust were bioassays performed in different laboratories?

Table S4. Comparison of bioassay results for the same bioassay performed in different laboratories.

Section S5. Initial screening of nuclear receptors and transcription factors

Figure S2. Limit of detection (LOD).

Figure S3. Screening of 25 nuclear receptors and 48 transcription factors with the FACTORIAL bioassay.

Section S6. Additional information on bioassay results

Figure S4. Presentation of bioassay results.

Figure S5. Results of bioassays indicative of induction of xenobiotic metabolism pathways.

Figure S6. Results of bioassays indicative of estrogenicity.

Figure S7. Results of bioassays indicative of glucocorticoid receptor (GR) activation.

Figure S8. Results of bioassays indicative of reactive modes of action.

Figure S9. Results of bioassays indicative of adaptive stress response pathways.

Table S5. Summary of all EC values.

Section S7. Monitoring treatment efficacy

Figure S10. Bioanalytical fingerprint of the water from the WRP process using reverse osmosis.

Figure S11. Bioanalytical fingerprint of the water treated with ozonation and biological activated carbon.

Figure S12. Percent treatment efficiency in the 13 bioassays that did not fall below limit of detection after treatment.

Figure S13. Bioanalytical fingerprint of the drinking water treatment.

Section S1. Additional information on collected water samples

The water samples were collected across a subset of sites in Southeast Queensland, Australia (Table S1), where previous water collections and bioanalytical characterization have taken place (Macova *et al.*, 2011).

Table S1. Description of samples and concentration of dissolved organic carbon (DOC).

Sample	Site description	DOC (mgC L ⁻¹)	Collected amount of water
Water Reclamation Plant 1			
Eff1	WWTP effluent from a municipal sewage treatment plant that uses activated sludge treatment, taken at the influent of the WRP	8.0 ± 0.1	14 L
MF	Water sample taken after microfiltration (MF) using filters disinfected by chloramination to avoid biofouling	7.7 ± 0.3	14 L
RO	Permeate after the reverse osmosis (RO) process	0.3 ± 0.1	28 L
AO	Final recycled water after RO and treatment with UV/H ₂ O ₂ (indirect potable reuse quality)	0.2 ± 0.2	28 L
Water Reclamation Plant 2			
Eff2	WWTP effluent from a municipal sewage treatment plant that uses activated sludge treatment, taken at the influent of the WRP	11.5 ± 0.2	14 L
O₃/BAC	Recycled water after ozonation and biologically activated carbon filtration (for industrial reuse and irrigation)	4.8 ± 0.4	28 L
RW	River water taken at the influent of a metropolitan drinking water treatment plant	5.6 ± 0.2	28 L
DW	Drinking water treated by coagulation, chlorination and chloramination	3.0 ± 0.4	28 L

SW	Collected after a rain event on the 25 th of January 2012 from a stormwater drain that receives runoff from a residential catchment.	4.2 ± 2.0 ^a	14 L
H₂O	Ultrapure water (milliQ water) run through the same SPE as all other water samples		28 L

^aAverage of eight measurements collected at the same site during 16/03/2011 to 17/04/2012, as SW collected 25/01/2013 yielded invalid DOC results. SPE = solid phase extraction, WRP = water reclamation plant, WWTP = wastewater treatment plant.

The water samples were collected in 1 L amber glass bottles and transported to the laboratory within 2 h where they were acidified to pH 3 with concentrated hydrochloric acid (HCl). Samples containing chlorine were quenched with sodium thiosulphate (1 g/L). Most samples underwent solid phase extraction (SPE) immediately. As the sample volume was high, some samples could not be processed immediately and these were cooled to 4°C and stored for less than a week before SPE was performed.

The SPE was performed according to Macova *et al.* (2011) with the sorbent material validated in (Leusch *et al.*, 2013). All samples were filtered with a 1.6 µm glass fiber filter (GF/A Whatman) before extraction. Fourteen one-liter batches of Eff1, MF, Eff2 and SW and 28 one-litre batches of RO, AO, O₃/BAC, RW, DW and Blank (defined in Table S1) were extracted by passing each through two 6 cc solid phase cartridges in series, first an Oasis® HLB (500 mg, Catalogue Number 186000115, Waters) followed by a Supelclean coconut charcoal cartridge (2 g, Catalogue Number 57144-U, Sigma-Aldrich). Both types of cartridges were individually preconditioned prior to extraction with 10 mL of 1:1 acetone:hexane mixture, followed by 10 mL methanol and 10 mL of 5 mM HCl in MilliQ water. This resulted in 28/14 pairs of cartridges per sample (28 pairs for the cleaner samples and controls, 14 pairs of the WWTP effluent samples, MF and SW).

All cartridges were sealed individually and kept at -20°C until elution. Before elution the cartridges were defrosted and dried completely under vacuum, then elution was carried out with 10 mL of methanol and 10 mL of acetone:hexane. The eluate of 8/4 pairs of cartridges per sample were combined and evaporated under purified nitrogen gas before being solvent exchanged to methanol at a final volume of 1 mL.

The SPE extracts were aliquoted and tested in four laboratories (ATG, GU, UA, UQ). The extracts were dried as described below to send to ATG and were sent as methanolic extracts to UA. After the initial positive results, the remaining 20/10 pairs of cartridges, which had been stored for 5 months, were eluted. The extracts were combined and aliquoted for the remaining 16 laboratories, and were evaporated under purified nitrogen gas before being solvent exchanged to dimethyl sulfoxide (DMSO) at a final volume of 2 µL for shipping. The 2 µL samples in Agilent high-recovery HPLC vials (Catalogue Number 5183-2030) were flushed with purified argon gas. The samples were shipped at room temperature with express mail to all laboratories, where they were reconstituted upon arrival (after 1 day (Australia) to 3-5 days (overseas)) with appropriate solvent and stored at -20 °C until bioanalysis.

For practicality while all water samples were collected and enriched on SPE cartridges together, the cartridges were eluted in two batches. The first batch was used in four laboratories (UQ, UA, GU and ATG). Only after the appropriateness of the 10 samples were assessed in this initial stage by comparison with historic data, the second batch was extracted, aliquoted and sent out to the remaining 16 laboratories. To assure that the storage of cartridges had not changed the samples, the Microtox assay (see below) was performed on both batches and there was good agreement (paired t-test, paring was effective with a P=0.0001 and r=0.9351 and a log-log linear regression with a r^2 of 0.9162).

Section S2. Additional information on the bioassay methods

In Table 1 of the main article, all bioassay methods are referenced. In some cases, small modifications were made to the protocols and these are listed in Table S2. If modifications were more extensive or the protocols were unpublished, these are detailed in the following paragraphs and referred to in Table S2.

Table S2. Modifications in the bioassay methods in comparison to the literature references. Only modified assays are included in this Table.

#	Laboratories	Bioassay	Method modification
5	CAPIM	CAR-yeast	Section S2-A.
9	BDS	CALUX-PPAR α	BDS: Assay performed at 1% DMSO and in 384-well format.
10	BDS, CSIRO	CALUX-PPAR γ	BDS: Assay performed at 1% DMSO and in 384-well format, CSIRO: Cells were lysed in 50 μ L of Triton lysis buffer. Luciferase assay substrate was prepared according to Brasier and Fortin (2001).
12	GU	PPAR γ - GeneBLAzer	Section S2-B.
13	GU	Anti-PPAR γ - GeneBLAzer	Section S2-B.
15	UQ, RECETOX	CAFLUX	RECETOX: seeded at 30000/well, 24h of exposure; cells washed with PBS, 100 μ l PBS added to each well, measured by fluorometer.
19	UFZ	DART cyp1a induction	Different exposure times (0-120 hpf) were used.
22	GU, CSIRO, BDS, IWW	ER-CALUX	BDS: Assay performed at 1% DMSO and in 384-well format, CSIRO: Cells were lysed in 50 μ L of Triton lysis buffer. Luciferase assay substrate was performed according to Brasier and Fortin (2001). IWW: (Richard, 2012).
30	RECETOX	hER α -HeLa-9903	Medium DMEM-F12 (Sigma Aldrich, USA), 10% dialyzed fetal calf serum treated with dextran coated charcoal. Each plate exposed to: medium, solvent control, 17 β -estradiol (1-500 pM) in triplicates, for 24h, 37° C.
34	UFZ	DART cyp19a1b	Different exposure time (0-120 hpf) was used.
35	UF, USF, UCR, SCCWRP	ER α - GeneBLAzer	96-Well format, Section S2-C.
37	GU, BDS,	AR-CALUX	BDS: Assay performed at 1% DMSO and in

#	Labora-tories	Bioassay	Method modification
	CSIRO		384 well format.
41	UF, USF, UCR, SCCWRP	AR-GeneBLAzer	96-Well format, Section S2-C.
43	RECETOX	MDA-kb2	Cells seeded at 50000/well; solvent control, medium and dihydrotestosterone (DHT 1 pM - 0.1 µM) tested on each plate.
44	RECETOX	Anti-MDA-kb2	Competing androgenic ligand: 0.1 nM DHT, agonist controls (0.01 µM and 0.1 nM DHT), medium, solvent control, and standard anti-androgen flutamide (10 nM - 10 µM) tested at each plate.
46	GU, BDS, CSIRO	GR-CALUX	BDS: Assay performed at 1% DMSO and in 384-well format, CSIRO: Cells were lysed in 50 µL of Triton lysis buffer. Luciferase assay substrate according to Brasier <i>et al.</i> (2001).
47	UA	GR Switchgear	Section S2-D.
49	RECETOX	GR-MDA-kb2	10 µM Flutamide was added to each sample dilution and agonist control (10 µM DHT) to inhibit androgenic activity; solvent control, medium and DHT (1 pM - 0.1 µM) tested on each plate.
50	GU, UF, USF, SCCWRP	GR-GeneBLAzer	96-Well format, see SI, Section S2-C; GU: ran as 384-well plate format, no change from original protocol.
51	GU	Anti-GR- GeneBLAzer	Agonist 0.4 nM mifepristone.
53	UF, USF, UCR, SCCWRP	PR-GeneBLAzer	96-Well format, Section S2-C.
54	GU, BDS, CSIRO	PR-CALUX	BDS: Assay performed at 1% DMSO and in 384-well format, CSIRO: Cells were lysed in 50 µL of Triton lysis buffer. Luciferase assay substrate according to Brasier <i>et al.</i> (2001).
58	BDS, GU	TR-CALUX	BDS: Assay performed at 1% DMSO and in 384-well format.
59	UQ	T-SCREEN	BDS: Assay performed at 1% DMSO and in 384-well format.
63	UQ	P19/A15	Section S2-E.
70	UA, IWW	Ames TA98	According to Xenometric manual (http://www.xenometrix.ch).
71	UA, IWW	Ames TA98+ S9	IWW: (Richard, 2012).
74	UQ, IWW	Ames TA100	UQ: none, IWW: (Richard, 2012).
79	UFZ	hspb11 induction	Different exposure time (0-120 hpf) was used.

#	Labora-tories	Bioassay	Method modification
		DART	
81	UA	Hypoxia-Switchgear	Section S2-F.
83	UQ	NF κ B - Geneblazer	96-well format, Section S2-C.
84	BDS	NF κ B -CALUX	Assay performed at 1% DMSO and in 384-well format.
85	GU	Jurkat E6.1 IkB	Section S2-G.
87	UA	Nrf2-keap	Section S2-H.
89	BDS	Nrf2-CALUX	Assay performed at 1% DMSO and in 384-well format.
91	BDS	p53-CALUX	Assay performed at 1% DMSO and in 384-well format.
92	BDS	p53-CALUX +S9	Assay performed at 1% DMSO and in 384-well format.
93	UF	p53-GeneBLAzer	96-well format, Section S2-B.
97	GU	Caco 2 NRU	Section S2-I.
98	CSIRO	RTG2 MTT	Exposure media were exchanged with media containing 0.5 mg/mL MTT rather than adding MTT solution directly to the wells. Incubation was for 3 hours at 22degC. MTT was solubilized with DMSO and absorbance determined at 540 nm.
101	GU	SK-N-SH cytotoxicity	Section S2-J.
102	GU	THP1 cytokine	Section S2-K.

S2-A. Two-hybrid CAR yeast assay

The experiments were performed according to Shiraishi *et al.* (2000) with the following modifications: Yeast cells that were introduced human constitutive androstane receptor (CAR) were cultured (30°C, overnight; Sanyo Incubator, Tokyo, Japan) in a modified SD medium supplemented with 0.88% glucose, lacking tryptophan and leucine. After centrifuge at 2000 rpm for 20 minutes, the medium was replaced by a fresh MSD medium. The yeast solution cell density was measured (595 nm), and, if necessary, cell density adjusted by diluting with MSD medium to readings to a constant 1.75 -1.85. MSD solution (60 μ L) was added to each well of the first row of a 96-well culture plate (Sumilon 96F disposable plates; Sumilon Bakelite Co., Tokyo, Japan). Thereafter, 2% DMSO / MSD solution (60 μ L) was automatically added (Nichiryo NSP-7000 Multi-

channel Auto Sampling System, Nichiryo Co., Tokyo, Japan) to each well of the 2nd - 8th rows of the plate. Six samples were run on each plate, with aliquots of each sample (60 µL) added to two neighboring wells of the 1st row of the plate. An aliquot was removed from each well of row 1 and added to row 2 to dilute 2-fold. This process was repeated from rows 2–7. No sample solution from row 7 was added to the 8th row. Thereafter, yeast solution (60 µL) was added to all wells, the plate shaken (30s; Taiyo S-2000 Automatic Mixer, Taiyo, Tokyo, Japan) and then incubated (30°C, 4 h).

After incubation, a mixed solution (80 µL) for inducing chemiluminescence and for enzymatic digestion (Aurora GAL-XE Reaction Buffer containing GalactaLux substrate, MP Biomedicals Inc., CA, USA and Zymolyase 100T diluted with Z buffer (a mixture of 21.5 g Na₂HPO₄·12H₂O; 6.2 g Na₂HPO₄·2H₂O; 0.75g KCl; 0.246 g MgSO₄·7H₂O in 1 L deionised water)) was then added to each well, and the plate incubated (37° C, 1 h; Ikemoto Scientific Technology Co, Tokyo, Japan). Thereafter, a light emission accelerator solution (50 µL; Aurora Accelerator, MP Biomedicals Inc., CA, USA) was added to each well, and the chemiluminescence produced by released β-galactosidase measured with a 96-well plate luminometer (Luminescencer-JNR AB-2100, ATTO Bioinstruments, Tokyo, Japan). 4-tert-octylphenol (Wako Pure Chemical Industries Ltd, Osaka, Japan) was used as positive control. A solvent (vehicle) control (DMSO, Nakalai Tesque Co., Kyoto, Japan) was also used.

S2-B. PPAR γ -GeneBLAzer assay

The commercially available PPAR γ -GeneBLAzer assay (Life Technologies, Vic, Australia) is based on a human embryonic kidney cell line (HEK 293H cells) modified to express a fusion protein combining the ligand binding domain of the human peroxisome proliferator-activated receptor γ (PPAR γ) fused with the DNA binding domain of the GAL4 gene, and stably transfected with a β-lactamase reporter gene downstream of a GAL4 activator sequence. When an agonist binds to the ligand-binding domain of the PPAR γ -GAL4 fusion protein, the protein binds to the activator sequence and stimulates expression of β-lactamase.

The division arrested (DA) kit was used here (cat no K1419, Life Technologies, Vic, Australia). In brief, the DA cell aliquot was thawed quickly in a 37°C water bath,

transferred to 10 mL of assay medium, and centrifuged at 200 ×g for 5 min. The supernatant was discarded, and the cell pellet reconstituted to a cell density of 9.4×10^5 cells/mL (determined using a Millipore Scepter Handheld Automated Cell Counter). Using a multi-channel pipette, 32 µL of assay medium was added to the "cell-free control" wells, and 32 µL of the cell suspension was added to all the other wells (30,000 cells/well) of a black wall clear bottom poly-D-lysine coated 384-well plate (cat no 354663, BD, NSW, Australia). In agonist mode, 8 µL of 5× 0.5% DMSO (solvent control), 5× rosiglitazone (reference compound, final concentration range from 7 pM to 2 µM) or 5× test samples were added to their respective wells (maximum 0.1% solvent in the final well for all test samples). In antagonist mode, 10× solutions of 0.5% DMSO (solvent control), 10× GW9662 (reference compound, final concentration range from 13 pM to 3.6 µM) or 10× test samples were pre-mixed 1:1 with 10× rosiglitazone agonist (for a final concentration in the well of 32 nM), and 8 µL of the resulting mix added to the respective wells for solvent control, reference compound or sample (maximum 0.1% solvent in the final well for all test samples).

The plate was then incubated for 16 h in a humidified 37°C/5% CO₂ incubator. At the end of incubation, 8 µL of 6× substrate mixture (provided in the kit) was added and the plate incubated for a further 2 h in the dark at room temperature. Fluorescence was then read with a plate reader (BMG Fluostar Omega; BMG Labtech, Vic, Australia) at 460 and 530 nm after excitation at 409 nm. Background fluorescence (determined in the cell-free control wells) was subtracted from all readings, and a β-lactamase expression ratio calculated by dividing the net fluorescence at 460 by net fluorescence at 530 nm. Samples were deemed as positive in agonist mode when they exceeded the 10 % effect concentration (EC₁₀; determined from the rosiglitazone standard curve) and in antagonist mode when they exceeded the 20 % inhibitory concentration (IC₂₀; determined from the GW9662 standard curve).

S2-C. GeneBLAZER assay panel

Reference chemicals were 17β-estradiol (E2), levonorgestrel (LEV), dexamethasone (DEX), and R1881 for the estrogen α, progesterone, glucocorticoid and androgen receptors (ER-α, PR, GR and AR), respectively. Chemicals were purchased from Sigma

Aldrich (E2, LEV, DEX) or PerkinElmer (R1881). Reference chemicals were diluted in assay-specific assay medium from a stock solution (E2: 2 µM, LEV: 20 µM, DEX, 0.1mM, R1881: 20 µM) prepared in DMSO. Nine dilutions, plus a DMSO-only control were made, and were based on previous optimization of reference compounds to include the entire linear range of fluorescent induction. Final DMSO concentration in diluted reference chemicals was 0.5%.

Water extracts were reconstituted in a total of 300 µL of DMSO. Four dilutions of each of the 10 water extracts were prepared by adding 5 µL of reconstituted water extract to assay-specific assay medium to the first dilution and then serially diluting 50 µL of first dilution and adding it to 47.5 µL of assay media and 2.5 µL of DMSO to the second and so forth. Final DMSO concentration in diluted reconstituted water extract was 0.5%.

Plate set-up was uniform across all laboratories, except in the case of USF, who did not have DMSO-control wells on the second plate. Reference chemicals and water extract dilutions were assayed simultaneously across two 96 well plates. Each reference chemical and water extract dilution was assayed in triplicate. Cell-free media, DMSO-control, and DMSO-free control were assayed in triplicate on each plate (except at USF).

The four reporter assays used in this assessment were GeneBLAzer® ER-Alpha, PR, GR and AR Division Arrested Assay Kits (Life Technologies, Carlsbad, CA). All kits were commercially bought except for the AR assay, which was manufactured for this assessment. Assays were bought as kits and optimized for use in a 96-well plate format, rather than the manufacturer suggested 384-well format. All procedures were performed in a Class II biological safety cabinet using sterile techniques. All media, chemicals and materials used in these assays were from manufacturer recommended sources. Modifications to the manufacturer's protocol were made to cell number, and media/dose volume to optimize the assay for 96-well format.

Cells stored under liquid nitrogen vapor were quickly thawed and transferred, drop-wise, into 10 mL of assay medium in a sterile 15-mL conical tube and centrifuged at $200 \times g$ for 5 minutes. Supernatant was aspirated and cell pellet was resuspended in 6 mL fresh assay medium. Cells were counted and diluted in assay medium to a density of 5.5×10^5 cells/mL. Ninety µL (50,000 cells) of cell suspension or assay medium (cell-free control wells) were added to each well. Ten µL of appropriate 10X reference chemical,

diluted water extract, or DMSO-added assay media were added to corresponding wells. Cells were incubated overnight (~ 16 h) at 37°C with 5% CO₂.

On the following day, 20 µL of six times concentrated loading solution, prepared according to the manufacturer's protocol, was added to each well. The plate was covered to protect from light and evaporation, and incubated at room temperature for 2 h in the dark. Fluorescent measurements were made according to manufacturer's instructions.

S2-D. Switchgear Assay for the Glucocorticoid Receptor

A commercially available GR assay kit (Switchgear Genomics, California) was used to evaluate the GR activity in water samples. The GR-Switchgear assay integrates the signal from four validated pathway-specific reporter vectors using the RenSP reporter gene. This is important, and unique to this assay, since there are numerous endogenous promoters for the gene and no single promoter can respond to all potential agonists and antagonists. Multiple validated house-keeping reporters, using the CLuc reporter gene, are also applied to monitor cell "health" during the assay, also unique to this particular assay. The assay is a transient-transfection assay, which means that plasmids containing the reporter genes are freshly transfected each time the assay is performed.

A human fibrosarcoma cell line (HT1080) was maintained in standard growth medium composed of 500 mL minimal essential media (MEM), 5 mL GlutaMax, 50 mL Fetal Bovine Serum (FBS) (Heat inactivated) and 5 mL Pen/Strep. Cells were thawed from liquid nitrogen and passaging was carried out in 75 cm² flasks every 2-3 days. The 2nd generation cells were used in this assay when they reached greater than 80% confluence. Cell density was controlled at 1×10⁵ cells/mL in stripped growth medium (charcoal stripped FBS, without antibiotics). The transfection reagent containing the 4 GR plasmid constructs and housekeeping constructs were thawed, mixed, and incubated for 30 min at room temperature. The transfection mix was then added to cell medium, thoroughly mixed, and 100 µL added to each well of a 96-well white tissue culture (TC) plate. In a separate 96-well clear TC plate, an aliquot of 100 µL was added to cells in 12 wells for visual monitoring of cell viability and growth. Both plates were incubated at 37°C in a CO₂ incubator for 12-16 h.

After overnight culture, the medium was replaced by 90 µL fresh stripped FBS growth medium and 10 µL of water sample extract diluted in 10% of stripped medium. After 18-24 h of exposure, 10 µL of the cell supernatant was transferred to a secondary white 96-well TC plate and both plates were frozen at -80°C. Substrate and buffer solutions were added after the plates were thawed. Luminescence was measured to determine for the luciferase reporter gene activity (LightSwitch Dual Assay System). DEX was used as the positive control and a negative control and solvent control were also included for quality control.

S2-E. p19/A15 assay for induction of the retinoic acid receptor (RAR)

p19/A15 Cells were grown in Dulbecco's MEM with sodium pyruvate and L-glutamine, high glucose, 10% FBS, 1% penicillin-streptomycin, 1.6% non essential amino acid (NEAA) was obtained from Gibco, Australia. Cells were grown in T75 flasks in 11 mL Dulbecco modified minimal essential medium (DMEM) and incubated at 37°C and 5% CO₂ and passaged every 2-3 days when cells were 70% confluent.

For an exposure experiment, cell concentration was adjusted to 100,000 cells/mL and 100 µL was transferred in each well of a white polystyrene tissue culture treated 96-well microplate (Corning). The plates were then incubated for 24h at 37°C and 5% CO₂ and dosed with the appropriate amount of chemical or extract. Each plate should include one serial dilution of atRA ($3.23 \cdot 10^{-11}$ M to $3.85 \cdot 10^{-18}$ M) or 9-cis RA ($2.43 \cdot 10^{-07}$ M to $2.89 \cdot 10^{-14}$ M) as positive control and one row of medium only. The plates were then covered with PCR-SP plate sealer from Axygen and incubated for 24 h before cytotoxicity or induction was assessed.

A typical experiment consists of two steps, where each step is performed in duplicate. First, a range finder with a serial (2-fold) dilution series was performed, where induction of RAR pathway and cytotoxicity were evaluated. Interference by cytotoxicity causes a suppression of the induction signal and such concentrations cannot be used for the induction data evaluation. Second, non-cytotoxic concentrations/dilutions of the water sample were selected and a linear dose-response curve was measured for induction only. Often the window between induction and cytotoxicity is small and no maximum

induction can be reached, therefore concentrations should be selected in a way that the maximum induction ratio is 5.

As a control, the cell viability was assessed with the MTS assay. MTS (tetrazolium) is bioreduced by cells into an aqueous, soluble formazan product by dehydrogenase enzymes found in metabolically active cells (Mosmann, 1983). When cells die, the ability to reduce these products is rapidly lost due to mitochondrial dysfunction. The absorbance of the formazan product at 490 nm can be measured directly from 96-well assay plates without additional processing, and the amount is directly proportional to the number of living cells in culture. After 24 h incubation the medium in each plate was replaced by 120 µL MTS (CellTiter 96® AQ_{ueous} One Solution Cell Proliferation Assay (Promega) with MTS and phenazine methosulfate as the electron coupling reagent) in Hyclone DMEM without phenol red (Thermo Scientific) and absorbance at 492 nm was read after 2 h incubation.

S2-F. Switchgear Assay for Hypoxia

The commercially available hypoxia assay kit (Switchgear Genomics, California) uses the HT1080 cell line with transient transfection of three reporter constructs including lactate dehydrogenase (LDHA) promoter, H1F1a promoter, housekeeping gene ACTB (ACTB_PROM) from Switchgear, in which H1F1a is a well-known hypoxia inducible transcription factor.

In brief, transfection reagent, which contained the three plasmid constructs (LDHA, H1F1a, and ACTB), were first thawed from -20°C and incubated for 30 min at room temperature. The human fibrosarcoma cell line HT1080 was thawed quickly in a 37°C water bath from the -80°C freezer, and the thawed cells were immediately added to the growth medium which was composed of Eagle's minimum essential medium (EMEM, ATCC #30-2003), 10% normal FBS, 1% of GlutaMax and 1% of PenStrep. To get 20,000 cells per well, cell density was maintained at 2.1×10^5 cells/mL. Then the transfection reagents were mixed with the cell and medium solution at a ratio of 5:95. Using a multi-channel pipette, 100 µL of the transfected cell mixture was aliquoted to each well of a white 96-well tissue culture plate. In a separate clear 96-well tissue culture plate, 100 µL of cells were aliquoted to 12 wells for visual monitoring of cell viability

and growth. Both plates were incubated at 37 °C in a CO₂ incubator for 12-16 h. After overnight culture, the medium was replaced by 90 µL of fresh charcoal-stripped FBS growth medium and 10 µL of sample, which was diluted in 10% of stripped medium in advance.

After 24 h of exposure, 10 µL of the supernatant was transferred to a secondary white 96-well tissue culture plate, and both of the plates were frozen in -80°C for better sensitivity. Substrate and buffer solution were then added after the plates were thawed, and luminescence was quantified as a measure of luciferase activity (LightSwitch Dual Assay System, available in the kit). Desferrioxamine (DFO) was used as the positive control. Negative control and solvent control were included for quality control.

S2-G. Jurkat E6.1 IkB

In the assay Jurkat E6.1, cells were resuspended in Roswell Park Memorial Institute medium (RPMI without phenol red supplemented with 5% charcoal-stripped fetal bovine serum) at 1×10⁶ cells/mL (determined using a Millipore Scepter Handheld Automated Cell Counter). Cells were then seeded at 200,000 cells/well by adding 200 µL of cell suspension to the 48 inner wells of a flat bottom standard 96-well plate, and the test samples were added in 50 µL of white media (maximum 0.1% final solvent concentration). The remaining wells were filled with 250 µL of phosphate buffered saline (PBS) to act as a humidity barrier, and the plate incubated for 24 h in a humidified 37°C/5% CO₂ incubator. A geometric dilution series of phorbol-12-myristate-13-acetate (PMA) was used as a reference compound, with final concentrations in the well ranging from 0.2 nM to 0.2 µM.

After incubation, the content of each well was gently mixed and 200 µL was transferred to a v-bottom 96-well plate. The plate was centrifuged at 300 ×g for 5 min, and 150 µL of the supernatant was discarded (paying particular attention not to disturb the cell pellet). The pellet was rinsed with 100 µL of warm sterile PBS and the plate centrifuged again at 300 ×g for 5 min. After centrifugation, 100 µL of the supernatant was discarded (again paying particular attention not to disrupt the cell pellet). IkB concentration in the cell pellet was then determined using a commercially available ELISA kit (IkBα Total InstantOne ELISA; cat no 85-86061, eBioscience), with minor modifications. In brief, cells were lysed with 1.5× lysis mix added in the v-bottom 96-

well plate directly and mixed by aspirating/dispensing with a multi-channel pipette, then placed on an orbital shaker at 300 RPM for 10 min at room temperature. Then, 50 µL of cell lysate were transferred into the InstantOne assay plate (provided with the kit) followed by 50 µL of IκB antibody cocktail (provided with the kit). A negative and positive IκB control, provided with the kit, were also tested with every ELISA run.

The plate was covered with an adhesive seal and incubated for 1 h at room temperature on a microplate shaker at 300 RPM. The wells were washed with 200 µL of wash buffer (provided with the kit), all liquid removed by inverting on a paper towel, and 100 µL of detection reagent (provided with the kit) was added to each well. The plate was incubated for 10 min at room temperature on a microplate shaker at 300 RPM, and the reaction stopped by adding 100 µL of stop solution. The absorbance of each well was then measured with a plate reader (BMG Fluostar Omega; BMG Labtech, Vic, Australia) at 450 nm.

S2-H. Nrf2-keap cell line

The human breast cancer cell line MDA-MB-231-745, which was transfected with the antioxidant response element (ARE) luciferase plasmid (Villeneuve *et al.*, 2008), was donated by Prof. Donna Zhang at Department of Pharmacy, the University of Arizona. The standard growth medium was composed of minimal essential medium (MEM, Life Tech, #11095-080), 10% FBS, 1% L-glutamine, 0.1% Gentamycin, 6 ng/mL Insulin, 2 mM HEPES and 1.5 µg/mL puromycin. Cells were thawed from liquid nitrogen and passaging was carried out in 75 cm² flasks every fourth day. The 4th generation cells at more than 80% confluence were used in this assay. Cell density was controlled at 2×10⁵ cells/mL. Using a 8-channel pipette, 100 µL of the cell solution were seeded into one white 96-well plates and one clear 96-well plate (cytotoxicity test).

After overnight culture in a CO₂ incubator for 16 h (5% CO₂, 90% humidity), the medium was replaced by 90 µL of fresh growth medium and 10 µL of sample, which was diluted in 10% of growth medium in advance. All samples were tested in triplicate including the medium blank and solvent blank. Tert-butylhydroquinone (tBHQ) was used as the positive control and the solvent used was methanol. After another 16 h of exposure in the CO₂ incubator, the medium in the white plates was removed, and washed with PBS. Twenty-five microlitres of Passive Lysis Buffer (PLB) was then added and the plates

were shaken for 15 min before luciferase analysis. Gen5 micro-plate reader with a delivery pump was used for the measurement and the luminescence was read directly by well after luciferase buffer (pH=7.8) was added. For the cytotoxicity test, after 16 h exposure the medium was replaced by 100 µL of clear fresh medium (without phenol red) and 20 µL of MTS solution (Promega, #G3580). Absorbance at 492 nm was read after 2 h of incubation.

S2-I. CaCo2 NRU assay

The Caco2-NRU (neutral red uptake) test is a measure of non-specific cytotoxicity. It is used to determine if the test sample impacts the viability of Caco2 (human epithelial colorectal adenocarcinoma) cells after 21h of exposure. Cell viability at the end of the incubation period is determined by adding neutral red, a dye that stains only live cells, and measuring the amount of dye taken up by the cell culture. The method was adapted from Konsoula and Barile (2005). In brief, Caco2 cells were grown in DMEM/F12 with phenol red supplemented with 8% FBS and 100 µM non-essential amino acids.

For the assay, cells were seeded at 20,000 cells/well in 100 µL assay medium (DMEM/F12 medium without phenol red supplemented with 5% stripped FBS (CD-FBS) and 100 µM non-essential amino acids) in 96-well plates and incubated for approximately 24h at 37°C 5% CO₂. When the cells reached confluence (usually about 24h), the medium was removed and replaced by 150 µL of fresh assay medium and 50 µL of assay medium containing the model compound or water extract to be tested (final methanol concentration in the assay plate ≤0.1%). After 21 h of incubation at 37° C 5% CO₂, the medium was removed, the wells rinsed with 150 µL PBS, and the PBS was replaced by 150 µL of neutral red media (50 µg/mL neutral red in assay media, made fresh). After a further 3h incubation at 37°C 5% CO₂, the medium was aspirated and replaced with 150 µL of neutral red desorbing fixative (1% acetic acid, 50% ethanol, in deionized water). The plate was placed on an orbital shaker at 600 rpm for 10 min at room temperature and absorbance was read in a plate absorbance reader (BMG FluoSTAR Omega) at 540 nm.

S2-J. Human neuroblastoma (SK-N-SH) cytotoxicity

Human neuroblastoma cells (SK-N-SH cells) were resuspended in white medium (DMEM/F12 without phenol red supplemented with 5% FBS, 1× non-essential amino

acids and 2 mM Glutamax; Life Technologies, Vic, Australia) at 1×10^5 cells/mL (determined using a Millipore Scepter Handheld Automated Cell Counter). Using a multi-channel pipette, 200 μ L of cell suspension was added to every well (20,000 cells/well) of a standard flat bottom 96-well plate, and the plate was incubated for 24 h in a humidified 37°C/5% CO₂ incubator.

The medium was then removed by aspiration and replaced with 200 μ L of fresh white media containing the test sample (maximum solvent concentration of 0.5%), and the plate incubated 21 h in a humidified 37°C/5% CO₂ incubator. The media was then aspirated, the wells rinsed with 150 μ L of PBS, the PBS aspirated, and 150 μ L of neutral red media (50 μ g/mL neutral red solution, prepared fresh) was added. The plate was then incubated a further 3 h in a humidified 37°C/5% CO₂ incubator.

At the end of the incubation, the medium was aspirated, the wells rinsed with 150 μ L PBS, the PBS was aspirated and 150 μ L of neutral red desorbing fixative (1% acetic acid, 50% ethanol, prepared in ultrapure water) was added. The plate was placed on an orbital shaker at 600 rpm for 10 min and the absorbance was read at 540 nm in a plate reader (BMG Fluostar Omega; BMG Labtech, Vic, Australia). DMSO was used as a reference compound, with an IC₁₀ and IC₅₀ of approximately 50 and 500 mM, respectively. Samples were deemed as "neurotoxic", when cytotoxicity exceeded IC₁₀ (determined from the DMSO standard curve).

S2-K. THP1 cytokine assay

The THP1 cytokine assay provides a measure of immunotoxicity. For this assay, we monitored interleukin 1 β (IL1 β). The assay was run in antagonist mode, by measuring the inhibition of the normal production of IL1 β by THP1 cells exposed to *E. coli* lipopolysaccharide (LPS) after exposure to the sample for 24h. The methods were adapted from Baqui *et al.* (1998). In brief, THP1 cells were cultured in growth medium (DMEM/F12 medium with phenol red supplemented with 8% FBS and 100 μ M non-essential amino acids).

For the assay, cells were seeded at 200,000 cells/well in 200 μ L of growth media (with 1 μ g/mL LPS in antagonist mode) and 50 μ L of assay medium containing the model compound or water extract to be tested (final methanol concentration in the assay plate \leq 0.1%). After 24 h incubation at 37°C 5% CO₂, cells were transferred to a V-bottom

96-well plate, centrifuged at 300×g for 5 min, and the supernatant was transferred to a fresh 96-well plate. IL1 β concentration in the supernatant was assayed by ELISA (Human IL1 β quantikine ELISA, RnD Systems), following the supplier's instruction.

Section S3. Additional information on data evaluation

The dose-metric of the concentration-effect curves is the relative enrichment factor REF, which is the combination of the enrichment of the extraction and the dilution in the bioassay (Eq. 1), thereby representing the enrichment (REF > 1) or dilution (REF < 1) of the original sample in each bioassay. The REF is expressed in the units of [L_{water sample}/L_{bioassay}].

$$\text{REF} = \text{dilution factor}_{\text{bioassay}} \cdot \text{enrichment factor}_{\text{SPE}} \quad (1)$$

The enrichment factor of the SPE enrichment factor_{SPE} was calculated using Eq. 2 from the volume of extracted water to the volume of resulting extract (in solvent).

$$\text{enrichment factor}_{\text{SPE}} = \frac{V_{\text{water}}}{V_{\text{extract}}} \quad (2)$$

The dilution factor of each bioassay was calculated using Eq. 3.

$$\text{dilution factor}_{\text{bioassay}} = \frac{\text{volume of extract added to bioassay}}{\text{total volume of bioassay}} \quad (3)$$

S3-A. EC₁₀ (10% effect concentration)

EC₁₀ values were reported for the cytotoxicity bioassays and for receptor-mediated effects and were obtained from a log-logistic fit of the concentration-effect curves (Figure 2A in the main article). The % effect was calculated with Eq. 4

$$\% \text{effect} = \frac{\text{signal}_{\text{sample}} - \text{signal}_{\text{control}}}{\text{signal}_{\text{max}} - \text{signal}_{\text{control}}} \quad (4)$$

Adjustable parameters were the slope and the effect concentration causing 50% reduction of maximum effect, EC₅₀.

$$\% \text{effect} = \frac{1}{1 + 10^{\text{slope}(\log \text{EC}_{50} - \log \text{concentration})}} \quad (5)$$

The EC₁₀, the effect concentration causing 10% reduction of cell viability, was derived from the EC₅₀ and the slope s (Eq. 6).

$$\log \text{EC}_{10} = \log \text{EC}_{50} + \frac{1}{s} \log \left(\frac{1}{9} \right) \quad (6)$$

In many cases, no full concentration-effect curves were obtained for the sample extracts. Partial concentration-effect curves can only be fitted if the slope is fixed at 1 or at the slope of the reference compound. Alternatively, because the lower portion of the log-logistic concentration effect curves is linear with respect to non-logarithmic concentrations, the EC₁₀ can also be derived from a linear concentration-effect curve with intercept zero up to 20% of maximum effect (Eq. 7, Figure 2C in the main article).

$$\% \text{effect} = \text{slope} \cdot \text{concentration} \quad (7)$$

$$\text{EC}_{10} = \frac{10\%}{\text{slope}} \quad (8)$$

The EC₁₀ values derived with the linear method agreed well with the log-logistic derivation and the final results of the samples were derived from the linear concentration-effect curves, although the EC₁₀ values of the reference compounds were from the full log-logistic fit.

S3-B. EC_{IR1.5} (effect concentration causing an induction ratio IR of 1.5)

The EC_{IR1.5} was derived for all reporter gene assays where no maximum response could be obtained. By nature of the endpoint the IR approach applies to genotoxicity and most

adaptive stress responses such as the oxidative stress response. In addition, a few of the endpoints assessed had no reference compounds, e.g., the FACTORIAL. For all endpoints where no reference compounds were tested, the EC_{IR1.5} was deduced.

The IR is the ratio of the measured signal (e.g., absorbance, relative light units (RLU), relative fluorescence units (RFU)) to its control value (Eq. 9). An analogous equation can be used for the number of revertants in the Ames assay, hence called revertant ratio (RR).

$$\text{IR} = \frac{\text{signal}_{\text{sample}}}{\text{signal}_{\text{control}}} \quad (9)$$

Concentration-effect (IR) curves would show the typical log-logistic form but the maximum is hard to establish due to cytotoxicity interference or it may not even exist (Figure 2B in the main article). Therefore only the linear portion of the concentration-effect curves was evaluated up to an IR of 5 (Eq. 10, Figure 2D in the main article).

$$\text{IR} = 1 + \text{slope} \cdot \text{concentration} \quad (10)$$

The assessment endpoint is the concentration that induces an IR of 1.5 (EC_{IR1.5}). The EC_{IR1.5} can be derived using the linear regression function with Eq. 11 (and analogously for the revertant ratio in the Ames test with Eq. 12).

$$\text{EC}_{\text{IR1.5}} = \frac{0.5}{\text{slope}} \quad (11)$$

$$\text{EC}_{\text{RR1.5}} = \frac{0.5}{\text{slope}} \quad (12)$$

The threshold of 1.5 was selected because (a) it is employed in several guideline documents, e.g., umuC genotoxicity assay, (b) it is very close to the limit of detection in many cases (control plus 3 standard deviations) (Escher *et al.*, 2012), (c) it is an interpolation not an extrapolation such as the EC₅₀, and (d) it can be applied if the maximum of the dose-response curve is not known. The disadvantage of using IR is that depending on the bioassay, maximum response can be at IR of 2 up to over 100. If the maximum IR reaches 6, then the EC_{IR1.5} is equivalent to the EC₁₀, if the maximum ER is 2, the EC_{IR1.5} is equivalent to the EC₅₀, and for IRs that level off at 100 or more, the EC_{IR1.5} is often close to the limit of detection.

For the 22 assays for which the maximum effect and EC₁₀ were derived, it was possible to calculate what % effect would be equivalent to IR 1.5 (Table S3). For 13 of these bioassays, the max IR fell between 3 and 15 allowing the EC₁₀ and the EC_{IR1.5} to be directly compared. For nine, the maximum IR was well above 15 up to 940 and in these cases the EC₁₀ would be inherently less responsive than the EC_{IR1.5}.

Table S3. Comparison of maximum IR with effect level at IR 1.5.

#	Laboratories	Bioassay	Maximum IR	Effect level at IR 1.5
9	BDS	CALUX-PPAR α	5	13%
10	BDS, CSIRO	CALUX-PPAR γ	20	3%
12	GU	PPAR γ GeneBLAzer	3	25%
15	UQ, RECETOX	AhR-CAFLUX	13, 2	4%, 50%
16	RECETOX	H4IIEluc	5 to 11	5 to 13 %
22	GU, CSIRO, BDS, IWW	ER-CALUX	15, 15, 5, 15	3, 3, 11, 3%
23	UQ	E-SCREEN	3 to 26	2 to 23%
24	Swiss, CSIRO, UA	YES	100	0.5%
30	RECETOX	hER α -HeLa-9903	3 to 5	13 to 25%
31	HK	MCF7-ERE	6	10%
35	UFL, USF, UCR, SCCWRP	ER-GeneBLAzer	24, 27, 5	3, 3, 13%
37	GU, BDS, CSIRO	AR-CALUX	10 to 60, 45, 20	1 to 60, 1, 3%
40	UA, CSIRO	YAS	100, 54	1, 1%
41	UFL, USF, UCR, SCCWRP	AR-GeneBLAzer	3	26%
43	RECETOX	MDA-kb2	6	10
46	GU, BDS, CSIRO	GR-CALUX	15 to 40, 15, 20	1 to 4, 4, 3%
47	UA	GR Switchgear	15	4%
50	GU, UFL, USF, UCR, SCCWRP	GR-GeneBLAzer	22, 25, 20, 50, 22	2, 3, 3, 1, 2%
53	UFL, USF, UCR, SCCWRP	PR-GeneBLAzer	6, 9, 3, 5	10, 6, 25, 15%
54	GU, BDS, CSIRO	PR-CALUX	50, 110, 940	1, 0.5, 0.05%
58	BDS, GU	TR-CALUX	940, 40	0.05, 1%
59	UQ	T-Screen	5	12%

S3-C. EC_{SR0.2} (effect concentration causing a suppression ratio SR of 0.2) for all antagonistic effects and chaperon dissociation.

A receptor-binding bioassay is run in antagonistic mode if the receptors are saturated or occupied with a constant concentration of an agonist (positive control). In an antagonistic mode experiment, varying concentrations of sample are added, while the concentration of the agonist is kept constant. If the signal of the agonist is suppressed the sample has an antagonistic effect (Figure S1). The suppression ratio SR is defined by Eq. 13. The analogous equation was used for endpoints that are based on chaperon dissociation (CD).

$$SR = 1 - \frac{\text{signal}_{\text{sample}} - \text{signal}_{\text{control}}}{\text{signal}_{\text{agonist}} - \text{signal}_{\text{control}}} \quad (13)$$

Signal_{agonist} refers to the signal (relative fluorescence/light units (RFU, RLU), etc.) measured in presence of the agonist (positive control), which is normally the highest signal obtained unless the agonist was not added at saturating concentrations and the sample also had an agonistic effect. If signal_{agonist} >> signal_{control}, Eq. 13 simplifies to Eq. 14.

$$SR = 1 - \frac{\text{signal}_{\text{sample}}}{\text{signal}_{\text{agonist}}} \quad (14)$$

In most cases no full concentration-effect curves were obtained for antagonistic effects. Therefore, we used only the initial linear part of the concentration-effect curves up to a suppression ratio of 0.3 (Figure S1). The EC_{SR0.2} is calculated from a linear regression through the zero point (Eq. 15). The 20% suppression level (SR 0.2) was chosen to derive the EC_{SR0.2} (Eq. 16) because the variability is typically larger than in the agonist mode and the 10% suppression level (SR 0.1) is often not above the variability of the controls, which would produce false-positive results.

$$SR = \text{slope} \cdot \text{concentration} \quad (15)$$

$$EC_{SR0.2} = \frac{0.2}{\text{slope}} \quad (16)$$

Analogously, an EC_{CD0.2} was defined for chaperone dissociation.

$$EC_{CD0.2} = \frac{0.2}{\text{slope}} \quad (17)$$

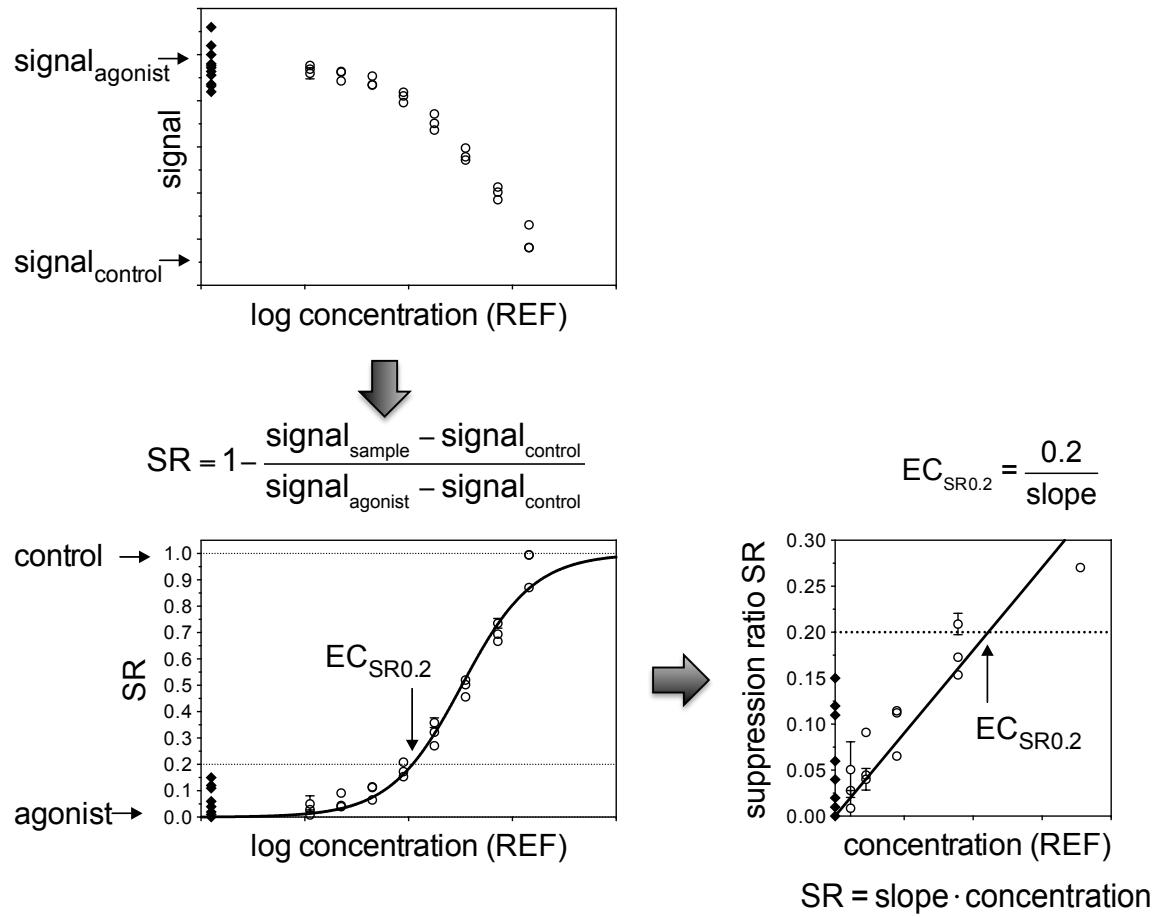


Figure S1. Derivation of $EC_{SR0.2}$.

Section S4. How robust were bioassays performed in different laboratories?

A number of bioassays were performed in multiple laboratories. In the following sections only mean results per bioassay are reported. For bioassays that had some positive and some negative results in two or more laboratories, the reported mean was obtained as follows:

- (a) If no effects were observed until the highest REF tested in one laboratory but the maximum experimental REF used in that laboratory was lower than the REF in another laboratory, then the results from the laboratory with lower maximum enrichment was ignored.
- (b) If the maximum REF were similar but some laboratories reported effects, others not, then the data without observed effect was not used to calculate the mean if it was only one laboratory out of three or four. If the test was only performed by two laboratories, and one was “>highest REF tested”, the other one was not, then the positive data were used.

The bioassays listed in Table S4 were performed multiple times and a repeated measures one-way ANOVA followed by the Bonferroni multiple comparison post-test was performed to assess if the results matched between the different laboratories and if it was legitimate to report mean values for each endpoint only (if only two laboratories were involved, then a paired t-test was used). We did not use ANOVA *per se* to quantify the distribution of the ten samples. Clearly, we cannot assume that these 10 samples follow a Gaussian distribution, and we did not target a true mean of all samples, but we can assume that if thousands of water samples of these types were tested they would follow a Gaussian distribution. Therefore a non-parametric test was not suitable for this analysis. ANOVA is suitable in this case as with the independent testing in two or more laboratories the circularity assumption is fulfilled.

Table S4. Comparison of the bioassay results for the same bioassay performed in different laboratories.

Bioassay	Labo- ratory	Test	# Sample s com- pared (>LOD)	Is there signi- ficant pair- ing?	P	r2 (r for t-test)	Bonferroni post test
Photo- synthesis inhibition	UQ, Swiss	Paired t-test	4	yes	0.014	0.973	
Algae growth inhibition	UQ, Swiss	Paired t-test	3	yes	0.039	0.993	
AhR- CAFLUX	UQ, RECETOX	Paired t-test	6	yes	0.008	0.894	
ER-CALUX	GU, BDS, CSIRO, IWW	Repeated measures ANOVA	3	yes	0.001	0.694	CSIRO vs. BDS signifi- cantly different
ER- GeneBLAzer	UF, USF, UCR, SCCWRP	Repeated measures ANOVA	4	yes	0.001	0.812	
GR- GeneBLAzer	UF, USF, UCR, SCCWRP	Repeated measures ANOVA	2+8 <LOD ^a	yes	<0.00 1	0.957	All laboratories not significantly different; the non-detects were included in the analysis because there were too few detects
GR-CALUX	BDS, CSIRO	Paired t-test	2	no	0.451	0.153	GU used lower REF ^b (and results were below LOD), therefore only BDS and CSIRO compared.
Microtox	UQ (1+2) Swiss	Repeated measures ANOVA	8	yes	0.046	0.531	All laboratories not significantly different

^aLOD = limit of detection, ^bREF = relative enrichment factor.

Also, data were log-transformed before the test was performed to assure that small values had equal weight. We used this test to obtain P values that test the null hypothesis of the population row means being equal.

Although the maximum amount of sample extract to be dosed was not prescribed, which lead to different laboratories testing different highest doses, we found generally good consistency between the results of the different laboratories.

Four laboratories performed the ER-CALUX and their results paired up although the CSIRO vs. BDS results were significantly different but the two other datasets lay well in the middle and therefore all four were averaged.

For the YES assay, two laboratories were below detection limits in all samples but the dataset of a third laboratory showed clear and differentiated results and high quality raw data. In this case therefore only the positive results of the third laboratory were used despite violation of the above rules.

Working with a semi-standardized protocol was beneficial as the four laboratories that performed the ER-GeneBLAzer assay achieved consistent results despite the fact that the assay was newly established in all laboratories. All ER-GeneBLAzer results were averaged.

The bioassays for androgenicity, *e.g.*, the AR-CALUX and the YAS, as well as for the progesterone receptor, PR-CALUX, and for the thyroid receptor, TR-CALUX, did not reach the effect threshold of 10% at the highest REF in any of the laboratories where the test was applied. The AR-GeneBLAzer and the PR-GeneBLAzer were performed in four and three laboratories, respectively. In a few cases individual samples were just above detection limit but in each case they were below detection limit in all other laboratories and they were therefore assigned as non-detects. It must be noted that this study was limited by the amounts distributed to each laboratory. Therefore it was not possible to do more repetitions, which one might normally do with results close to the detection limit. The GR-GeneBLAzer assay was consistent between four laboratories showing positive results in the same two samples and with all other samples below the significance threshold of effect of 10%. The results were averaged.

The GR-CALUX was one of the few assays where the different laboratories did not agree although the relative effect pattern was consistent. One laboratory tested much lower concentrations than the others and did not detect any effect, the two other

laboratories showed two samples that were positive (MF and Eff2) but a third sample (Eff1) that was cytotoxic in one laboratory and inducing in another lab. The effects of the two positive samples differed more than in other bioassays and pairing could not be established. Therefore only the data set from the laboratory that reported highest activation and no cytotoxicity was reported in the final table.

For reactive toxicity, we tested a large variety of Ames strains with slightly different properties and therefore were unable to narrow down specific patterns, although similar samples caused induction. IWW duplicated some of the tests but obtained no responses due to smaller REFs applied and these datasets were omitted as other laboratories had positive results with the same strains. For adaptive stress response pathways, the only assay performed in duplicate was the p53-CALUX, which was not responsive.

Section S5. Initial screening of nuclear receptors and transcription factors

The FACTORIAL bioassays were applied here for the first time to screen water samples. The raw water samples did not show any effects (data not shown), the responses discussed in this section relate to water samples after enrichment by SPE to a REF of 4. As no reference compounds were measured and the maximum response was not known, only induction ratios could be calculated from the raw response data. For more than 90% of the tested endpoints, the limit of detection (LOD; IR of control plus 3 times standard deviation) fell below an IR of 1.5 (Figure S2).

A fixed threshold has the advantage that datasets of variable size can be compared, while the LOD is dependent on the number of datapoints from which it is derived. For example, in Figure S2 the LOD calculations for the nuclear receptors (NR) was based on 24 control data points while the LOD for the transcription factors (TF) was based on 48 datapoints. The latter yielded much lower variability and lower LODs. Therefore we opted to use the IR 1.5 as a threshold of effect for the FACTORIAL bioassays but also for all other bioassays where no % effect could be calculated.

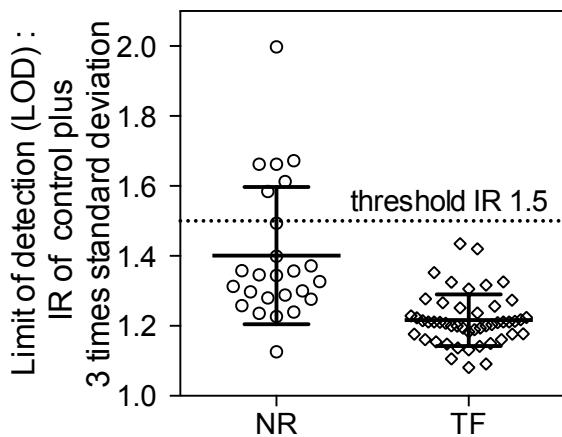


Figure S2. Limit of detection (LOD) calculated from the induction ratio (IR) of control plus 3 times standard deviation for the 25 nuclear receptors (NR) and 48 transcription factors (TF) of the FACTORIAL bioassay.

All samples were screened with the FACTORIAL bioassays and the results are depicted in Figure S3. The radar plots in Figure S3 depict the induction ratios IR directly

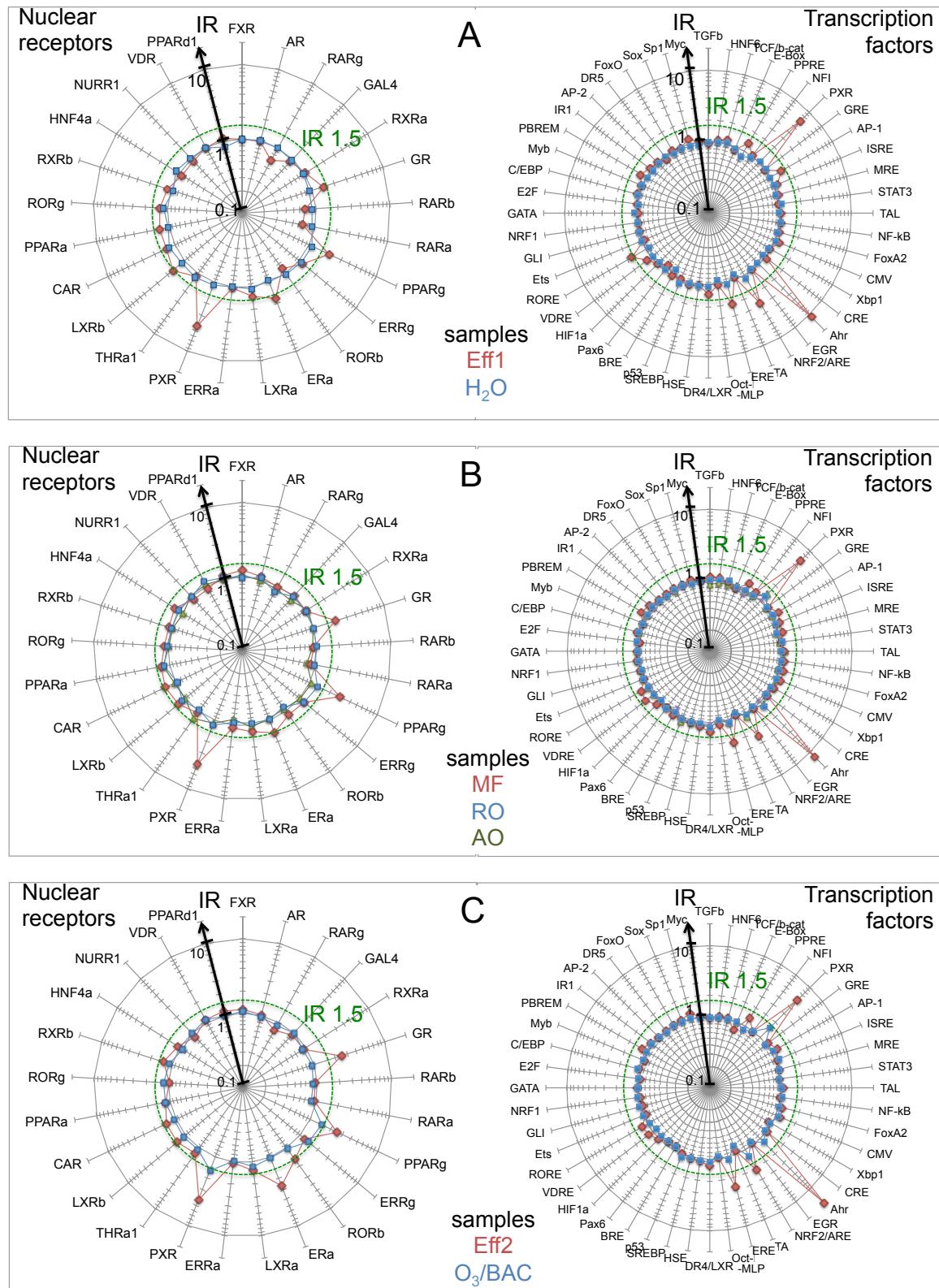
measured after incubation of the HepG2 cells with water samples at an REF of 4. A high IR relates to a high effect, the controls are IR = 1 and the threshold of effect was defined as IR = 1.5 as this is also the IR for the derivation of the EC_{IR1.5}.

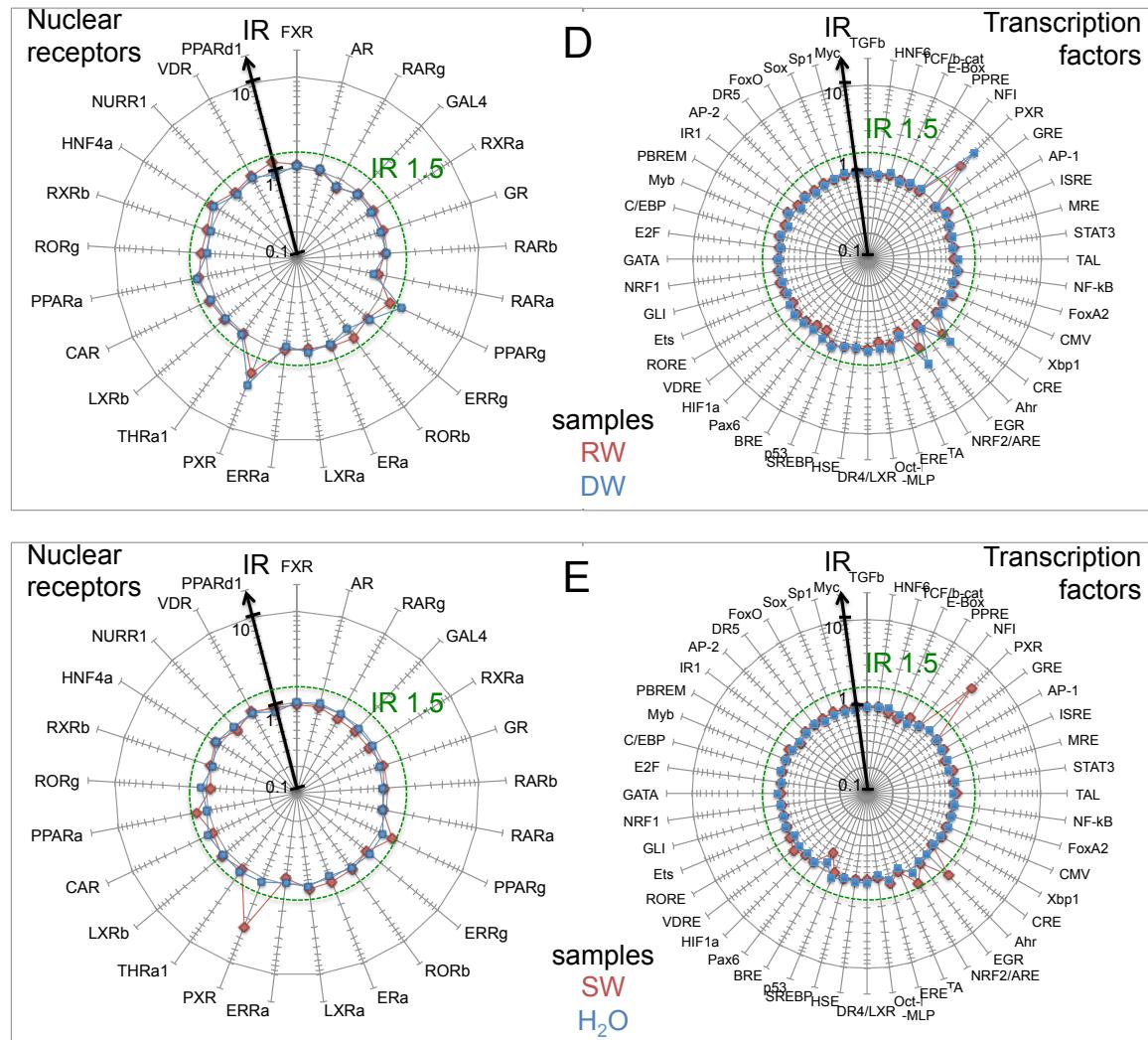
The highest induction was seen for the pregnane X receptor PXR both in the NR and TF assay and in all samples but the blank (H_2O) (Figure S3). As expected the estrogen receptor ER α was activated but not the estrogen related receptors ERR α and ERR γ in the NR assay and the estrogen response element ERE was activated in the TF assay (Figure S3).

In the NR assay, the peroxisome proliferator activated receptor PPAR γ was active but with a lower IR around or below 2 in the samples Eff1, Eff2, MF. The glucocorticoid receptor responded weakly in the NR assay (IR of 1.4 at an REF of 4, i.e. just below the threshold of effect) but showed no response in the TF assay.

The highest induction in the TF assay was observed for the arylhydrocarbon response element, which does not come as a surprise as a large number of chemicals activate this xenobiotic metabolism pathway. Activity measured in Eff1, Eff2 and MF disappeared after further treatment, and drinking water treatment marginally increased the effect.

The next highest activity was caused by the response element associated with the PXR and this is consistent with the high activity in the NR assay. Third in activity was the antioxidant response element (ARE) that is activated through the Keap-Nrf2 pathway. Then came the response elements for ER (ERE), the retinoic acid receptor (RAR)-related orphan receptor (RORE) and response element for the PPAR (PPRE; IR 1.4 just below effect threshold).





*Figure S3. Screening of 25 nuclear receptors and 48 transcription factors with the FACTORIAL bioassay. The induction ratios (IR) are depicted on the radar scale and the effect threshold of IR 1.5 is depicted with a green dashed circle. Activity profile of the induction ratios IR of nuclear receptors (left) and transcription factors (right) in HepG2 cells incubated for 24h with water samples at a REF of 4; (A) wastewater treatment plant effluent (*Eff1*) and the blank (milliQ water) (B) microfiltration (MF), reverse osmosis (RO), advanced oxidation (AO), (C) *Eff2* and ozone/biologically activated carbon (O_3/BAC), (D) river water (RW) and drinking water (DW), (E) stormwater (SW) and laboratory blank (H_2O).*

Section S6. Additional information on the bioassay results

Details of the bioassay results are depicted in plots that are structured as shown in Figure S4. The EC values were plotted in an inverse scale so that the most toxic ECs were on the top and the least toxic ECs on the bottom (Figure S4-A). The samples were grouped according to the treatment processes (Figure S4-B). Sensitivity cannot be compared directly because of the two different endpoints (EC_{10} and $EC_{IR1.5}$) but these plots give an indication about the responsiveness and thus the suitability of the bioassays for water quality assessment.

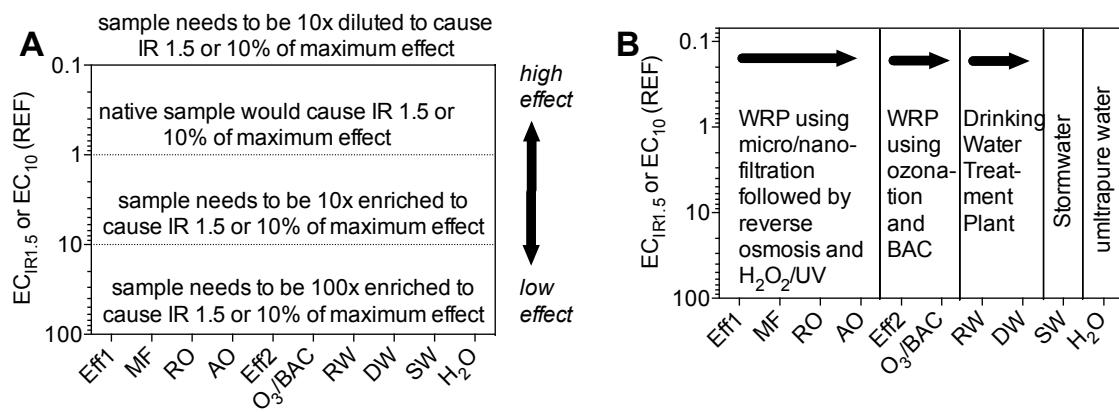


Figure S4. Presentation of bioassay results (EC = effect concentration, IR = induction ratio, REF = relative enrichment factor, for sample abbreviation see Table S1).

S6-A. Induction of xenobiotic metabolism pathways

The three bioassays for PXR and the six bioassays for AhR all showed positive responses in less treated samples and negative responses in recycled water and the blank (Figure S5). For the PXR (Figure S5-A), the FACTORIAL assays were most responsive with an $EC_{IR1.5}$ below a REF of 1, *i.e.*, this effect would be observable in the ambient water sample. The HG5LN-hPXR (Seimandi *et al.*, 2005; Lemaire *et al.*, 2006) reporter gene assay has been applied widely in water quality monitoring including for testing of wastewater, surface water and reclaimed water (Mahjoub *et al.*, 2009; Creusot *et al.*, 2010; Kinani *et al.*, 2010; Mnif *et al.*, 2010; Mnif *et al.*, 2011). This assay was responsive to the same samples as the two PXR-FACTORIAL endpoints.

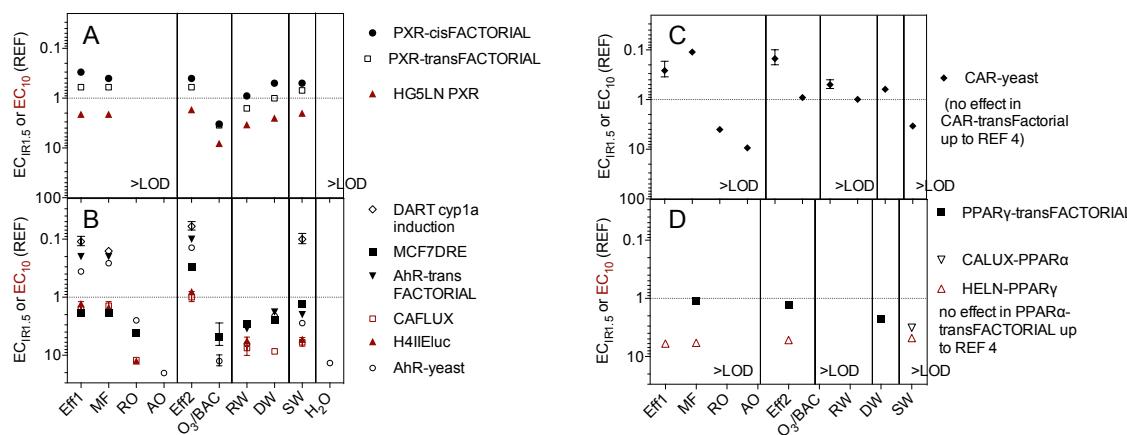


Figure S5. Results of bioassays indicative of induction of xenobiotic metabolism pathways. The red symbols are EC_{10} values the black symbols are $EC_{IR1.5}$ values.

The six AhR-related bioassays all showed consistent effect-patterns for the different samples and similar range of sensitivity although quantitative comparison is not possible because two bioassays yielded EC_{10} and four $EC_{IR1.5}$ values (Figure S5-B). The AhR-yeast had a remarkable dynamic range of EC over more than two orders of magnitude. A very responsive endpoint for induction of AhR was the CYP1a transcription in zebrafish embryos measured by RT-PCR. Only four samples (Eff1, MF, Eff2 and SW) were tested with RT-PCR but they all responded at low REF. The resulting $EC_{IR1.5}$ ranged from 0.06 to 0.16 REF (*i.e.*, responsive even in diluted samples).

CAR was tested in two bioassays but only the CAR-yeast gave a response already at low REFs (Figure S5-C). For PPAR, only two of seven bioassays gave signals in the four most polluted samples (Figure S5-D). Both active PPAR assays related to PPAR γ (PPAR γ -transFACTORIAL and HELN-PPAR γ). We detected no PPAR antagonism but samples were only tested up to a REF of 2.

S6-B. Endocrine disruption

All bioassays for estrogenicity were active in four to five samples. Blanks, RO and AO did not induce any estrogenic *in vitro* effects, and only two bioassays gave very low effects for samples DW and O₃/BAC (Table S5, Figure S6). The results are consistent with a previous interlaboratory comparison study of five different bioassays for estrogenicity (GWRC, 2008), where the effects observed in the ER-CALUX, the YES, the E-SCREEN and the T47KBluc (not assessed here) were highly correlated. Leusch *et*

al. (2010) also showed that the yeast-based assays have higher detection limits and therefore are not suitable for highly treated water but for the present applications they could still be used as the REF could be increased without cytotoxicity occurring and EC values obtained for MF and Eff2 were in the same range as for other endpoints. Again it must be cautioned that a quantitative comparison between EC_{10} and $EC_{IR1.5}$ is not possible but as Figure S6 demonstrates they are in the same range of relative sensitivity.

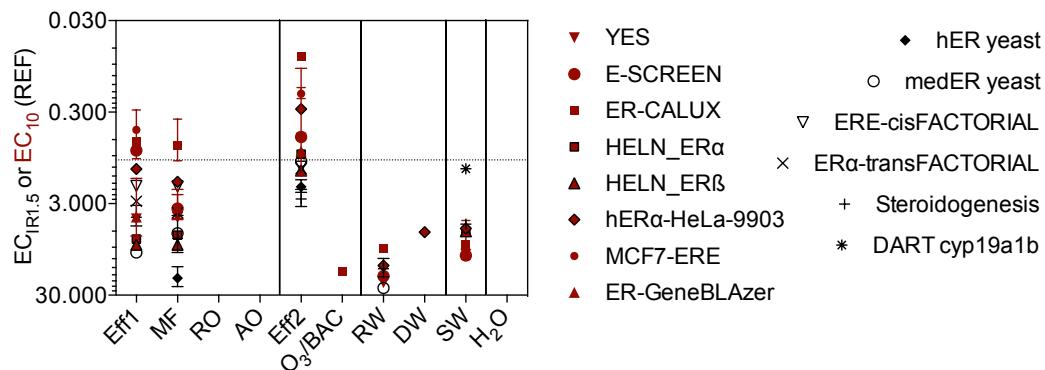


Figure S6. Results of bioassays indicative of estrogenicity. The red symbols are EC_{10} values the black symbols are $EC_{IR1.5}$ values.

The anti-ER test quantifies how the sample influences the effect of a constant concentration of estradiol that is typically spiked at concentrations that would elicit 50 to 80% of maximum effect. If the effect of the constant concentration of estradiol was suppressed and the sample was not cytotoxic, the sample can be considered to act as an anti-estrogen. No anti-estrogenic activity could be detected with the anti-ER-CALUX in any of the samples. It should be noted that agonistic and antagonistic activity may be occurring simultaneously in these assays, masking any such individual response of these activities.

The steroidogenesis pathway represents the biosynthesis route of steroid hormones from cholesterol via a battery of oxidative enzymes (Zhang *et al.*, 2005). The H295R steroidogenesis assay is an OECD validated bioassay used to evaluate the endocrine disrupting effects by chemicals via non-receptor mediated mechanisms. As this pathway directly affects the hormone system function, we have classified it with the receptor-mediated hormonal effects rather than with xenobiotic metabolism. The steroidogenesis assay showed increased concentrations of estrone and estradiol, which could be associated with a decreased estradiol metabolism, as well as increase in

progesterone and 17 α -hydroxyprogesterone, which is most likely due to an inhibitory effect on CYP21A. A similar effect was observed when oil sand product water was assessed with the steroidogenesis assay: the raw water increased the estradiol levels and the effect disappeared after ozonation (He *et al.*, 2010). In the same way that the effect of sample Eff2 went below the limit of detection when it was ozonated. The effect pattern of the sample Eff1 was similar to what has been observed when dosing with bisphenol A (Zhang *et al.*, 2011).

In relation to activation of the glucocorticoid receptor, GR-CALUX was the most responsive of these assays, followed by GR-transFACTORIAL (Figure S7). Their EC values were roughly ten times lower than the GR-Switchgear and GR-MDA-kb2 assays. The GR-GeneBLAzer was positive in MF, Eff2 and SW but the potency did not correlate well with the other assays.

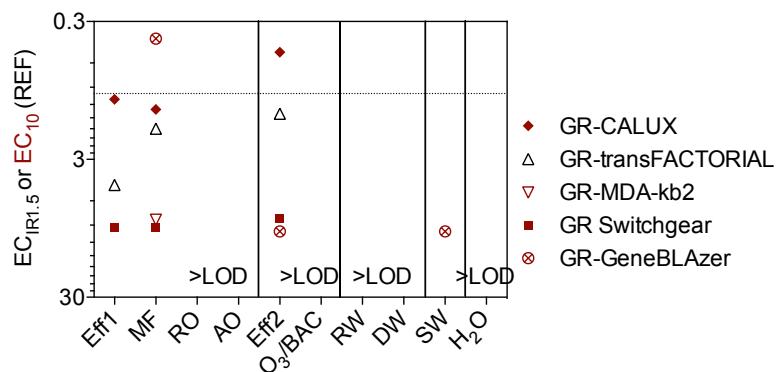


Figure S7. Results of bioassays indicative of glucocorticoid receptor (GR) activation. The red symbols are EC₁₀ values the black symbols are EC_{1R1.5} values.

No assay indicative of modulation of the thyroid hormone system showed response to any of the water samples, even at high REF. The T-Screen is a cell proliferation assay where the cells only proliferate in the presence of thyroid hormones (Gutleb *et al.*, 2005). This assay has been mainly applied for chemical dose-response tests and only few of the environmentally relevant chemicals showed activity. Accordingly it was not surprising that no effects were detected in the water samples. Many of the thyroid-active chemicals appear to need metabolic activation, and thus a combination with a system for metabolic activation may be beneficial (Taxvig *et al.*, 2011). Inoue *et al.* (2009) observed activity lower than 10% of maximum effect at a REF of 100 with a two-hybrid yeast assay in surface waters; a greater response (>10%) was observed for a

WWTP influent but these effects disappeared in the effluent (Inoue *et al.*, 2011). In a different yeast-based assay, Li *et al.* (2011) did not observe any TR agonistic effect in water samples, and attributed anti-TR activity to phthalates (Li *et al.*, 2010).

The P19/A15 cell line was developed by transfecting an embryonic mouse carcinoma cell line with a plasmid carrying the retinoic acid response element (Novak *et al.*, 2007). This cell line has not been tested with water samples prior to this study and did not show any effects with water samples but the water samples enhanced the effect of constant concentrations of 9-cis retinoic acid slightly (data not shown). This effect could be caused by mixture effect or by the organic micropollutants acting as solubilizer for the very lipophilic RA.

S6-C. Reactive toxicity

Three samples were active in the micronucleus assay, Eff2, RW and DW (Figure S8). This is a different profile as compared to the receptor-mediated modes of action, where the DW typically did not show any response and the activity in the DW sample is presumably due to disinfection by-products formed during chlorination.

Both the SOS chromotest, based on induction of SOS repair in *Escherichia coli* (Quillardet *et al.*, 1982), and the umuC assays with *Salmonella typhimurium* (Oda *et al.*, 1985) are reporter gene assays, while the Ames test uses histidine-deficient *S. typhimurium* that can only grow if a reverse mutation occurs. The umuC and Ames tests utilize different strains of *S. typhimurium*, so the seven genotoxicity assays in Table 1 (main article) in fact only represents four different assay types. All umuC assays showed activity only at high REFs around 20 (Figure S8). The SOS chromotest gave very similar responses as the umuC. The Ames assay responded generally at lower REFs but suffered from high variability between the different bacterial strains (Figure S8). One problem was that several samples (*e.g.*, RO, AO, SW and H₂O) showed detectable yet inconsistent activity in the Ames assay, which was not apparent in the other genotoxicity assays. The dynamic range of these genotoxicity assays was relatively small and effects were only observed at relatively high enrichments (REF up to 20).

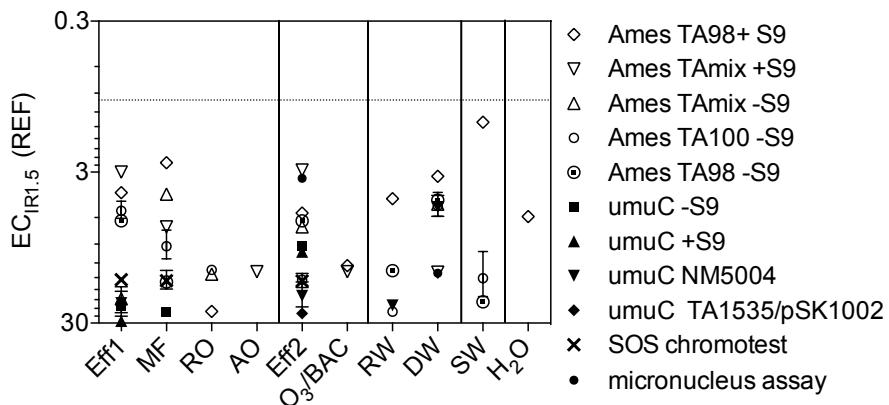


Figure S8. Results of bioassays indicative of reactive modes of action.

S6-D. Adaptive stress response

These bioassays are discussed in more detail in the main manuscript.

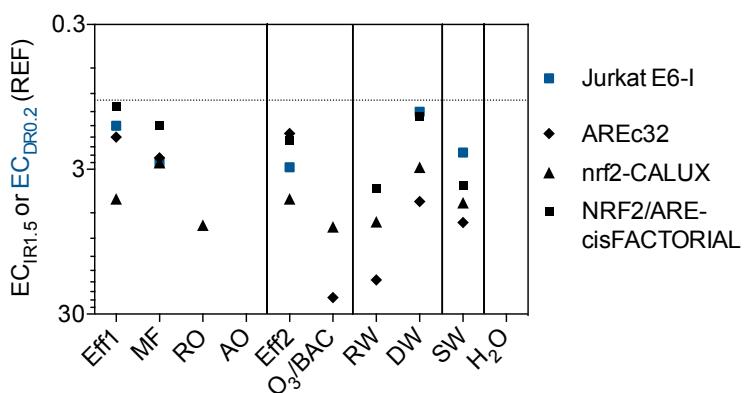


Figure S9. Results of bioassays indicative of adaptive stress response pathways.

Table S5. Summary of all EC values measured in all water samples (see Table S1 for definition of sample abbreviations). The errors represent the propagated standard error of the concentration-effect curve if only one laboratory performed the assay, and standard deviation of the mean results from different laboratories if the bioassay was performed by several laboratories (see also Section S4 for details on treatment of bioassays that were measured by multiple laboratories).

#	Class of MOA ^a	Lab(s)	Bioassay	EC (REF)	Eff1	MF	RO	AO	Eff2	O ₃ /BAC	RW	DW	SW	H ₂ O
1	XM	ATG	PXR-cisFACTORIAL	EC _{IR1.5}	0.3	0.4	>4	>4	0.4	3.3	0.9	0.5	0.5	>4
2	XM	ATG	PXR-transFACTORIAL	EC _{IR1.5}	0.6	0.6	>4	>4	0.6	3.5	1.6	1.0	0.7	>4
3	XM	IRCM	HG5LN PXR	EC ₁₀	2.1 ±0.01	2.1 ±0.01	>12	>12	1.7 ±0.01	8.1 ±0.08	3.4 ±0.01	2.5 ±0.02	2.0 ±0.01	>12
4	XM	ATG	CAR-transFACTORIAL	EC _{IR1.5}	>4	>4	>4	>4	>4	>4	>4	>4	>4	>4
5	XM	CAPIM	CAR-yeast	EC _{IR1.5}	0.3 ±0.1	0.1±0.01	4.0 ±0.2	9.4 ±0.6	0.2 ±0.1	0.9 ±0.1	0.5 ±0.1	1.0 ±0.1	0.6 ±0.1	3.4 ±0.2
6	XM	ATG	PPARα-transFACTORIAL	EC _{IR1.5}	>4	>4	>4	>4	>4	>4	>4	>4	>4	>4
7	XM	ATG	PPARγ-transFACTORIAL	EC _{IR1.5}	2.0	1.1	>4	>4	1.3	>4	>4	2.3	>4	>4
8	XM	IRCM	HELN-PPARγ	EC ₁₀	>6	5.8 ±0.1	>12	>12	5.2 ±0.1	>12	>12	>12	4.8 ±0.1	>12
9	XM	BDS	CALUX-PPARα	EC ₁₀	>30	>30	>30	>30	>30	>30	>30	>30	3.2	>30
10	XM	BDS, CSIRO	CALUX-PPARγ	EC ₁₀	>30	>30	>30	>30	>30	>30	>30	>30	>30	>30
11	XM	HK	MCF7-PPAR	EC _{IR1.5}	n.t.	n.t.	>20	>20	n.t.	n.t.	17.0 ±11.3	n.t.	n.t.	>20
12	XM	GU	PPARγ GeneBLAzer	EC ₁₀	>2	>2	>2	>2	>2	>2	>2	>2	>2	>2
13	XM	GU	Anti-PPARγ GeneBLAzer	EC _{SR0.2}	>2	>2	>1	>1	>2	>1	>1	>1	>2	>1

#	Class of MOA ^a	Lab(s)	Bioassay	EC (REF)	Eff1	MF	RO	AO	Eff2	O ₃ /BAC	RW	DW	SW	H ₂ O
14	XM	CAPIM	AhR-yeast	EC _{IR1.5}	0.36 ±0.01	0.26 ±0.02	2.46 ±0.03	20.2 ±0.5	0.14 ±0.01	12.6 ±2.7	6.4 ±0.6 2.14 ±0.13	2.81 ±0.13	13.5 ±0.3	
15	XM	UQ, RECET OX	AhR-CAFLUX 24h	EC ₁₀	1.6 ±0.4	1.4 ±0.2	12.1 ±0.5	>30	1.0 ±0.2	>30	7.5 ±2.6	8.6	6.0 ±1.0	>30
16	XM	RECET OX	H4IIEluc	EC ₁₀	1.3 ±0.1	1.5 ±0.2	12.5 ±1.6	>14	0.8 ±0.1	>14	5.6 ±0.6	>14	5.3 ±0.7	>27
17	XM	HK	MCF7-DRE	EC _{IR1.5}	1.9 ±0.3	1.9 ±0.3	4.1 ±0.5	>20	0.3 ±0.1	4.8 ±2.0	2.9 ±0.3	2.5 ±0.3	1.3 ±0.2	>18
18	XM	ATG	AhR- cisFACTORIAL	EC _{IR1.5}	0.2	0.2	>4	>4	0.1	>4	3.5	1.8	2.0	>4
19	XM	UFZ	DART cypla induction	EC _{IR1.5}	0.11 ±0.02	0.16 ±0.02	-	-	0.06 ±0.01	-	-	-	0.10 ±0.02	-
20	Specific MOA	UQ, Swiss	Algae photosynthesis inhibition	EC ₁₀	2.2 ±0.6	2.6 ±0.3	>20	>20	6.3 ±3.8	>20	>20	>20	6.5	>20
21	Specific MOA	UQ	Acetylcholin- esterase inhibition	EC ₁₀	1.9±0.1	>2	>2	>2	3.2 ±0.2	>2	>2	>2	>2	>2
22	Specific MOA: ER	GU, CSIRO, BDS, IWW	ER-CALUX	EC ₁₀ ^b	0.63 ±0.34	0.69 ±0.33	>25	>25	0.07 ±0.06	16.6	9.3	>25	8.5 ±3.9	>25
23	Specific MOA: ER	UQ	E-SCREEN	EC ₁₀	0.79 ±0.02	3.39 ±0.02	>30	>30	0.56 ±0.02	>30	18.6 ±0.4	>30	11.0 ±0.3	>30
24	Specific MOA: ER	Swiss, CSIRO, UA	Yeast Estrogen Screen (YES)	EC ₁₀ ^c	4.40 ±0.05	4.26 ±0.05	>30	>30	0.58 ±0.01	>30	22.0 ±0.1	>30	10.7 ±0.1	>30
25	Specific MOA: ER	CAPIM	hER yeast	EC _{IR1.5}	>30	19.5 ±4.8	>30	>30	2.0 ±0.3	>30	>30	>30	>30	>30
26	Specific MOA: ER	CAPIM	medER yeast	EC _{IR1.5}	10.3 ±1.0	6.3 ±0.6	>30	>30	1.0 ±0.1	>30	25.0 ±3.7	>30	>30	>30

#	Class of MOA ^a	Lab(s)	Bioassay	EC (REF)	Eff1	MF	RO	AO	Eff2	O ₃ /BAC	RW	DW	SW	H ₂ O
27	Specific MOA: ER	IRCM	HELN-ER α	EC ₁₀	7.29 ±0.13	6.61 ±0.05	>12	>12	0.86 ±0.01	>12	>12	>12	>6	>12
28	Specific MOA: ER	IRCM	HELN-ER β	EC ₁₀	8.42 ±0.15	8.42 ±0.11	>12	>12	1.30 ±0.02	>12	>12	>12	>6	>12
29	Specific MOA: ER	ATG	ERE-cisFACTORIAL	EC _{IR1.5}	1.91	1.94	>4	>4	1.18	>4	>4	>4	>4	>4
30	Specific MOA: ER	RECET OX	hER α -HeLa-9903	EC ₁₀	1.26 ±0.01	1.73 ±0.02	>27	>27	0.28 ±0.01	>25	14.2 ±0.2	6.1 ±0.1	5.6 ±0.1	>27
31	Specific MOA: ER	HK	MCF7-ERE	EC ₁₀	0.47 ±0.01	>10	>20	>20	0.19 ±0.03	>20	>20	>20	>10	>20
32	Specific MOA: ER	ATG	ER α -transFACTORIAL	EC _{IR1.5}	2.82	3.77	>4	>4	1.24	>4	>4	>4	>4	>4
33	Specific MOA: ER	NJU	Steroidogenesis	EC _{IR1.5}	4.2 ±1.1	7.2 ±3.1	>20	>20	2.7 ±0.6	>20	15.5 ±3.5	>20	5.0 ±0.7	>20
34	Specific MOA: ER	UFZ	DART cyp19a1b (aromatase)	EC _{IR1.5}	>1.2	>2.2	-	-	>2.2	-	-	-	0.89 ±0.31	-
35	Specific MOA: ER	UF, USF, UCR, SCCWR P	ER-GeneBLAzer	EC ₁₀ ^b	4.2 ±2.6	3.3 ±1.2	>20	>20	0.8 ±0.7	>20	>20	>20	>20	>20
36	Specific MOA: ER	CSIRO, GU	Anti-ER-CALUX	EC _{SR0.2}	>8	>8	>15	>15	>8	>15	>15	>15	>8	>15
37	Specific MOA: BDS,	GU, BDS,	AR-CALUX	EC ₁₀	>30	>30	>30	>30	>30	>30	>30	>30	>30	>30

#	Class of MOA ^a	Lab(s)	Bioassay	EC (REF)	Eff1	MF	RO	AO	Eff2	O ₃ /BAC	RW	DW	SW	H ₂ O
	AR	CSIRO												
38	Specific MOA: AR	IRCM	HELN-AR	EC ₁₀	>6	>6	>12	>12	>6	>12	>12	>12	>6	>12
39	Specific MOA: AR	HK	MCF7-ARE	EC _{IR1.5}	>10	>10	>20	>20	>10	>20	>20	>20	>10	>20
40	Specific MOA: AR	UA, CSIRO	Yeast Androgen Screen (YAS)	EC ₁₀	>30	>30	>30	>30	>30	>30	>30	>30	>30	>30
41	Specific MOA: AR	UF, USF, UCR, SCCWR P	AR-GeneBLAzer	EC ₁₀	>10	>10	>20	>20	>10	>20	>20	>20	>10	>20
42	Specific MOA: AR	ATG	AR-transFACTORIAL	EC _{IR1.5}	>4	>4	>4	>4	>4	>4	>4	>4	>4	>4
43	Specific MOA: AR/GR	RECET OX	MDA-kb2	EC ₁₀	2.3 ±0.1	1.7 ±0.1	24.6 ±0.4	>27	1.2 ±0.1	>27	>27	>27	>13	>27
44	Specific MOA: AR/GR	RECET OX	Anti-MDA-kb2	EC _{SR0.2}	>13	>13	>30	>30	>13	>30	6.2 ±7.0	>30	8.1 ±11.2	>30
45	Specific MOA: AR	CSIRO, GU	Anti-AR-CALUX	EC _{SR0.2}	7.7 ±2.0	6.4 ±1.0	>15	>15	2.9 ±0.2	>15	11.0 ±0.7	>15	6.6 ±1.1	>15
46	Specific MOA: GR	GU, BDS, CSIRO	GR-CALUX	EC ₁₀ ^d	1.1 ±0.1	1.3 ±0.1	>30	>30	0.5 ±0.1	>30	>30	>30	>30	>30
47	Specific MOA: GR	UA	GR Switchgear	EC ₁₀	9.4 ±0.1	9.4 ±0.1	>20	>20	8.1 ±0.2	>20	>20	>20	>10	>20

#	Class of MOA ^a	Lab(s)	Bioassay	EC (REF)	Eff1	MF	RO	AO	Eff2	O ₃ /BAC	RW	DW	SW	H ₂ O
48	Specific MOA: GR	ATG	GR-transFACTORIAL	EC _{IR1.5}	4.6	1.8	>4	>4	1.4	>4	>4	>4	>4	>4
49	Specific MOA: GR	RECETOX	GR-MDA-kb2 (AR suppressed with Flutamide)	EC ₁₀	>14	8.2 ±0.1	>27	-	>14	-	-	-	-	-
50	Specific MOA: GR	GU, UF, USF, UCR, SCCWRP	GR-GeneBLAzer	EC ₁₀ ^b	>10	12.6 ±3.0	>20	>20	>10	>20	>20	>20	>10	>20
51	Specific MOA: GR	GU	Anti-GR-GeneBLAzer	EC _{SR0.2}	>2	>2	>1	>1	>2	>1	>1	>1	>2	>1
52	Specific MOA: GR	GU	Anti-GR-CALUX	EC _{SR0.2}	>2	>2	>1	>1	>2	>1	>1	>1	>2	>1
53	Specific MOA: PR	UF, USF, UCR, SCCWRP	PR-GeneBLAzer	EC ₁₀	>10	>20	>20	>20	>20	>20	>20	>20	>10	>20
54	Specific MOA: PR	GU, BDS, CSIRO	PR-CALUX	EC ₁₀	>30	>30	>30	>30	>30	>30	>30	>30	>30	>30
55	Specific MOA: PR	GU	Anti-PR-CALUX	EC _{SR0.2}	>2	>2	>1	>1	>2	>1	>1	>1	>2	>1
56	Specific MOA: PR	NJU	Steroidogenesis, induction of progesterone	EC _{IR1.5}	5.4 ±2.6	4.2 ±1.6	>10	>10	2.0 ±0.3	>10	>10	>10	>10	>10
57	Specific MOA: PR	NJU	Steroidogenesis, induction of 17α OH-progesterone	EC _{IR1.5}	2.7 ±0.3	6.3 ±0.8	>10	>10	1.1 ±0.1	>10	>10	>10	>10	>10

#	Class of MOA ^a	Lab(s)	Bioassay	EC (REF)	Eff1	MF	RO	AO	Eff2	O ₃ /BAC	RW	DW	SW	H ₂ O
58	Specific MOA: GU TR	BDS, GU	TR-CALUX	EC ₁₀	>30	>30	>30	>30	>30	>30	>30	>30	>30	>30
59	Specific MOA: TR	UQ	T-Screen	EC ₁₀	>30	>30	>30	>30	>30	>30	>30	>30	>30	>30
60	Specific MOA: TR	ATG	THRα1-transFACTORIAL	EC _{IR1.5}	>4	>4	>4	>4	>4	>4	>4	>4	>4	>4
61	Specific MOA: TR	IRCM	HELN-TR	EC ₁₀	>6	>6	>12	>12	>6	>12	>12	>12	>6	>12
62	Repro	HK	MCF7-RARE	EC _{IR1.5}	>10	>10	>20	>20	>20	>20	>20	>20	>20	>20
63	Repro	UQ	P19/A15	EC _{IR1.5}	>30	>30	>30	>30	25.3 ±4.7	>30	>30	>30	>30	>30
64	Repro	ATG	RORβ-transFACTORIAL	EC _{IR1.5}	>4	>4	>4	>4	>4	>4	>4	>4	>4	>4
65	Repro	CAPIM	hRAR-Yeast Assay	EC _{IR1.5}	4.6 ±1.0	3.6 ±0.3	>30	>30	8.0 ±1.2	>30	>30	>30	>30	>30
66	Reactive MOA	RCEES, UQ	umuC TA1535/pSK1002	EC _{IR1.5} ^b	22.6 ±0.7	25.3 ±1.6	>30	>30	17.5 ±11.8	>30	>30	>30	>30	>30
67	Reactive MOA	UQ	umuC TA1535/pSK1002 +S9	EC _{IR1.5}	29.0 ±1.4	31.3 ±2.2	>30	>30	10.1 ±0.2	>30	>30	>30	>30	>30
68	Reactive MOA	RCEES	umuC NM5004	EC _{IR1.5}	22.1 ±3.6	>30	>30	>30	19.8 ±3.7	>30	23±3	5.1 ±0.8	>30	>30
69	Reactive MOA	RECET OX	SOS chromotest	EC _{IR1.5}	15.5 ±1.3	15.7 ±2.2	>27	>27	15.8 ±0.1.7	>27	>27	>27	>13	>27
70	Reactive MOA	UA, IWW	Ames TA98 -S9	EC _{RR1.5} ^c	6.3	16.0	>30	>30	6.3	>30	13.5	4.6	21.7	>30
71	Reactive MOA	UA, IWW	Ames TA98+ S9	EC _{RR1.5} ^c	4.1	2.6	25.1	>30	5.6	12.5	4.5	3.2	1.4	5.9
72	Reactive MOA	UA, IWW	Ames TAmix -S9	EC _{RR1.5}	20.7	4.2	14.2	>30	6.9	>30	>30	4.9	>30	>30

#	Class of MOA ^a	Lab(s)	Bioassay	EC (REF)	Eff1	MF	RO	AO	Eff2	O ₃ /BAC	RW	DW	SW	H ₂ O
73	Reactive MOA	UA, IWW	Ames TAmix +S9	EC _{RR1.5} ^e	3.0	6.9	>30	13.7	2.9	13.7	>30	13.8	>30	>30
74	Reactive MOA	UQ	Ames TA100 -S9	EC _{RR1.5} ^e	5.4 ±0.7	9.3 ±2.0	13.3 ±0.01	>30	15.5 ±23.2	>30	25.2 ±0.01	5.0 ±0.9	15.1 ±5.0	>30
75	Reactive MOA	AWQC	Micronucleus assay	EC _{RR1.5} ^e	>20	>20	>30	>30	2.6	>30	20.9	9.0	>20	>30
76	Reactive MOA	CSIRO	ROS formation in RTG2 cells	EC _{RR1.5}	9.0 ±0.2	9.4 ±0.2	>30	>30	4.9 ±0.1	>30	>30	>30	11.9 ±0.7	>30
77	Reactive MOA	UQ	Protein damage <i>E.coli</i> GSH+/-	EC _{IR1.5}	>30	>30	>30	>30	>30	>30	>30	>30	>30	>30
78	ASR	ATG	HSE-cisFACTORIAL	EC _{IR1.5}	>4	>4	>4	>4	>4	>4	>4	>4	>4	>4
79	ASR	UFZ	hspb11 induction in DART after 120h	EC _{IR1.5}	>1.2	>4.2	-	-	>2.2	-	-	-	>2.2	-
80	ASR	ATG	HIF-1a-cisFACTORIAL	EC _{IR1.5}	>4	>4	>4	>4	>4	>4	>4	>4	>4	>4
81	ASR	UA	Hypoxia-Switchgear	EC _{IR1.5}	>10	>10	>20	>20	>10	>20	>20	>20	>10	>20
82	ASR	ATG	NFκB-cisFACTORIAL	EC _{IR1.5}	>4	>4	>4	>4	>4	>4	>4	>4	>4	>4
83	ASR	UQ	NFκB-Geneblazer	EC _{IR1.5}	>20	13.8 ±1.3	>20	>20	17.1 ±2.2	>20	>20	>20	>20	>20
84	ASR	BDS	NFκB-CALUX	EC _{IR1.5}	>30	>30	>30	>30	>30	>30	>30	>30	>30	>30
85	ASR	GU	Jurkat E6.1 IκB	EC _{CD0.2}	0.8	1.4	>2	>2	1.6	>2	>2	0.7	1.3	>2
86	ASR	UQ	AREc32	EC _{IR1.5} ^f	1.8 ±0.1	2.5 ±0.1	>30	>30	1.7 ±0.1	23.1 ±1.3	17.4 ±0.5	5.0 ±0.2	7.0 ±0.3	>30
87	ASR	UA	Nrf2-keap	EC _{IR1.5}	>10	>10	>20	>20	>10	>20	>20	>20	>10	>20
88	ASR	ATG	Nrf2/ARE-cisFACTORIAL	EC _{IR1.5}	1.1 ±0.1	1.5 ±0.1	>4	>4	1.9 ±0.1	>4	4.1 ±0.7	1.3 ±0.1	3.9 ±1.1	>4
89	ASR	BDS	Nrf2-CALUX	EC _{IR1.5}	4.8 ±0.1	2.7 ±0.1	7.3 ±0.4	>30	4.8 ±0.4	7.5 ±0.5	6.9 ±0.4	2.9 ±0.2	5.1 ±0.3	>30
90	ASR	ATG	p53-	EC _{IR1.5}	>4	>4	>4	>4	>4	>4	>4	>4	>4	>4

#	Class of MOA ^a	Lab(s)	Bioassay	EC (REF)	Eff1	MF	RO	AO	Eff2	O ₃ /BAC	RW	DW	SW	H ₂ O
cisFACTORIAL														
91	ASR	BDS, IWW	p53-CALUX	EC _{IR1.5}	>30	>30	>30	>30	>30	>30	>30	>30	>30	>30
92	ASR	BDS	p53-CALUX +S9	EC _{IR1.5}	>30	>30	>30	>30	>30	>30	>30	>30	>30	>30
93	ASR	UF	p53-GeneBLAzer	EC _{IR1.5}	>10	>10	>20	>20	>10	>20	>20	>20	>10	>20
94	CT	UQ	AREc32 cell viability	EC ₁₀	15.5	>30	>30	>30	28.8	>30	>30	>30	>30	>30
95	CT	GU	Caco 2 NRU	EC ₁₀	7.20	10.80	>20	>20	3.90	>20	>20	>20	>20	>10
96	CT	CSIRO	RTG2 MTT	EC ₁₀	>30	>30	>30	>30	17.6 ±0.2	>30	>30	>30	>30	>30
97	CT	UFZ	DART 48h lethality	EC ₁₀	>10	>10	>10	>30	5.0	>30	>30	>30	6.5	>30
98	CT	UFZ	DART 120h sublethal	EC ₁₀	1.2	4.9	>10	>30	2.3	>30	>30	>30	2.2	>30
99	CT	GU	SK-N-SH cytotoxicity	EC ₁₀	>1	>1	>2	>2	>1	>2	>2	>2	>2	>2
100	CT	GU	THP1 cytokine	EC ₁₀	>1	>1	>2	>2	>1	>2	>2	>2	>2	>2
101	CT	UQ	algae growth inhibition	EC ₁₀	5.4 ±0.6	4.5 ±0.2	17.3	>20	7.7 ±1.0	>20	>20	14.1	15.4	>20
102	CT	UQ, Swiss	<i>Vibrio fischeri</i> (Microtox)	EC ₁₀ ^{b,g}	1.3 ±2.1	0.7 ±0.6	2.7 ±4.5	10.7 ±19.1	0.4 ±0.3	1.4 ±1.6	1.8 ±1.8	0.7 ±0.7	1.7 ±2.1	14 ±22
103	CT	RCEES	<i>Photobacterium phosphoreum</i> T3	EC ₁₀	0.4	0.5	2.7	6.6	0.2	0.5	0.5	0.2	0.3	1.7

refers to the number in the heatmap. ^aMOA= mode of action, XM = xenobiotic metabolism, Repro = reproductive and developmental effects, ASR = adaptive stress response, CT = cytotoxicity. ^bBioassays that were performed by several laboratories and the error denotes the standard deviation of the mean of different laboratories' results; ^conly CSIRO results; ^donly BDS; ^eonly UA, ^fprevious published (Escher et al., 2013). ^gPreviously published (Tang et al., 2013).

Section S7. Monitoring treatment efficacy

In this section we discuss the bioassays in the light of their suitability to serve as process monitoring tools. Of course one cannot say *a priori* that a bioassay is “good” or “sensitive” if it still measures an effect in treated waters. Effects can disappear if all chemicals that are responsive in this endpoint are well removed in the particular treatment process. The dynamic range between the effect of the product water and the blank is decisive for the suitability of an assay for assessing treatment efficacy (provided that reproducibility, repeatability and sensitivity have been already established with reference chemicals).

S7-A. Advanced water treatment plant using reverse osmosis

The investigated water reclamation plant (WRP) uses microfiltration followed by reverse osmosis and finished with advanced oxidation. The micropollutant flow in this plant has been characterized in detail in previous work by both chemical and bioanalytical tools (Escher *et al.*, 2011; Macova *et al.*, 2011). In the present study, we selected only four sampling points before and after critical treatment steps, the WRP inflow (WWTP effluent, sample Eff1), after microfiltration (sample MF), after reverse osmosis (sample RO) and after advanced oxidation combining hydrogen peroxide and UV irradiation (sample AO).

Effects were detected in Eff1 in 51 of 101 bioassays (Figure S10, red symbols and line, excluding the bacterial cytotoxicity assays). Subsequent treatment steps greatly reduced the effect burden caused by micropollutants. After MF (Figure, S10, blue symbols and line) 52 bioassays tested positive (not exactly the same ones), but RO decreased the number of positives sharply to 11. After AO, only three bioassays tested positive and these also tested positive in the ultrapure water blank.

In 12 bioassays, the effect after MF increased by more than 20%, *i.e.*, more than the variability of the assay response. In 20 bioassays the effect remained constant ($\pm 20\%$) and in 15 bioassays the effect decreased substantially already in the MF step. The five times increase in effect in the Ames TA_{mix}-S9 is presumably an artifact of the large variability of the results of this endpoint. To avoid biofouling, the MF membranes are

chloraminated, which causes the formation of disinfection by-products that can cause effects in some of the bioassays for reactive modes of action (Neale *et al.*, 2012).

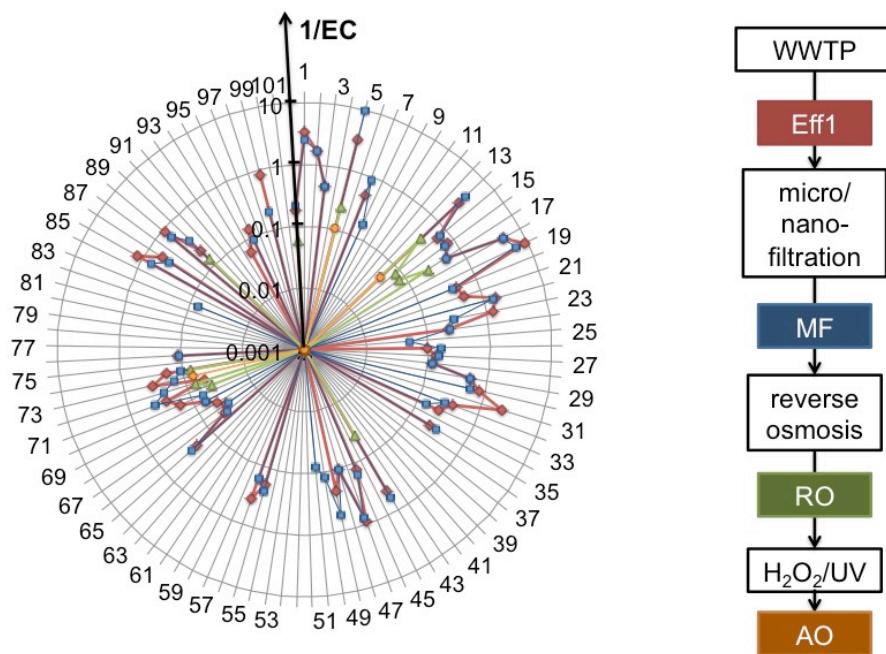


Figure S10 Bioanalytical fingerprint of the water from the WRP process using reverse osmosis. The red diamonds represent the Eff1, blue squares are MF, green triangles RO and yellow circles AO. The numbers refer to the bioassay numbers in Table 1 or Table S5.

The effects were greatly reduced after reverse osmosis (Figure S10). In 31 of 52 bioassays the effect disappeared to below detection limit and in most bioassays the effective response was reduced by an order of magnitude. There was no preferred type or group of effects removed. The bioassays with high variability, *e.g.*, the Ames assay, seem unsuitable for reliable assessment. The bioassays that showed reduction of effect but are sufficiently sensitive to respond after RO, are best suited as indicator bioassays. These include algae growth inhibition, the xenobiotic metabolism indicators, AhR-CAFLUX, H4IIIEluc and MCF7-DRE. Of specific receptor-mediated modes of action, MDA-kb2 and hER α -HeLa-9903 were able to show the dynamics of treatment. In the group of adaptive stress responses, the AREc32 and Nrf2-CALUX showed a distinct reduction

pattern but were still above the LOD in RO water and are thus suitable as sensitive screening tool for process control.

S7-B. Water reclamation plant using ozonation and biologically activated carbon filtration

The second investigated WRP produces recycled water from secondary treated wastewater plant effluent using ozonation and activated carbon filtration (van Leeuwen *et al.*, 2003). The WRP has a capacity of 10 ML d⁻¹ and provides water to industry for non-potable uses. Whilst the plant provides water for non-potable applications, it has been designed to meet drinking water standards. The treatment process incorporates biological denitrification, pre-ozonation, coagulation/ flocculation/dissolved air flotation-sand filtration (DAFF), ozonation, biological activated carbon treatment and ozone disinfection.

The removal efficiency of micropollutants has been analysed in detail in a series of studies that combined chemical analysis with bioanalytical tools (Macova *et al.*, 2010; Reungoat *et al.*, 2010; Reungoat *et al.*, 2011; Reungoat *et al.*, 2012a; Reungoat *et al.*, 2012b). In the present study, two points in the treatment train were targeted: the WRP influent (secondary treated effluent, Eff2) and the product water after ozonation and biological activated carbon treatment (O₃/BAC).

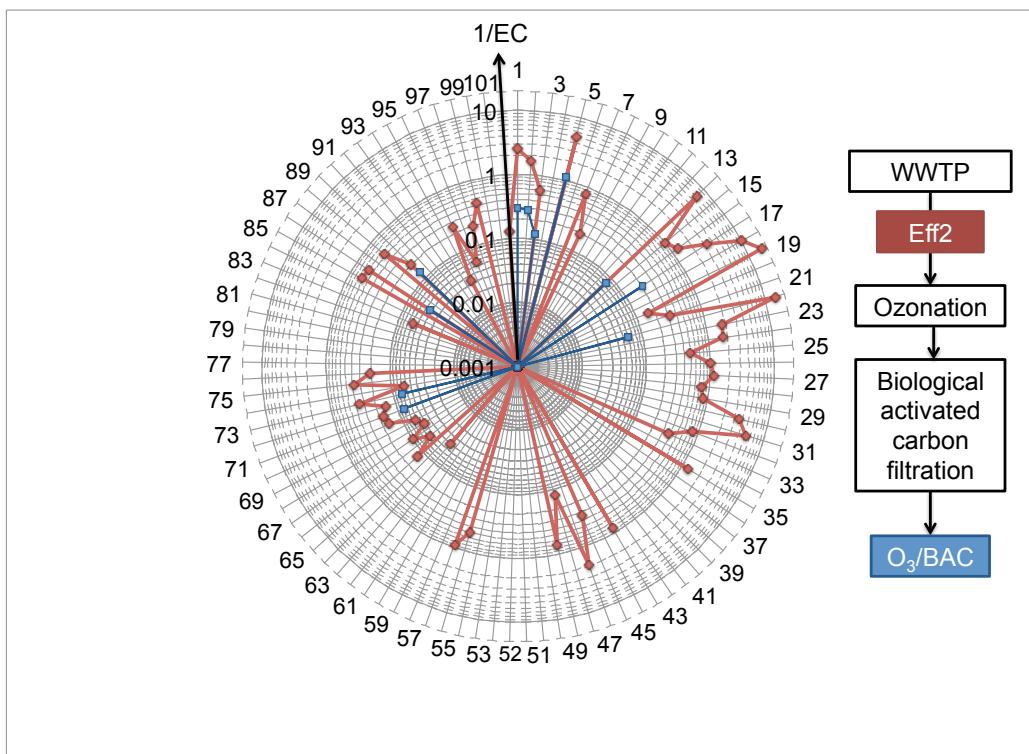


Figure S11. Bioanalytical fingerprint of the water treated with ozonation and biological activated carbon.

As is shown in Figure S11, 58 bioassays were above detection limit in the Eff2, similar to what was found with Eff1. The treatment reduced the number of responses to 11 and the effects in these positive bioassays were also greatly reduced (Figure S12). As was the case for the other facility, the bioassays that still showed an effect in the treated water are suitable as indicator bioassays to benchmark treatment efficiency.

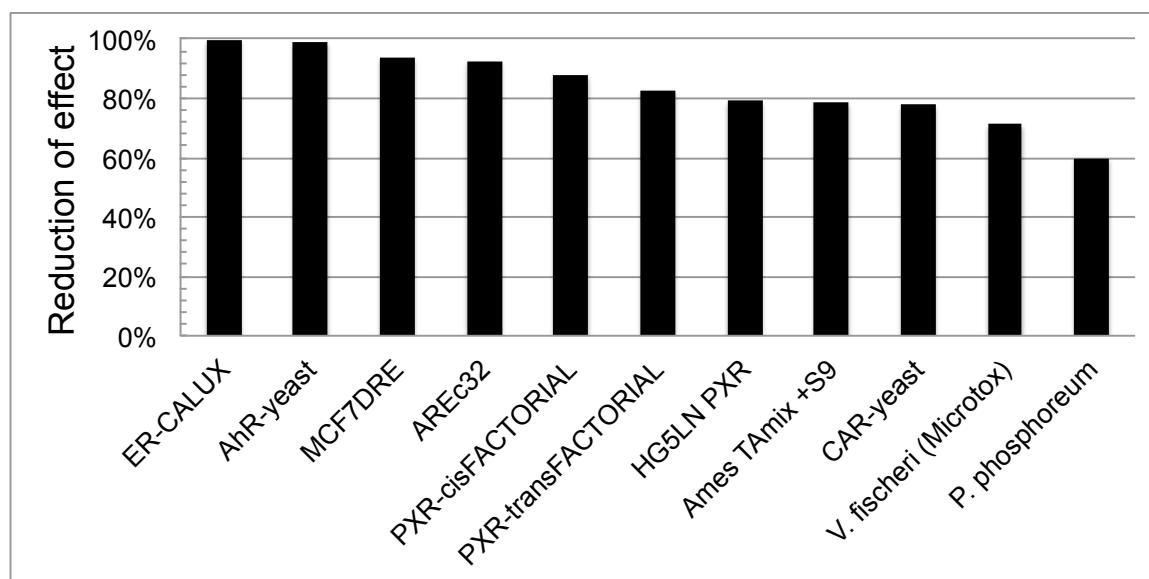


Figure S12. Percent treatment efficiency in the bioassays that did not fall below limit of detection (LOD) after treatment.

S7-C. Drinking water treatment plant

For comparison, we assessed treatment in a drinking water (DW) treatment plant. This plant has also been evaluated previously and applies coagulation and filtration followed by chlorination and finishing with chloramination (Macova *et al.*, 2011; Neale *et al.*, 2012). Here, only the feed water and the final drinking water were evaluated. The feed water is drawn from a river (RW) and the levels of micropollutants and effects (Figure S13) were low (Tang *et al.*, 2013). In RW and DW, 25 and 22 of 101 bioassays were positive but only 17 positive bioassays in DW were identical to those positive in the RW, for the remaining positives different biological endpoints were triggered in RW and DW.

The effects in the E-SCREEN, the AhR-CAFLUX and the MCF7-DRE remained the same or were reduced indicating that chlorination degraded or did not change existing micropollutants but did not produce specifically acting compounds. However drinking water treatment with chlorination and chloramination increased the non-specific and reactive toxicity (Figure S13) due to the formation of disinfection by-products, which is consistent with previous findings and chemical analysis of formed disinfection by-products (Neale *et al.*, 2012).

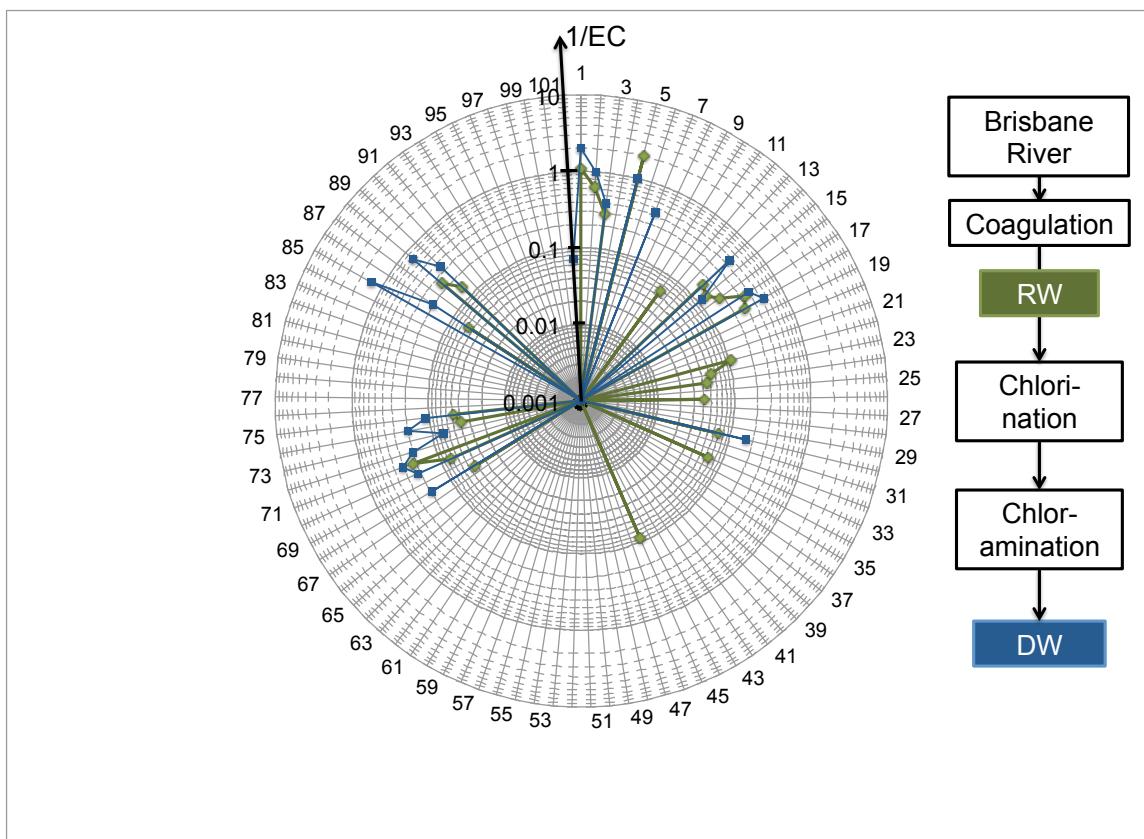


Figure S13. Bioanalytical fingerprint of the drinking water treatment.

Of the bioassays that increased in the toxicity, only one represents a specific mode of action, the hER α -HeLa-9903. As expected, the majority of these positive assays targeted xenobiotic metabolism, reactive modes of action and adaptive stress responses. Increase was most pronounced in the reactive modes of action (Ames TA98+ and -S9, Ames TA100 -S9, umuC NM5004 and micronucleus assay). There was a detectable but small increase by up to a factor of 2 for the bioassays indicative of xenobiotic metabolism, with a preference for the PXR (HG5LN PXR, PXR-transFACTORYL, PXR-cisFACTORYL). The responses in all three bioassays for the oxidative stress response (Nrf2-CALUX, Nrf2/ARE-cisFACTORYL and AREc32) increased by a factor of 2.4 to 4.2.

References

- Baqi, A., Meiller, T.F., Chon, J.J., Turng, B.F. and Falkler, W.A. (1998). Granulocyte-macrophage colony-stimulating factor amplification of interleukin-1 beta and tumor necrosis factor alpha production in THP-1 human monocytic cells

- stimulated with lipopolysaccharide of oral microorganisms. *Clinical and Diagnostic Laboratory Immunology*, **5**(3): 341-347.
- Brasier, A.R. and Fortin, J.J. (2001). Nonisotopic Assays for Reporter Gene Activity *Curr Protoc Mol Biol*: Chapter 9:Unit9.7B. doi: 10.1002/0471142727.mb0471140907bs0471142729.
- Creusot, N., Kinani, S., Balaguer, P., Tapie, N., LeMenach, K., Maillot-Marechal, E., Porcher, J.M., Budzinski, H. and Ait-Aissa, S. (2010). Evaluation of an hPXR reporter gene assay for the detection of aquatic emerging pollutants: screening of chemicals and application to water samples. *Analytical and Bioanalytical Chemistry*, **396**(2): 569-583.
- Escher, B.I., Dutt, M., Maylin, E., Tang, J.Y.M., Toze, S., Wolf, C.R. and Lang, M. (2012). Water quality assessment using the AREc32 reporter gene assay indicative of the oxidative stress response pathway. *Journal of Environmental Monitoring*, **14**: 2877-2885
- Escher, B.I., Lawrence, M., Macova, M., Mueller, J.F., Poussade, Y., Robillot, C., Roux, A. and Gernjak, W. (2011). Evaluation of contaminant removal of reverse osmosis and advanced oxidation in full-scale operation by combining passive sampling with chemical analysis and bioanalytical tools. *Environmental Science & Technology*, **45**: 5387-5394.
- Escher, B.I., van Daele, C., Dutt, M., Tang, J.Y.M. and Altenburger, R. (2013). Most oxidative stress response in water samples comes from unknown chemicals: the need for effect-based water quality trigger values. *Environmental Science & Technology*, **47**: 7002-7011.
- Gutleb, A.C., Meerts, I.A.T.M., Bergsma, J.H., Schriks, M. and Murk, A.J. (2005). T-Screen as a tool to identify thyroid hormone receptor active compounds. *Environmental Toxicology and Pharmacology*, **19**(2): 231-238.
- GWRC (2008). Tools to detect estrogenicity in environmental waters. F. D. L. Leusch, Global Water Research Coalition (GWRC) and Water Environment Research Foundation, Alexandria, VA, USA, pp. 74.
- He, Y.H., Wiseman, S.B., Zhang, X.W., Hecker, M., Jones, P.D., El-Din, M.G., Martin, J.W. and Giesy, J.P. (2010). Ozonation attenuates the steroidogenic disruptive effects of sediment free oil sands process water in the H295R cell line. *Chemosphere*, **80**(5): 578-584.
- Inoue, D., Nakama, K., Matsui, H., Sei, K. and Ike, M. (2009). Detection of agonistic activities against five human nuclear receptors in river environments of Japan using a yeast two-hybrid assay. *Bulletin of Environmental Contamination and Toxicology*, **82**(4): 399-404.
- Inoue, D., Nakama, K., Sawada, K., Watanabe, T., Matsui, H., Sei, K., Nakanishi, T. and Ike, M. (2011). Screening of agonistic activities against four nuclear receptors in wastewater treatment plants in Japan using a yeast two-hybrid assay. *Journal of Environmental Sciences-China*, **23**(1): 125-132.
- Kinani, S., Bouchonnet, S., Creusot, N., Bourcier, S., Balaguer, P., Porcher, J.M. and Ait-Aissa, S. (2010). Bioanalytical characterisation of multiple endocrine- and dioxin-like activities in sediments from reference and impacted small rivers. *Environmental Pollution*, **158**(1): 74-83.

- Konsoula, R. and Barile, F.A. (2005). Correlation of *in vitro* cytotoxicity with paracellular permeability in Caco-2 cells. *Toxicology in Vitro*, **19**(5): 675-684.
- Lemaire, G., Mnif, W., Pascussi, J.M., Pillon, A., Rabenoelina, F., Fenet, H., Gomez, E., Casellas, C., Nicolas, J.C., Cavailles, V., Duchesne, M.J. and Balaguer, P. (2006). Identification of new human pregnane X receptor ligands among pesticides using a stable reporter cell system. *Toxicological Sciences*, **91**(2): 501-509.
- Leusch, F.D.L., De Jager, C., Levi, Y., Lim, R., Puijker, L., Sacher, F., Tremblay, L.A., Wilson, V.S. and Chapman, H.F. (2010). Comparison of five *in vitro* bioassays to measure estrogenic activity in environmental waters. *Environmental Science & Technology*, **44**(10): 3853-3860.
- Leusch, F.D.L., Khan, S.J., Gagnon, M.M., Quayle, P., Trinh, T., Coleman, H., Rawson, C., Chapman, H., Blair, P., Nice, H. and Reitsema, T. (2013). Assessment of wastewater and recycled water quality: A comparison of lines of evidence from *in vitro*, *in vivo* and chemical analyses. *Water Research* <http://dx.doi.org/10.1016/j.watres.2013.1010.1056>.
- Li, N., Ma, M., Rao, K. and Wang, Z. (2011). *In vitro* thyroid disrupting effects of organic extracts from WWTPs in Beijing. *Journal of Environmental Sciences*, **23**(4): 671-675.
- Li, N., Wang, D.H., Zhou, Y.Q., Ma, M., Li, J.A. and Wang, Z.J. (2010). Dibutyl phthalate contributes to the thyroid receptor antagonistic activity in drinking water processes. *Environmental Science & Technology*, **44**(17): 6863-6868.
- Macova, M., Escher, B.I., Reungoat, J., Carswell, S., Chue, K.L., Keller, J. and Mueller, J.F. (2010). Monitoring the biological activity of micropollutants during advanced wastewater treatment with ozonation and activated carbon filtration. *Water Research*, **44**(2): 477-492.
- Macova, M., Toze, S., Hodgers, L., Mueller, J.F., Bartkow, M.E. and Escher, B.I. (2011). Bioanalytical tools for the evaluation of organic micropollutants during sewage treatment, water recycling and drinking water generation. *Water Research*, **45**: 4238-4247.
- Mahjoub, O., Leclercq, M., Bachelot, M., Casellas, C., Escande, A., Balaguer, P., Bahri, A., Gomez, E. and Fenet, H. (2009). Estrogen, aryl hydrocarbon and pregnane X receptors activities in reclaimed water and irrigated soils in Oued Souhil area (Nabeul, Tunisia). *Desalination*, **246**(1-3): 425-434.
- Mnif, W., Dagnino, S., Escande, A., Pillon, A., Fenet, H., Gomez, E., Casellas, C., Duchesne, M.-J., Hernandez-Raquet, G., Cavaillès, V., Balaguer, P. and Bartegi, A. (2010). Biological analysis of endocrine-disrupting compounds in Tunisian sewage treatment plants. *Archives of Environmental Contamination and Toxicology*, **59**(1): 1-12.
- Mnif, W., Hassine, A.I.H., Zidi, I., Dagnino, S., Bouaziz, A., Fenet, H., Hamouda, Y.H., Balaguer, P. and Bartegi, A. (2011). Occurrence of estrogenic and pregnane X receptor specific activities in Tunisian sewage treatment plants using a panel of bioassays. *Immuno-Analyse & Biologie Specialisee*, **26**(1): 1-5.
- Mosmann, T. (1983). Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays. *Journal of Immunological Methods*, **65**(1-2): 55-63.

- Neale, P.A., Antony, A., Bartkow, M., Farre, M., Heitz, A., Kristiana, I., Tang, J.Y.M. and Escher, B.I. (2012). Bioanalytical assessment of the formation of disinfection by-products in a drinking water treatment plant. *Environmental Science & Technology*, **46**(18): 10317–10325.
- Novak, J., Benisek, M., Pachernik, J., Janosek, J., Sidlova, T., Kiviranta, H., Verta, M., Giesy, J.P., Blaha, L. and Hilscherova, K. (2007). Interference of contaminated sediment extracts and environmental pollutants with retinoid signaling. *Environmental Toxicology and Chemistry*, **26**(8): 1591-1599.
- Oda, Y., Nakamura, S.-I., Oki, I., Kato, T. and Shinagawa, H. (1985). Evaluation of the new system (umu-test) for the detection of environmental mutagens and carcinogens. *Mutation Research*, **147**(5): 219-229.
- Quillardet, P., Huisman, O., Dari, R. and Hofnung, M. (1982). SOS Chromotest, a direct assay of induction of an SOS function in *Escherichia coli* K-12 to measure genotoxicity. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences*, **79**(19): 5971-5975.
- Reungoat, J., Escher, B., Macova, M., Farre, M., Argaud, F.X., Rattier, M., Gernjak, W. and Keller, J. (2012a). Wastewater Reclamation using Ozonation combined with Biological Activated Carbon Filtration. Urban Water Security Research Alliance <http://www.urbanwateralliance.org.au/publications/UWSRA-tr69.pdf>, Brisbane, Australia.
- Reungoat, J., Escher, B.I., Macova, M., Argaud, F.X., Gernjak, W. and Keller, J. (2012b). Ozonation and biological activated carbon filtration of wastewater treatment plant effluents. *Water Research*, **46**(3): 863-872.
- Reungoat, J., Escher, B.I., Macova, M. and Keller, J. (2011). Biofiltration of wastewater treatment plant effluent: Effective removal of pharmaceuticals and personal care products and reduction of toxicity. *Water Research*, **45**(9): 2751-2762.
- Reungoat, J., Macova, M., Escher, B.I., Carswell, S., Mueller, J.F. and Keller, J. (2010). Removal of micropollutants and reduction of biological activity in a full scale reclamation plant using ozonation and activated carbon filtration. *Water Research*, **44**(2): 625-637.
- Richard, J. (2012). Toxicological in vitro evaluation of organic micropollutants before and after oxidative waste water treatment. Dept. of Chemistry. Essen, Germany, Universität Duisburg-Essen, <http://duepublico.uni-duisburg-essen.de/servlets/DocumentServlet?id=31124>.
- Seimandi, M., Lemaire, G., Pillon, A., Perrin, A., Caravan, I., Voegel, J.J., Vignon, F., Nicolas, J.C. and Balaguer, P. (2005). Differential responses of PPAR alpha, PPAR delta, and PPAR gamma reporter cell lines to selective PPAR synthetic ligands. *Analytical Biochemistry*, **344**(1): 8-15.
- Shiraishi, F., Shirashi, H., Nishikawa, J., Nishihara, T. and Morita, M. (2000). Development of a simple operational estrogenicity assay system using the yeast two-hybrid system. *Journal of Environmental Chemistry*, **10**(1): 57-64.
- Tang, J.Y.M., McCarty, S., Glenn, E., Neale, P.A., Warne, M.S. and Escher, B.I. (2013). Mixture effects of organic micropollutants present in water: towards the development of effect-based water quality trigger values for baseline toxicity. *Water Research* **47**: 3300-3314.

- Taxvig, C., Olesen, P.T. and Nellemann, C. (2011). Use of external metabolizing systems when testing for endocrine disruption in the T-screen assay. *Toxicology and Applied Pharmacology*, **250**(3): 263-269.
- van Leeuwen, J., Pipe-Martin, C. and Lehmann, R.M. (2003). Water Reclamation At South Caboolture, Queensland, Australia. *Ozone: Science & Engineering: The Journal of the International Ozone Association*, **25**: 107-120.
- Villeneuve, N.F., Du, Y., Wang, X.J., Sun, Z. and Zhang, D.D. (2008). High-throughput screening of chemopreventive compounds targeting Nrf2. 2008 3rd IEEE International Conference on Nano/Micro Engineered and Molecular Systems, Vols 1-3, New York, NY, USA, IEEE.
- Zhang, X.W., Chang, H., Wiseman, S., He, Y.H., Higley, E., Jones, P., Wong, C.K.C., Al-Khedhairy, A., Giesy, J.P. and Hecker, M. (2011). Bisphenol A Disrupts Steroidogenesis in Human H295R Cells. *Toxicological Sciences*, **121**(2): 320-327.
- Zhang, X.W., Yu, R.M.K., Jones, P.D., Lam, G.K.W., Newsted, J.L., Gracia, T., Hecker, M., Hilscherova, K., Sanderson, J.T., Wu, R.S.S. and Giesy, J.P. (2005). Quantitative RT-PCR methods for evaluating toxicant-induced effects on steroidogenesis using the H295R cell line. *Environmental Science & Technology*, **39**(8): 2777-2785.