2

Development of a bioanalytical test battery for water quality monitoring: Fingerprinting identified micropollutants and their contribution to effects in surface water

6

7   Peta A. Neale[a, b], Rolf Altenburger[c], Selim Aït-Aïssa[d], François Brion[d], Wibke Busch[c],

8   Gisela de Aragão Umbuzeiro[e], Michael S. Denison[f], David Du Pasquier[g], Klára Hilscherová[h],

9   Henner Hollert[i], Daniel A. Morales[e], Jiří Novák[h], Rita Schlichting[c], Thomas-Benjamin Seiler[i],

10   Helene Serra[d], Ying Shao[i], Andrew J. Tindall[g], Knut Erik Tollefsen[j], Timothy D. Williams[k],

11   Beate I. Escher[b, c, l]*

12

13   [a]Australian Rivers Institute, Griffith School of Environment, Griffith University, Southport QLD

14   4222, Australia

15   [b]The University of Queensland, National Research Centre for Environmental Toxicology (Entox),

16   Brisbane QLD 4108, Australia

17   [c]UFZ – Helmholtz Centre for Environmental Research, 04318 Leipzig, Germany

18   [d]Institut National de l'Environnement Industriel et des Risques INERIS, Unité d'Ecotoxicologie,

19   60550 Verneuil-en-Halatte, France

20   [e]Faculdade de Tecnologia, FT-UNICAMP, Universidade Estadual de Campinas, Limeira, SP

21   13484-332, Brazil

22   [f]Department of Environmental Toxicology, University of California, Davis, California 95616,

23   United States

24   [g]WatchFrog, Bâtiment Genavenir 3, 1 rue Pierre Fontaine, 91000 Evry, France

1 hMasaryk University, Research Centre for Toxic Compounds in the Environment (RECETOX),

2 Kamenice 753/5, 62500 Brno, Czech Republic

3 iDepartment of Ecosystem Analysis, Institute for Environmental Research, RWTH Aachen

4 University, 52074 Aachen, Germany

5 jNorwegian Institute for Water Research NIVA, Gaustadalléen 21, N-0349 Oslo, Norway

6 kSchool of Biosciences, The University of Birmingham, Birmingham B15 2TT, UK

7 lEberhard Karls University Tübingen, Environmental Toxicology, Center for Applied Geosciences,

8 72074 Tübingen, Germany

9

10

11 *corresponding author: beate.escher@ufz.de; Ph: +49 341 235 1244

12 **Abstract**

13 Surface waters can contain a diverse range of organic pollutants, including pesticides,

14 pharmaceuticals and industrial compounds. While bioassays have been used for water quality

15 monitoring, there is limited knowledge regarding the effects of individual micropollutants and their

16 relationship to the overall mixture effect in water samples. In this study, a battery of *in vitro*

17 bioassays based on human and fish cell lines and whole organism assays using bacteria, algae,

18 daphnids and fish embryos was assembled for use in water quality monitoring. The selection of

19 bioassays was guided by the principles of adverse outcome pathways in order to cover relevant

20 steps in toxicity pathways known to be triggered by environmental water samples. The effects of 34

21 water pollutants, which were selected based on hazard quotients, available environmental quality

22 standards and mode of action information, were fingerprinted in the bioassay test battery. There was

23 a relatively good agreement between the experimental results and available literature effect data.

24 The majority of the chemicals were active in the assays indicative of apical effects, while fewer

25 chemicals had a response in the specific reporter gene assays, but these effects were typically

26 triggered at lower concentrations. The single chemical effect data were used to improve published

1   mixture toxicity modeling of water samples from the Danube River. While there was a slight

2   increase in the fraction of the bioanalytical equivalents explained for the Danube River samples, for

3   some endpoints less than 1% of the observed effect could be explained by the studied chemicals.

4   The new mixture models essentially confirmed previous findings from many studies monitoring

5   water quality using both chemical analysis and bioanalytical tools. In short, our results indicate that

6   many more chemicals contribute to the biological effect than those that are typically quantified by

7   chemical monitoring programs or those regulated by environmental quality standards. This study

8   not only demonstrates the utility of fingerprinting single chemicals for an improved understanding

9   of the biological effect of pollutants, but also highlights the need to apply bioassays for water

10   quality monitoring in order to prevent underestimation of the overall biological effect.

11

12   **Keyword:** *In vitro;* cell-based bioassay, fish embryo toxicity test; ToxCast; mixture toxicity

1. Introduction

Chemical pollution in rivers has been identified as a major threat to ecosystem and public health (Malaj et al. 2014, Vörösmarty et al. 2010). Busch et al. (2016) identified 426 different chemicals present in European rivers, including the Danube and Rhine. While the huge number of chemicals present in surface waters indicates that targeted chemical analysis alone is unsuitable for understanding the overall chemical burden, it is still the main approach used for water quality monitoring, e.g., in the European Union Water Framework Directive (WFD) (European Commission 2011, 2012).

A diverse set of *in vivo, in vitro* and ecological indicators were proposed as monitoring tools within the WFD (Wernersson et al. 2015). *In vivo* bioassays have a long tradition of application in effluent assessment and water quality monitoring studies (Escher and Leusch 2012). In contrast, *in vitro* cellular bioassays have mainly been applied to assess technical water treatment processes, such as sewage treatment (Prasse et al. 2015), advanced water treatment (Leusch and Snyder 2015) and drinking water treatment (Neale et al. 2012). In most applications, *in vitro* bioassays are not being used as a direct link to the ecological health of aquatic organisms, but rather as a complementary analytical tool to detect and quantify chemicals via their effects within environmental mixtures. The EU Project SOLUTIONS proposes to connect both approaches and employ bioassays for water quality monitoring whilst linking them to chemical assessment (Altenburger et al. 2015). Cell-based bioassays have also been proposed as part of the first tier screening step of a new conceptual framework for monitoring water contaminants in California (Maruya et al. 2016).

Programs such as the United States Environmental Protection Agency (US EPA) Toxicity Forecaster (ToxCast) and Toxicity Testing in the 21st Century (Tox21) have screened a large number of chemicals in more than a thousand *in vitro* assays (US EPA 2015). These programs focus on human health assessment (Tice et al. 2013), with less attention on effects relevant to environmental risk assessment. More recently, Schroeder et al. (2016) proposed the application of

the high-throughput screening tools used in ToxCast and Tox21 for environmental surveillance and water quality monitoring initiatives.

Consequently, the goal of the current study was to assemble a battery of bioassays that covers both the expected effects and safeguards against overlooking others with unidentified modes of action. Cell-based assays based on different steps of cellular toxicity pathways were applied, as well as whole organism assays indicative of both apical and specific effects. Cellular effects are key parts of adverse outcome pathways (Ankley et al. 2010), with the studied bioassays covering induction of xenobiotic metabolism, receptor-mediated effects, reactive modes of action, induction of adaptive stress response pathways and cell viability. Assays using bacteria (*Aliivibrio fischeri* and *Salmonella typhimurium*), algae (*Chlamydomonas reinhardtii*), crustaceans (*Daphnia magna*), fish (embryonic *Danio rerio* and *Oryzias latipes*) and amphibians (embryonic *Xenopus laevis*) were applied, while cellular responses were evaluated using cultured fish cells (*Danio rerio*), as well as mammalian cells (human and rat cell lines), allowing comparison with the existing US EPA ToxCast database (US EPA 2015). Effect data from peer-reviewed literature and the US EPA ToxCast database were collected and compared with our experimental effect data. *In vitro* and whole organism specific effects were compared to determine if the studied pollutants acted as baseline toxicants, meaning they would result in non-specific effects, or would produce specific effects in the whole organism assays. The generated effects data were finally applied to improve mixture toxicity modeling for environmental water samples.

2. Assembling a bioanalytical test battery for surface water quality monitoring

Ideally, a bioanalytical test battery for water quality monitoring should be motivated by effects found typically in water and include assays covering a wide range of environmentally relevant modes of action and different stages of cellular toxicity pathways, as well as low-complexity whole-organism effects (Figure 1). To narrow down the large number of available bioassays to a smaller list of indicator bioassays, a balance must be struck between the desire to cover all possible effects

and practicability issues. Broad coverage of modes of action, inclusion of the contributions from all chemicals and relevance for ecological health through the alignment of the bioassays to relevant steps of adverse outcome pathways are desirable (Schroeder et al. 2016), as well as a focus on effects that have been previously observed in water samples (Escher et al. 2014). Practicability, assay robustness, applicability for less specialized laboratories and the possibility to run the assays in a high-throughput mode for low-volume tests were further considered (Escher and Leusch 2012). Reducing the sample volume requirements for each test is also important as it facilitates routine monitoring by decreasing the total volume required at each sampling location.

A large screening study using more than a hundred individual *in vitro* bioassays as well as a multifactorial assay that quantifies the activation of 25 nuclear receptors and 45 transcription factors demonstrated that the estrogen receptor (ER), and to a lesser degree the glucocorticoid receptor (GR), were amongst the most responsive hormonal nuclear receptors in wastewater and surface water testing (Escher et al. 2014). Therefore, ER reporter gene assays were a prominent component of the test battery in this study and we included assays using human and fish cell lines. Furthermore, surface water samples have also been shown to inhibit the androgen receptor (AR) (Escher et al. 2014, Jálová et al. 2013). Given the high relevance of hormone receptors, we added three transgenic assays that have been developed in recent years to assess hormone-mediated effects in early life-stage organisms (Brion et al. 2012, Fini et al. 2007, Spirhanzlova et al. 2016). Nuclear receptors triggering the activation of metabolism and other biological effects such as the aryl hydrocarbon receptor (AhR), the pregnane X receptor (PXR) and the peroxisome- proliferator-activated receptor (PPARγ) were even more prominently activated by water sample extracts (Escher et al. 2014) and play a prominent role in the proposed test battery.

We also included reporter gene assays for activation of adaptive stress responses in the present study (Table 1, Figure 1). Adaptive stress responses are key events (Simmons et al. 2009) and very recent work has elucidated the key event relationships of the toxicologically relevant Nrf2- and p53- mediated adaptive stress responses in response to reference chemicals (Hiemstra et al.

2017, Wink et al. 2014, Wink et al. 2017). Nrf2, which activates the oxidative stress response, was most responsive in water samples in a large number of studies (Escher et al. 2012, Escher et al. 2014, Escher et al. 2015b, König et al. 2017, Neale et al. 2015, Neale et al. 2017), while p53 activity was only occasionally detected in water samples (e.g., Yeh et al. 2014). A recent study using the Attagene battery as part of the ToxCast high-throughput screening assays identified this same transcription factor Nrf2 as being relevant for surface water quality monitoring and also identified the hypoxia pathway as relevant (Schroeder et al. 2016). A reporter gene assay for hypoxia was not included in the present effort but will be added to the test battery in the near future.

Mutagenicity, a reactive mode of action, is another important endpoint of ecological and human health relevance. With the exception of the p53 transcription factor assay, which is often masked by cytotoxicity when applied to surface water samples (e.g., Neale et al. 2017), available reporter gene assays are not suitable for detecting DNA damage. Instead, we included two protocols of the classic Ames assay in the present study (Table 1, Figure 1). The Ames assay detects gene mutations and has been widely used for environmental and wastewater assessment (Claxton et al. 2010, Reifferscheid et al. 2012, Umbuzeiro et al. 2016).

Whole organism assays indicative of apical effects, such as algal growth inhibition, *Daphnia* immobilization and fish embryo toxicity (FET), are more widely used for water quality assessment than cellular assays to date and can provide information about effects on mortality, growth, development and reproduction (Di Paolo et al. 2016, Wernersson et al. 2015). They are comprehensive as they cover the effects from multiple toxicity pathways leading to the same apical endpoint. Consequently, whole organism assays integrate the mixture effects of all chemicals that are present in a sample, depending on their effect potency. Therefore, they constitute an important complement to the specific bioassays. In this study we have included the Microtox test, a rapid assay based on bioluminescence inhibition of bacteria that has been demonstrated to be a useful pre-screening tool for water samples (Tang et al. 2013), the algal growth inhibition test, the immobilization test with *Daphnia magna* and the fish embryo toxicity test with *Danio rerio* as

typical representatives of apical endpoints and whole organism tests that are still legally considered to be *in vitro* test systems (Table 1, Figure 1). This latter aspect is not only an important consideration for animal ethics but also because only *in vitro* assays can be scaled up to high-throughput. In fact, several of the applied bioassays are already running on robotic systems in 384 well (cell-based assays) or 96 well (FET, algae) format, though they can also be run in a low-throughput mode, making the panel of indicator bioassays very versatile.

The panel of bioassays selected here is one possible example of a test battery design, but the reasoning provided above should be considered when designing any fit-for-purpose monitoring test battery. For example, the number of bioassays may be reduced for routine monitoring applications, whereas evaluation of highly impacted sites may require expansion of the number of bioassays to cover unusual responses triggered in whole organism endpoints. As specific and selective reporter gene assays will not capture all relevant modes of action, it is important to complement these endpoints with whole organism assays indicative of apical effects and to ensure that the bioassay battery covers different events/steps in selected toxicity pathways (Figure 1).

3. Materials and methods

3.1 Chemical selection

A total of 34 water pollutants were selected for the current study (Table 2). Thirty-two of the chemicals were selected from the list of 214 chemicals identified as relevant river pollutants by Busch et al. (2016) due to their presence in European river systems and hazard quotients $> 10^{-4}$. The hazard quotients were calculated based on the 95th percentile measured environmental concentration and measured or predicted algal, daphnid and fish 5th percentile acute effect concentration data. Further details are available in Busch et al. (2016). In addition, the pharmaceutical flutamide and the fungicide picoxystrobin were also included to represent a potent anti-androgen and a respiration inhibitor, respectively. Picoxystrobin, which inhibits respiration by blocking electron transport, was

1 added to complement the pesticide dinoseb, which also inhibits respiration but via uncoupling i.e.,

2 by protonophoric shuttle mechanisms.

3   The studied chemicals included pharmaceuticals, pesticides and industrial compounds, and

4 the selection process was based on scoring (a) their rank in the hazard quotient list, (b) the

5 availability of an environmental quality standard (EQS) from the WFD or at least a published

6 proposed EQS, (c) a unique mode of action that is not covered by a higher ranked chemical, (d)

7 diversity of chemical use groups, (e) a specific mode of action that is covered by the test battery, (f)

8 a specific mode of action expected to lead to enhanced toxicity in the whole organism bioassays and

9 (g) lack of rapid biodegradation based on BIOWIN^{TM} (US EPA 2008). The final scores for

10 prioritization are given in Table S1 of the Supplementary Information.

11   A minimum score of three points was required to qualify a chemical for experimental

12 analysis and preference was given to chemicals ranked in the top 100 of Busch et al. (2016). An

13 exception was hexadecylpyridinium, which had a score of two, but was still included as it was the

14 only surface active antiseptic compound and was ranked 2nd in Busch et al. (2016) based on its

15 hazard quotient. The non-steroidal anti-androgenic compound flutamide only had a score of 2

16 because no information about its degradability was available in BIOWIN^{TM}. All other compounds,

17 with the exception of hexadecylpyridinium, were not readily biodegradable according to

18 BIOWIN^{TM}, which meets the expectation that more recalcitrant chemicals will be found in surface

19 water. The antibiotic sulfamethoxazole was the only antibiotic selected despite its low rank in

20 Busch et al. (2016) because no other sulfonamides were ranked higher and they are an important

21 antibiotic group. Apart from these exceptions, all other chemicals were selected based on their high

22 score. If chemicals with a similar structure and function had an equal score, the chemical ranked

23 higher in Busch et al. (2016) was included.

24

25 *3.2 Data mining*

1 Effect data for the 34 selected chemicals were collected from either the peer reviewed literature or

2 the US EPA ToxCast database (US EPA 2015) for the studied bioassays. The effect data were

3 categorized based on data availability and quality (Figure S1). A chemical was considered active if

4 it had a response in the studied assay, while a chemical was considered inactive if it was tested in a

5 studied assay and had no effect (or was inactive up to cytotoxic concentrations). When experimental

6 data were unavailable for the studied assay, but available in another assay covering the same mode

7 of action, chemicals were assigned the class of likely active or likely inactive, based on whether

8 they produced an effect or not. If no experimental data were available, but the mode of action of the

9 studied chemical and bioassays theoretically matched, the chemicals were assigned as potentially

10 active or potentially inactive. Chemicals were considered as 'inactive' if they were only active in

11 the whole organism assays indicative of apical effects at effect concentrations (EC) greater than 1

12 mM. The 1 mM cut-off was based on the highest tested chemical concentration in many assays in

13 the ToxCast database. In some cases, no information could be found and the chemical was

14 designated as 'no information available'.

15

16 *3.3 Bioanalysis*

17 Bioassays applied in this study are listed in Table 1. A summary of experimental conditions, test

18 media and quality controls are provided in Table S2. Detailed standard operating procedures of the

19 bioassays, also detailing whether cytotoxicity controls were performed, are presented in the SI. All

20 chemicals were run with at least two independent replicates in each assay, with each chemical

21 assessed over a range of concentrations. Concentration-effect curves for each assay's positive

22 reference compound are shown in Figure S2. The assays were run by eight different laboratories

23 and it should be noted that not all 34 selected chemicals were run in all assays, but the average

24 coverage was 71% with the exact numbers tested given in Table 1 and the percentage tested given

25 in Table S2. The coverage depended on the capacity of each laboratory and prior knowledge of the

26 mechanism of action. For example, the algae growth inhibition assay focused on chemicals with

1 known effects on photosynthesis or other modes of action likely to lead to substantial toxicity, with

2 only 50% of the chemicals tested, while the Microtox assay was applied to all 34 chemicals. The

3 chemical concentration ranges studied in the different assays are provided in Table S3, with the

4 maximum tested concentrations selected based on the physicochemical properties of the studied

5 chemicals including solubility.

6

7 *3.4 Data evaluation*

8 EC values were derived from concentration-effect curves, with different models applied depending

9 on the assay type. All EC values are expressed in molar units. The data evaluation, which aimed to

10 be as simple and standardized as possible across the entire panel of bioassays, was developed in a

11 previous study (Escher et al. 2014). Briefly, for assays where a maximum effect could be reached,

12 such as induction of xenobiotic metabolism, hormone-mediated effects and apical effects, the

13 experimental % effect was calculated using Equation 1, where $signal_{sample}$ is the signal of the tested

14 chemical, $signal_{control}$ is the signal of the solvent control and $signal_{max}$ is the maximum response of

15 the positive reference compound in the assay. Signal refers to measured luminescence or

16 fluorescence and is specific for each assay. Details are given in the SOPs of the assays in the SI.

17

$$\%effect = \frac{signal_{sample} - signal_{control}}{signal_{max} - signal_{control}} 100\%$$

18 (1)

19

20 Both linear and log-sigmoidal concentration-effect curves were applied to determine EC values for

21 % effect data and only linear concentration-effect curves were applied to induction ratio (IR) effect

22 data.

23 Sigmoidal log concentration-effect curves (Equation 2) were applied for the reference

24 compounds in the reporter gene assays and for all chemicals in the bioassays indicative of apical

1   effects. The concentration causing 50% effect (EC_{50}) was derived from the fit applied to the

2   experimental data using Equation 2, where the slope and the EC_{50} value were the adjustable

3   parameters.

4

$$\% \text{ effect} = \frac{100\%}{1 + 10^{\text{slope}(\log EC_{50} - \log \text{concentration})}}$$

5   (2)

6

7   Log-logistic concentration-effect curves are linear with respect to linear concentrations up to an

8   effect level of approximately 40% (Escher et al. 2014). Linear concentration-effect curves

9   (Equation 3) were applied to determine the concentration causing 10% effect (EC_{10}) (Equation 4) in

10 the induction of xenobiotic metabolism, hormone-mediated effects and Ames fluctuation test assays

11 because many of the tested compounds did not reach 50% effect in these assays and because in

12 some cases cytotoxicity can mask the specific effect at high concentrations in reporter gene assays.

13 Cytotoxicity was measured in parallel to induction for many cell-based reporter gene assays and

14 concentrations that caused more than 10% cytotoxicity were excluded from the data evaluation in

15 the reporter gene assays because they would produce false positive ("cytotoxicity burst" (Judson et

16 al. 2016)) or false negative results (masking of effect by cytotoxicity).

17

$$\% \text{ effect } = \text{concentration x slope}$$

18   (3)

19

20

$$EC_{10} = \frac{10\%}{\text{slope}}$$

21   (4)

22

1 EC values derived from both linear and log-sigmoidal curves are shown in Figure S3 for the AhR

2 CALUX, HG5LN-hPXR, anti-MDA-kb2 and MELN assays, with both approaches yielding similar

3 EC values. Linear concentration-effect curves are a necessity when cytotoxicity occurs at higher

4 concentrations, but are also advantageous for the calculation of bioanalytical equivalent

5 concentrations (BEQ), which requires in principle parallel log-sigmoidal concentration-effect

6 curves, but this is not a restriction for linear concentration-effect curves as was discussed in more

7 detail in Neale et al. (2015).

8

9 For those assays where a maximum effect could not be defined, such as the adaptive stress response

10 (AREc32) and Ames microplate agar assays, an IR was calculated using Equation 5. The

11 concentration causing an induction ratio of 1.5 ($EC_{IR1.5}$) was determined for linear concentration-

12 effect curves (Equation 6) up to an IR of 5 using Equation 7.

13

14

$$IR = \frac{signal_{sample}}{signal_{control}}$$

15 (5)

16

$$IR = 1 + concentration \times slope$$

17 (6)

18

$$EC_{IR1.5} = \frac{0.5}{slope}$$

19

20 (7)

21

1   The MDA-kb2 assay was also conducted in antagonist mode (anti-MDA-kb2) and a suppression

2   ratio (SR) was calculated for anti-MDA-kb2 using Equation 8, where signal$_{agonist}$ is the response of

3   the agonist, which is typically the highest signal in the assay. A linear concentration-effect curve

4   analogous to Equation 3 was used to derive the effect concentration causing a SR of 0.2 (EC$_{SR0.2}$)

5   (Equation 9).

6

$$SR = 1 - \frac{signal_{sample} - signal_{control}}{signal_{agonist} - signal_{control}}$$

7   (8)

$$EC_{SR0.2} = \frac{0.2}{slope}$$

8   (9)

9

10

11   The described data evaluation methods could not be applied to the ChgH-GFP (spiked mode only)

12   or the THbZip-GFP (XETA) assays. Therefore, analysis of variance (ANOVA) or nonparametric

13   testing was applied to assess whether the signal of the sample was statistically different from the

14   control using either Dunnett's multiple comparison test (assuming Gaussian distribution) or Dunn's

15   multiple comparison test (assuming non-Gaussian distribution). The lowest observable effect

16   concentration (LOEC) was reported.

17

18   *3.5 Quantitative Structure- Activity Relationship*

19   Experimental EC$_{50}$ values for chemicals in assays indicative of apical effects (Microtox, algal

20   growth inhibition, *Daphnia* immobilization test and FET assay) were compared with predicted EC$_{50}$

21   values using QSARs for baseline toxicity from the literature. The aim of this analysis was to

22   determine if a chemical had a specific or non-specific effect in the whole organism assay.

The liposome water partition constant K_{lipw} was used in the QSAR instead of the octanol-water partition constant K_{ow} to account for speciation and diversity of the chemicals as K_{lipw} are applicable for polar and nonpolar baseline toxicants (Escher and Schwarzenbach 2002). For ionizable compounds, the K_{lipw} was replaced by the ionization-corrected liposome-water distribution ratio (D_{lipw}) with D_{lipw} for the studied chemicals provided in Table S4. The Microtox QSAR was developed by Escher et al. (2017) (Equation 10), the algal growth inhibition (Equation 11) and *Daphnia* immobilization test (Equation 12) were rescaled from K_{ow} according to Escher and Schwarzenbach (2002). The algal growth inhibition QSAR was based on *Chlorella vulgaris*, but the experimental data were derived for *Chlamydomonas reinhardtii*. The FET QSAR was developed by Klüver et al. (2016) (Equation 13).

$$\log\left(\frac{1}{EC_{50}}\right)(\text{Microtox}) \; (M) = 0.75 \cdot \log D_{lipw} + 0.97$$

$$\tag{10}$$

$$\log\left(\frac{1}{EC_{50}}\right)(\text{algal growth inhibition}) \; (M) = 0.91 \cdot \log D_{lipw} + 0.63$$

$$\tag{11}$$

$$\log\left(\frac{1}{EC_{50}}\right)(\textit{Daphnia} \text{ immobilization test}) \; (M) = 0.77 \cdot \log D_{lipw} + 1.89$$

$$\tag{12}$$

$$\log\left(\frac{1}{EC_{50}}\right)(\text{FET}) \; (M) = 0.99 \cdot \log D_{lipw} + 0.78$$

$$\tag{13}$$

The toxic ratio (TR_i) was derived using Equation 14 from Verhaar et al. (1992), with a chemical with a TR of >0.1 to <10 considered a baseline toxicant, while a chemical with a TR>10 was considered to have a specific effect in the assay.

$$TR_i = \frac{EC_{50\text{-baseline QSAR},i}}{EC_{50\text{-experimental},i}}$$

1   (14)

2   *3.6. Mixture toxicity modeling*

3   To compare different chemicals in one bioassay and as a basis for mixture toxicity modeling (Neale

4   et al. 2015), the relative effect potency (REP_i) was calculated from the effect concentration of the

5   reference compound divided by the effect concentration of compound i (Equation 15). The REP_i can

6   be defined for any effect concentration EC_x for any matching effect level x but it is only

7   independent of effect level for linear concentration-effect curves (Neale et al. 2015) or if the slopes

8   of the sigmoidal log-concentration-effect curves are similar for the reference compound and the

9   compound of interest, i (Villeneuve et al. 2000).

10

$$REP_i = \frac{EC_x\,(ref)}{EC_x\,(i)}$$

11   (15)

12

13   In mixture toxicity modeling we compare the bioanalytical equivalent concentrations from

14   bioassays (BEQ_{bio}, Equation 16) with the bioanalytical equivalent concentrations from chemical

15   analysis (BEQ_{chem}, Equation 17) for environmental samples using the detected concentration of an

16   individual chemical (C_i) and its REP_i (Neale et al. 2015).

17

$$BEQ_{bio} = \frac{EC_x\,(ref)}{EC_x\,(sample)}$$

18   (16)

$$BEQ_{chem} = \sum_{i=1}^{n} REP_i \cdot C_i$$

16

1   (17)

2

3 **4. Results**

4 *4.1 Availability of effect data in literature*

5 A variable picture of available effect data emerged for the 34 selected chemicals in the 20 studied

6 assays (Figure S4, Table S5). Several of the studied compounds, including fipronil, carbendazim,

7 bisphenol A, propiconazole, triclocarban, chlorophene, benzo(a)pyrene, benzo(b)fluoranthene and

8 genistein, were expected to be active, likely active or potentially active in more than 50% of the

9 studied assays. In contrast, no information was available regarding the effect of telmisartan,

10 hexadecylpyridinium, clofibric acid and mefenamic acid in 50% or more of the selected assays.

11 Data availability tended to be more comprehensive for the conventional ecotoxicology tests, such as

12 the *Daphnia* immobilization test and the FET assay, while there was no information available

13 regarding the effect for 50% or more of these relevant water contaminants for the HG5LN-hPXR,

14 ZELH-zfERalpha, ZELH-zfERbeta2 and Cyp19a1b-GFP assays. The US EPA ToxCast database

15 proved to be a valuable tool, with effect or likely effect data available for 33 of the 34 studied

16 chemicals for the AhR CALUX, PPARγ-bla, MDA-kb2 and AREc32 assays. Overall, the data

17 mining exercise highlights the lack of effect data for many of the detected water pollutants,

18 emphasizing the importance of fingerprinting their biological effects.

19

20 *4.2 Experimental effect data*

21 The experimental EC values for the studied chemicals are reported in Table 3, with a summary of

22 the active and inactive chemicals shown in Figure 2A and all concentration-effect curves provided

23 in Figures S5 to S27. Twenty-four of the representative chemicals were run in 10 or more assays,

24 with bisphenol A (70%), diazinon (55%), cyprodinil (50%) and triphenylphosphate (50%) active in

25 50% or more of the applied assays. In contrast, hexadecylpyridinium (20%), 2,4-

26 dichlorophenoxyacetic acid (18%), telmisartan (15%) and triclocarban (17%) were active in 20% or

less of the studied assays, while carbendazim was only active in the FET assay, despite being predicted to be active, likely active or potentially active in 50% of the assays and ranked at 16th in the list of potentially bioactive chemicals by Busch et al. (2016). The comparison between the experimental data and expected activity based on the data mining exercise is shown in Figure 2B. While only a qualitative assessment was possible, between 39% and 100% of the experimental effect data matched the expected activity, with over 60% similarity observed for most assays (Table S6). The observed effects of the chemicals for each assay class are described below.

4.2.1 Induction of xenobiotic metabolism

Assays indicative of activation of AhR (AhR CALUX (ID 1)), activation of PXR (HG5LN-hPXR (ID 2)) and binding to PPARγ (PPARγ-bla (ID 3)) were applied in the current study. Seven of the 18 tested chemicals were active in the AhR CALUX assay, while 20 of the 34 chemicals were active in the HG5LN-hPXR assay, which fits with previous findings by Martin et al. (2010) that many environmental chemicals can activate AhR and PXR. In contrast, only 4 out of the 18 tested chemicals were active in the PPARγ-bla assay, with the pharmaceutical telmisartan the only active compound that matched with the available effect data from the US EPA ToxCast database (US EPA 2015).

4.2.2 Hormone receptor-mediated effects

In vitro assays indicative of activation and inhibition of AR (MDA-kb2 (ID 4), anti-MDA-kb2 (ID 5)), activation of ER (MELN (ID 6), ZELH-zfERalpha (ID 7), ZELH-zfERbeta2 (ID 8)) and activation of GR (GR CALUX (ID 9)) were included in the test battery. Twenty-four compounds were run in the MDA-kb2 assay, with only three chemicals, benzo(a)pyrene, benzo(b)fluoranthene and genistein, inducing more than 10% effect in the assay. The MDA-kb2 assay contains both AR and GR (Wilson et al. 2002), and the active samples were also analyzed in the presence of the AR antagonist flutamide, which suppressed the response, confirming that the three chemicals were

indeed AR agonists (Figure S9). In contrast, 12 of the 25 tested chemicals were active in the anti-MDA-kb2 assay, with seven of these active chemicals also reported to be active in the US EPA ToxCast database (US EPA 2015). Between 2 to 9 of the 34 tested chemicals were active in the activation of ER assays, with the human based MELN assay proving more responsive than the zebrafish based ZELH-zfERalpha and ZELH-zfERbeta2 assays. In contrast, none of the 20 tested chemicals had a response in the GR CALUX assay, despite the fact that some of the studied compounds were predicted by QSARs to be potentially active based on their chemical mode of action. The fact that none of the tested chemicals were active fits with previous findings by Leusch et al. (2014), who found that the majority of tested environmental chemicals were inactive in the GR CALUX assay.

In addition to the cell-based assays, embryonic fish and tadpole-stage amphibian assays were applied to assess whether the representative chemicals can interfere with the endocrine system in whole organisms. Among the 26 tested compounds, several, including hexadecylpyridinium, diclofenac, chlorpyrifos and chlorophene, caused 100% mortality of transgenic zebrafish Cyp19a1b-GFP embryos (ID 10), while bisphenol A and genistein induced aromatase in a concentration-dependent manner in the developing brain. The concentrations inducing 100% mortality were lower than in the FET assays (ID 19 and 20), which may be related to the longer exposure duration in the cyp19a1b-GFP assay (96 h) compared to the FET assays (48 h). The ChgH-GFP assay (ID 11) provided information about estrogenic signaling in medaka embryos (unspiked mode), as well as anti-estrogenic signaling and aromatase activity when co-exposed to testosterone (spiked mode) (Spirhanzlova et al. 2016), while the THbZIP-GFP (XETA) assay (ID 12) detected chemicals that act as thyroid agonists in tadpoles (unspiked mode), as well as chemicals that interfere with thyroid receptors and thyroid hormone transport and metabolism when exposed to thyroid hormone triiodothyronine (spiked mode) (Fini et al. 2007). Of the 24 tested chemicals, EC_{10} values could only be derived for 3 chemicals, bisphenol A, triphenylphosphate and chlorophene, in unspiked mode in the ChgH-GFP assay, while 17 of the tested chemicals were

active in spiked mode. In contrast, 7 and 5 of the 20 tested chemicals were active in spiked and unspiked mode, respectively, in the THbZIP-GFP (XETA) assay.

4.2.3 Reactive modes of action

Two bacterial assays, the Ames fluctuation test (ID 13) and Ames microplate agar (ID 14) assays, were applied to assess mutagenicity in 22 and 12 tested chemicals, respectively. In both assays, benzo(a)pyrene was the only active chemical, despite a large number of the tested chemicals predicted to be potentially active based on mode of action analyses.

4.2.4 Adaptive stress response pathway

Induction of the oxidative stress response was assessed using the AREc32 assay (ID 15), with 9 of the 18 tested chemicals active in the assay. All active chemicals were predicted to be likely active based on the ARE GeneBLAzer assay in the ToxCast database (US EPA 2015), with cytotoxicity masking induction for four of the chemicals (fipronil, diclofenac, carbendazim and perfluoroctanoic acid) predicted to be likely active.

4.2.5 Apical effects in whole organisms

Between 47 to 88% of the tested chemicals were active in the whole organism assays indicative of apical effects. Diclofenac, bisphenol A, chlorophene and triclosan all caused an effect in bacteria, algae, crustaceans and embryonic fish (Table 3). Furthermore, 4-nonylphenol, cyprodinil, diazinon, diuron, propiconazole and triphenylphosphate also produced a response in all apical assays, with the exception of the Microtox assay (ID 16), which has a solubility cutoff for baseline toxicants with high melting points as described in more detail by Escher et al. (2017). Benzo(b)fluoranthene had no effect in any of the assays up to the highest tested concentration, while effects were only observed for metoprolol at concentrations above 1×10^{-3} M in the algae growth inhibition (ID 17) and *Daphnia* immobilization (ID 18) assays. Fipronil, triclocarban and benzo(a)pyrene only

1   produced a positive response in the *Daphnia* immobilization assay, with no observed effect at the

2   highest concentrations tested in the other apical assays.

3

4   **5. Discussion**

5   *5.1 Suitability of the bioassay test battery to detect representative pollutants*

6   The majority of assays indicative of induction of xenobiotic metabolism, adaptive stress responses

7   and apical effects in whole organisms were able to detect a large number of studied chemicals

8   compared to assays indicative of hormone receptor-mediated effects or reactive modes of action. In

9   many cases, the same chemical was active in a number of assays, but at different concentrations

10   (Figure 2A). For example, consistent with mechanistic considerations, the endocrine-active

11   chemical bisphenol A induced an effect at lower concentrations in the MELN, anti-MDA-kb2 and

12   Cyp19a1b-GFP assays compared to the AREc32 assay and apical endpoints. While some modes of

13   action, such as endocrine disruption, were well covered by assays indicative of specific effects,

14   other modes of action were covered indirectly by the whole organism assays indicative of apical

15   effects. This suggests that a single bioassay cannot be representative for all effects but the applied

16   bioassay battery was generally suitable to detect the effect of the selected pollutants.

17   The discrepancy between expected and observed activity in the same bioassay (Figure 2B)

18   could be attributed to several factors including the quality of the expected activity data and

19   solubility issues. For example, the experimental results from the GR CALUX assay had the lowest

20   correspondence with the expected activity, but many of the chemicals expected to be potentially

21   active were based on a theoretical match with mode of action from QSAR predictions, rather than

22   being based on experimental data. Furthermore, many of the studied pollutants showed less effect

23   than predicted in the Microtox assay due to the solubility cutoff for chemicals with high melting

24   points, which is discussed in more detail in Escher et al. (2017) and which might also apply for

25   other bioassays.

1   Bioassays are typically applied to enriched water samples, with solid-phase extraction

2   commonly used. In order to assess whether the studied assays were suitable for detecting the

3   individual compounds from the list of studied water contaminants at environmentally relevant

4   concentrations, occurrence data were compared with the EC values. Busch et al. (2016) reported the

5   95th percentile of all measured environmental concentrations (MEC95) from six studies on

6   European rivers, with the maximum MEC95 reprinted in Table S7. Based on the MEC95 values the

7   MELN assay could potentially detect bisphenol A and genistein in water samples after 17 and 15-

8   fold enrichment, respectively, while the *Daphnia* immobilization assay could detect chlorpyrifos

9   and diazinon in water samples after 17 and 3-fold enrichment, respectively (Table S7). A larger

10   number of assays, including AhR CALUX, HG5LN-hPXR, PPARγ-bla, Anti-MDA-kb2, ZELH-

11   zfERalpha and ZELH-zfERbeta2, are potentially able to detect more of the studied chemicals with

12   up to 500-fold enrichment. However, it is important to note that in environmental samples,

13   bioassays will not detect single compounds, but rather mixtures of compounds. Therefore, it does

14   not mean that a particular bioassay is unsuitable for water quality monitoring if a single chemical is

15   not detected. It is rather likely that none of the known pollutants cause an effect alone, but instead it

16   is the mixture effect that will be detected by the assay. This is termed a "something from nothing"

17   effect (Silva et al. 2002), which has been demonstrated numerous times in defined mixture

18   experiments where chemicals mixed at concentrations below their observable effect level show an

19   effect in combination.

20

21   *5.2 Comparing cellular effects with whole organism specific effects*

22   Three *in vitro* assays indicative of activation of ER in a human cell line (MELN) and zebrafish cell

23   lines (ZELH-zfERalpha and ZELH-zfERbeta2) and two whole organism assays indicative of ER-

24   regulated effects in early life-stage fish (Cyp19a1b-GFP (zebrafish) and ChgH-GFP (medaka) were

25   applied in the current study, allowing a comparison of estrogenic effects between different cell lines

26   and   different   organisms.   While   9   of   the   tested   chemicals   (diazinon,   bisphenol A,

1   triphenylphosphate, 4-nonylphenol, chlorpyrifos, benzo(a)pyrene, benzo(b)fluoranthene,

2   terbutylazine and genistein) were active in the MELN assay, only two, bisphenol A and genistein,

3   had a response in the ZELH-zfERalpha and ZELH-zfERbeta2 assays. This could be due to either

4   species-specific differences in sensitivity or selectivity, as has previously been observed in

5   environmental samples for the studied assays (Sonavane et al. 2016), and/or due to the higher

6   metabolic capacity of the hepatic zebrafish cell line, which may potentially lead to more

7   biotransformation of the test compounds than the MELN assay, which is based on a breast cancer

8   cell line (Le Fol et al. 2015). At the organism level, the Cyp19a1b-GFP assay results were very

9   consistent with the ZELH-zfERalpha and ZELH-zfERbeta2 assays, with bisphenol A and genistein

10   able to induce the expression of ER-regulated aromatase in the developing brain. Other bisphenol

11   compounds have also been shown to induce estrogenic responses in the studied zebrafish *in vitro*

12   and whole organism assays (Le Fol et al. 2017). This emphasizes the relevance of applying fish-

13   specific *in vitro* assays, as well as early life-stage organism assays, for environmental risk

14   assessment.

15   The ChgH-GFP assay using medaka embryos also responded to bisphenol A, but genistein,

16   which was active in all other estrogenic assays, did not induce a response in the ChgH-GFP assay.

17   This has also been observed in a previous study and was attributed to the lower sensitivity of

18   medaka to genistein, with reverse transcriptase-polymerase chain reaction (RT-PCR) indicating no

19   change in choriogenin H or vitellogenin transcription in the presence of genistein (Scholz et al.

20   2005). In addition to bisphenol A, EC_{10} values could be determined for triphenylphosphate and

21   chlorophene in unspiked mode, with triphenylphosphate also active in the MELN assay. The

22   observed differences between the zebrafish and medaka assays may be due to differences in species

23   sensitivity to (xeno)-estrogens. It is noteworthy that both transgenic models presented differences in

24   sensitivity to the reference compound 17α-ethinylestradiol (Figure S2) and to some xenoestrogens

25   (e.g. bisphenol A and genistein) (Figures S14 and S15), thus highlighting the different intrinsic

26   sensitivities in the estrogenic response. The exposure duration (96 h for the Cyp19a1b-GFP assay

and 24 h for the ChgH-GFP assay), the tissue context (brain versus liver) and the metabolic capacity of the models may account for these differences. The variability in responsiveness and sensitivity of the many different estrogenicity assays is well known (Kunz et al. 2017) and can be used to investigate the exact toxic mechanisms and to differentiate true effects from assay interferences (Browne et al. 2015, Judson et al. 2015). In turn this does not mean that for water quality monitoring that many different assays indicative of the one endpoint should be used or that one assay should be favored over others. However, when undertaking mixture toxicity modeling or when deriving effect-based trigger values, one should use data from the same bioassay as the BEQ value will be specific for each assay (Escher et al. 2015a).

5.3 Are the representative compounds acting specifically or as baseline toxicants on the apical endpoints?

The majority of tested chemicals were active in the whole organism assays indicative of apical effects given these endpoints cover effects of multiple toxicity pathways. To determine whether the tested chemicals were baseline toxicants in the studied assays ($0.1 < TR < 10$) or if they had a specific mode of action ($TR > 10$), the derived EC_{50} values were compared with EC_{50} values predicted by baseline toxicity QSARs (Figure 3), with the TR calculated using Equation 14 (Table S8).

The majority of studied chemicals acted as baseline toxicants ($01 < TR < 10$) for the 30-min bioluminescence inhibition assay with *A. fischeri* (Microtox assay, Figure 3A). The exceptions were 1,2-benzisothiazolinone, which had a TR of 488, and dinoseb, which had a TR of 22. 1,2-Benzisothiazolinone is a biocide and soft electrophile, while the pesticide dinoseb is a potent uncoupler, meaning that it can interfere with cellular energy transduction (Escher et al. 1996), and both can have a specific effect on bacteria, as supported by the higher TR. The antibiotic sulfamethoxazole had a TR of 4.8, thus was not considered as having a specific effect, which was also observed by Tang et al. (2013), although antibiotics are often specifically acting in bacteria.

24

This may be due to the 30 min exposure duration, with previous studies showing that antibiotics often only have a specific effect on bacteria after longer exposure durations when growth becomes important (Backhaus et al. 1997).

Despite being based on a different algal species, the experimental data for *Chlamydomonas reinhardtii* fit well with the *Chlorella vulgaris* baseline toxicity QSAR predictions (Figure 3B). The one chemical expected to be acting specifically was the photosystem II inhibitor diuron, which had a TR of 852, which confirms earlier studies identifying diuron to be specifically acting on growth in green algae (Neuwoehner et al. 2008).

Five chemicals, chlorpyrifos, diazinon, fipronil, cyprodinil and triclocarban, had a TR > 10 in the *Daphnia* immobilization assay (Figure 3C). The most potent chemicals, chlorpyrifos (TR=123818) and diazinon (TR=96043), are both neuroactive chemicals that inhibit acetylcholinesterase (AChE), with previous studies showing that daphnids are particularly sensitive to the AChE inhibitors (Vaal et al. 2000).

Many chemicals act as baseline toxicants in the FET assay (Ellison et al. 2016, Klüver et al. 2016), with the majority of pollutants in our test set showing baseline toxicity. Three compounds, carbendazim, 1,2-benzisothiazolinone and mefenamic acid, had a TR > 10 in the FET assay (Figure 3D). The fungicide carbendazim had a TR of 1775 and has previously been shown to have a specific effect in the FET assay (Schmidt et al. 2016). In contrast, the TR was less than 0.1 for 4-nonylphenol, which may be due to its hydrophobicity. The EC_{50} values of some of the more hydrophobic compounds, including 4-nonylphenol and triphenylphosphate, were lower in the FET assay conducted in glass vials than the FET assay run in 96 well polystyrene plates, despite both assays using a 48 h exposure period. Chemical sorption to the plastic may have reduced the bioavailable chemical concentration in the 96 well plate assay. This limitation could potentially be overcome through passive dosing, which has been recently applied to the FET assay using silicone O-rings (Vergauwen et al. 2015), though this approach has yet to be scaled down to the 96 well format.

1

5.4 Application of the fingerprinting data for mixture toxicity modeling

3 Recent work has focused on trying to understand mixture effects of known and analyzed chemicals

4 compared to the overall risk of known mixtures in wastewater treatment plant effluent and surface

5 water and it appears that a small number of chemicals determine the overall risk (Backhaus and

6 Karlsson 2014, Munz et al. 2017). In case studies on water treatment and surface water, some

7 industrial compounds, including benzothiazoles and fragrance chemicals, may be present at high

8 concentrations, but they were often less potent than other chemicals and consequently contributed

9 little to the observed biological effects (Tang et al. 2014). In another study, pesticides present at

10 lower concentrations than pharmaceuticals were found to be the drivers of toxicity in aquatic

11 organisms (Munz et al. 2017).

12 Another approach is to additionally quantify the contribution of detected chemicals to the

13 observed effect in a bioassay. This helps to not only apportion the toxicity to the various known

14 constituents of the mixture, but also to estimate the contribution of unknown chemicals in the

15 mixture effect of the entire water sample. This approach is termed iceberg modeling, because it

16 allows a quantification of the unknown chemicals contribution to the effect of the mixture without

17 having to identify them. Both modeling approaches for mixture effects can be combined as we have

18 demonstrated in previous studies (Neale et al. 2015, Neale et al. 2017, Tang et al. 2013, Tang et al.

19 2014) but this approach is limited by the lack of effect data for commonly detected chemicals. As a

20 result, it is often unclear whether the detected chemicals are active, and potentially contributing to

21 the effect, or inactive in the applied assays. This knowledge gap can be overcome by fingerprinting

22 the effects of relevant chemicals in bioassays. Consequently, the generated effect data from the

23 current study can be applied for improved mixture toxicity modeling.

24 If the majority of chemicals in mixtures are acting in a concentration additive manner, which

25 has been confirmed for most reporter gene assays and is also supported by the design principle of a

26 reporter gene assay that is based on a single mode of action, BEQ values from bioanalysis (BEQ_{bio},

Equation 16) can be compared to BEQ values from chemical analysis (BEQ_{chem}, Equation 17) (Neale et al. 2015). This approach has been applied to a range of water types, using initially estrogenicity assays (Aerni et al. 2004, Leusch et al. 2010, Murk et al. 2002) but later other reporter gene assays, such as those that detect additional hormonal effects (König et al. 2017), PXR and AhR activity (Creusot et al. 2010) or adaptive stress responses (Escher et al. 2013, Tang et al. 2014).

It is more daring to apply the iceberg-modeling concept to apical endpoints in whole organism tests because, as demonstrated in Section 5.3, single chemicals can have very high TR, which would mean that diverse modes of action are involved and that the mixture model of concentration addition would not necessarily apply. However, in those studies simulating environmental mixtures using the Microtox assay, water pollutants were typically well described by the mixture toxicity model of concentration addition (Escher et al. 2017, Tang et al. 2013), therefore the iceberg-modeling approach should also be feasible for this endpoint. Likewise for the investigated water pollutants, the TRs in the FET assays were fairly low, which means that potential specific effects were not of high potency; therefore, it can be expected that deviations from the model of concentration addition would not be substantial.

In contrast, for algae and daphnia, some of the water pollutants that were tested here had high TRs and therefore the assumption of concentration addition can be challenged. In practice, we observed that concentration addition was a robust tool for algal toxicity when considering typical water pollutants at concentrations encountered in environmental water samples (Tang et al. 2013), but this remains to be confirmed for daphnids. Ongoing experiments with defined mixtures of the water pollutants presented here, carried out within the framework of the EU project SOLUTIONS (*www.solutions-project.eu*) will hopefully shed more light on these questions.

Using the EC values fingerprinted in the current study, mixture toxicity modeling of Danube River samples previously presented in Neale et al. (2015) was revised for the HG5LN-hPXR,

27

MELN and FET assays. The revised REP_i values are shown in Table S9, while the REP_i values used previously can be found in Table 3 of Neale et al. (2015).

The percentage of effect explained by individual detected chemicals for both existing and revised mixture toxicity modeling is shown in Figure 4. EC values in the HG5LN-hPXR assay are now available for 17 of the detected chemicals, but even with the additional chemicals, only 0.2% of PXR activation can be explained. Previously, there was no information about the effect of genistein in the HG5LN-hPXR assay, but using the new effect data genistein was found to contribute to over half of the explained effect in some of the samples. A similar picture emerges for the FET assay; although EC values are now available for 19 chemicals, they account for less than 0.4% of the observed effect. Since a large number of chemicals can produce a response in both of these assays, the low fraction of effect that can be explained is not surprising, even with EC values for up to 19 chemicals. In contrast, much of the observed effect in the MELN assay at some sites can already be explained, with the new EC values having a negligible effect as diazinon and terbutylazine are weakly active in the assay. It should be noted that the contribution of some of the detected chemicals changed with the revised EC values, with some chemicals found to be more or less potent than the previously published. For example, estrone was found to be more potent in the MELN assay than previously reported by Pillon et al. (2005). This exercise highlights the importance of applying effect-based tools for water quality assessment as targeted chemical analysis alone often provides a limited view of the chemical burden.

6. Conclusions

A battery of bioassays covering different modes of action was assembled in the current study to detect the effects of representative water pollutants. It is important to stress that the exact type of bioassay is not essential but a diverse panel of bioassays that includes apical endpoints is essential, as well as specific bioassays indicative of crucial steps in toxicity pathways relevant for

micropollutants occurring in surface water. This is why the selection of bioassays was also guided by what types of effects were detected in the surface water samples.

All studied chemicals were active in at least one of the applied assays, with the industrial compound bisphenol A active in 70% of the studied assays. As expected, more chemicals were active in the assays indicative of apical effects, but the concentrations at which some chemicals induced an effect were much lower in assays that were based on a specific pathway/mechanism. For example, genistein and bisphenol A were more responsive in the ER mediated assays compared to the whole organism assays. These patterns illustrate how important it is to combine bioassays with apical endpoints and specific pathway endpoints to comprehensively capture the hazard potential of micropollutants in surface water.

There was reasonable agreement between the experimental results and the expected activity, though the data mining exercise highlighted the lack of available data, particularly for some recently developed mechanism-based assays, which is a common limitation of such investigations. Effect data for individual chemicals are required as input parameters for mixture toxicity modeling and the data generated in the current study will be applied in ongoing and future studies to assess the mixture effects of representative pollutants in river and other water samples.

References

Aerni, H.R., Kobler, B., Rutishauser, B.V., Wettstein, F.E., Fischer, R., Giger, W., Hungerbuhler, A., Marazuela, M.D., Peter, A., Schonenberger, R., Vogeli, A.C., Suter, M.J.F., Eggen, R.I.L., 2004. Combined biological and chemical assessment of estrogenic activities in wastewater treatment plant effluents. Anal. Bioanal. Chem. 378(3), 688-696.

Altenburger, R., Aït-Aïssa, S., Antczak, P., Backhaus, T., Barcelo, D., Seiler, T.-B., Brion, F., Busch, W., Chipman, K., de Alda, M.L., de Aragao Umbuzeiro, G., Escher, B.I., Falciani, F., Faust, M., Focks, A., Hilscherová, K., Hollender, J., Hollert, H., Jager, F., Jahnke, A., Kortenkamp, A., Krauss, M., Lemkine, G.F., Munthe, J., Neumann, S., Schymanski, E.L., Scrimshaw, M., Segner, H., Slobodnik, J., Smedes, F., Kughathas, S., Teodorovic, I., Tindall, A.J., Tollefsen, K.E., Walz, K.-H., Williams, T.D., Van den Brink, P.J., van Gils, J., Vrana, B., Zhang, X., Brack, W., 2015. Future water quality monitoring - Adapting tools to deal with mixtures of pollutants in water resource management. Sci. Total Environ. 512-513, 540-551.

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., Serrrano, J.A., Tietge, J.E., Villeneuve, D.L., 2010. Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. Environ. Toxicol. Chem. 29(3), 730-741.

Backhaus, T., Froehner, K., Altenburger, R., Grimme, L.H., 1997. Toxicity testing with *Vibrio fischeri*: A comparison between the long term (24 h) and the short term (30 min) bioassay. Chemosphere 35(12), 2925-2938.

1   Backhaus, T., Karlsson, M., 2014. Screening level mixture risk assessment of pharmaceuticals in

2   STP effluents. Water Res. 49, 157-165.

3   Balaguer, P., Francois, F., Comunale, F., Fenet, H., Boussioux, A.M., Pons, M., Nicolas, J.C.,

4   Casellas, C., 1999. Reporter cell lines to study the estrogenic effects of xenoestrogens. Sci.

5   Total Environ. 233(1-3), 47-56.

6   Basketter, D.A., Rodford, R., Kimber, I., Smith, I., Wahlberg, J.E., 1999. Skin sensitization risk

7   assessment: a comparative evaluation of 3 isothiazolinone biocides. Contact Derm. 40(3),

8   150-154.

9   Brennan, J.C., He, G.C., Tsutsumi, T., Zhao, J., Wirth, E., Fulton, M.H., Denison, M.S., 2015.

10   Development of species-specific Ah receptor-responsive third generation CALUX cell lines

11   with enhanced responsiveness and improved detection limits. Environ. Sci. Technol. 49(19),

12   11903-11912.

13   Brion, F., Le Page, Y., Piccini, B., Cardoso, O., Tong, S.K., Chung, B.C., Kah, O., 2012. Screening

14   estrogenic activities of chemicals or mixtures *in vivo* using transgenic (cyp19a1b-GFP)

15   zebrafish embryos. Plos One 7(5), e36069.

16   Browne, P., Judson, R.S., Casey, W.M., Kleinstreuer, N.C., Thomas, R.S., 2015. Screening

17   chemicals for estrogen receptor bioactivity using a computational model. Environ. Sci.

18   Technol. 49(14), 8804-8814.

19   Busch, W., Schmidt, S., Kuhne, R., Schulze, T., Krauss, M., Altenburger, R., 2016. Micropollutants

20   in European rivers: A mode of action survey to support the development of effect-based

21   tools for water monitoring. Environ. Toxicol. Chem. 35(8), 1887-1899.

22   Claxton, L.D., Umbuzeiro, G.D.A., DeMarini, D.M., 2010. The *Salmonella* mutagenicity assay:

23   The stethoscope of genetic toxicology for the 21st century. Environ. Health Perspect. 118,

24   1515-1522.

25   Cosnefroy, A., Brion, F., Maillot-Marechal, E., Porcher, J.M., Pakdel, F., Balaguer, P., Aït-Aïssa,

26   S., 2012. Selective activation of zebrafish estrogen receptor subtypes by chemicals by using

stable reporter gene assay developed in a zebrafish liver cell line. Toxicol. Sci. 125(2), 439-449.

Creusot, N., Kinani, S., Balaguer, P., Tapie, N., LeMenach, K., Maillot-Maréchal, E., Porcher, J.M., Budzinski, H., Aït-Aïssa, S., 2010. 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. 396(2), 569-583.

DeMarini, D.M., Dallas, M.M., Lewtas, J., 1989. Cytotoxicity and effect on mutagenicity of buffers in a microsuspension assay. Teratogen. Carcin. Mut. 9(5), 287-295.

Di Paolo, C., Ottermanns, R., Keiter, S., Aït-Aïssa, S., Bluhm, K., Brack, W., Breitholtz, M., Buchinger, S., Carere, M., Chalon, C., Cousin, X., Dulio, V., Escher, B.I., Hamers, T., Hilscherová, K., Jarque, S., Jonas, A., Maillot-Marechal, E., Marneffe, Y., Nguyen, M.T., Pandard, P., Schifferli, A., Schulze, T., Seidensticker, S., Seiler, T.B., Tang, J., van der Oost, R., Vermeirssen, E., Zounkova, R., Zwart, N., Hollert, H., 2016. Bioassay battery interlaboratory investigation of emerging contaminants in spiked water extracts - Towards the implementation of bioanalytical monitoring tools in water quality assessment and monitoring. Water Res. 104, 473-484.

Ellison, C.M., Piechota, P., Madden, J.C., Enoch, S.J., Cronin, M.T.D., 2016. Adverse outcome pathway (AOP) informed modeling of aquatic toxicology: QSARs, read-across, and interspecies verification of modes of action. Environ. Sci. Technol. 50(7), 3995-4007.

Escher, B.I., Snozzi, M., Schwarzenbach, R.P., 1996. Uptake, speciation, and uncoupling activity of substituted phenols in energy transducing membranes. Environ. Sci. Technol. 30(10), 3071-3079.

Escher, B.I., Schwarzenbach, R.P., 2002. Mechanistic studies on baseline toxicity and uncoupling of organic compounds as a basis for modeling effective membrane concentrations in aquatic organisms. Aquat. Sci. 64(1), 20-35.

1   Escher, B.I., Dutt, M., Maylin, E., Tang, J.Y.M., Toze, S., Wolf, C.R., Lang, M., 2012. Water

2   quality assessment using the AREc32 reporter gene assay indicative of the oxidative stress

3   response pathway. J. Environ. Monit. 14(11), 2877-2885.

4   Escher, B.I., Leusch, F.D.L., 2012. Bioanalytical tools in water quality assessment, IWA

5   Publishing, London.

6   Escher, B.I., van Daele, C., Dutt, M., Tang, J.Y.M., Altenburger, R., 2013. Most oxidative stress

7   response in water samples comes from unknown chemicals: The need for effect-based water

8   quality trigger values. Environ. Sci. Technol. 47(13), 7002-7011.

9   Escher, B.I., Allinson, M., Altenburger, R., Bain, P.A., Balaguer, P., Busch, W., Crago, J.,

10   Denslow, N.D., Dopp, E., Hilscherová, K., Humpage, A.R., Kumar, A., Grimaldi, M.,

11   Jayasinghe, B.S., Jarosova, B., Jia, A., Makarov, S., Maruya, K.A., Medvedev, A., Mehinto,

12   A.C., Mendez, J.E., Poulsen, A., Prochazka, E., Richard, J., Schifferli, A., Schlenk, D.,

13   Scholz, S., Shiraish, F., Snyder, S., Su, G.Y., Tang, J.Y.M., van der Burg, B., van der

14   Linden, S.C., Werner, I., Westerheide, S.D., Wong, C.K.C., Yang, M., Yeung, B.H.Y.,

15   Zhang, X.W., Leusch, F.D.L., 2014. Benchmarking organic micropollutants in wastewater,

16   recycled water and drinking water with *in vitro* bioassays. Environ. Sci. Technol. 48(3),

17   1940-1956.

18   Escher, B.I., Neale, P.A., Leusch, F.D.L., 2015a. Effect-based trigger values for *in vitro* bioassays:

19   Reading across from existing water quality guideline values. Water Res. 81, 137-148.

20   Escher, B.I., Neale, P.A., Leusch, F.D.L., 2015b. Effect-based trigger values for in vitro bioassays:

21   Reading across from existing water quality guideline values. Water Res. 81(0), 137-148.

22   Escher, B.I., Baumer, A., König, M., Kühnert, C., Mayer, P., Klüver, N., 2017. General baseline

23   toxicity QSAR for non-polar, polar and ionisable chemicals in the bioluminescence

24   inhibition assay with *Aliivibrio fischeri*. Env. Sci. Process. Impact 19(3), 414-428.

European Commission, 2011. Common implementation strategy for the water framework directive (2000/60/EC). Guidance document No. 27. Technical guidance for deriving environmental quality standards, Brussels.

European Commission, 2012. Proposal for a Directive of the European Parliament and the Council amending Directives 2000/60/EC and 2008/105/EC as regards priority substances in the field of water policy, COM/2011/0876 final - 2011/0429 (COD), Brussels.

European Commission, 2013. Directive 2013/39/EU of the European Parliament and the Council of 12. August 2013 amending Directives 2000/60/EC and 2008/105/EC as regards priority substances in the field of water policy. Off. J. Eur. Union, L226/221-L226/217.

Fini, J.B., Le Mevel, S., Turque, N., Palmier, K., Zalko, D., Cravedi, J.P., Demeneix, B.A., 2007. An *in vivo* multiwell-based fluorescent screen for monitoring vertebrate thyroid hormone disruption. Environ. Sci. Technol. 41(16), 5908-5914.

Hiemstra, S., Niemeijer, M., Koedoot, E., Wink, S., Pip, J.E., Vlasveld, M., de Zeeuw, E., van Os, B., White, A., van de Water, B., 2017. Comprehensive Landscape of Nrf2 and p53 Pathway Activation Dynamics by Oxidative Stress and DNA Damage. Chem. Res. Toxicol. 30(4), 923-933.

Invitrogen, 2010. GeneBLAzer® PPAR gamma 293H DA and PPAR gamma-UAS-bla 293H Cell-based Assay Protocol, Invitrogen Corporation, Carslbad, CA, USA.

Jálová, V., Jarošová, B., Bláha, L., Giesy, J.P., Ocelka, T., Grabic, R., Jurčíková, J., Vrana, B., Hilscherová, K., 2013. Estrogen-, androgen- and aryl hydrocarbon receptor mediated activities in passive and composite samples from municipal waste and surface waters. Environ. Int. 59, 372-383.

Judson, R., Houck, K., Martin, M., Richard, A.M., Knudsen, T.B., Shah, I., Little, S., Wambaugh, J., Setzer, R.W., Kothya, P., Phuong, J., Filer, D., Smith, D., Reif, D., Rotroff, D., Kleinstreuer, N., Sipes, N., Xia, M.H., Huang, R.L., Crofton, K., Thomas, R.S., 2016.

Analysis of the effects of cell stress and cytotoxicity on *in vitro* assay activity across a diverse chemical and assay space. Toxicol. Sci. 152(2), 323-339.

Judson, R.S., Magpantay, F.M., Chickarmane, V., Haskell, C., Tania, N., Taylor, J., Xia, M.H., Huang, R.L., Rotroff, D.M., Filer, D.L., Houck, K.A., Martin, M.T., Sipes, N., Richard, A.M., Mansouri, K., Setzer, R.W., Knudsen, T.B., Crofton, K.M., Thomas, R.S., 2015. Integrated model of chemical perturbations of a biological pathway using 18 *in vitro* high-throughput screening assays for the estrogen receptor. Toxicol. Sci. 148(1), 137-154.

Klüver, N., Vogs, C., Altenburger, R., Escher, B.I., Scholz, S., 2016. Development of a general baseline toxicity QSAR model for the fish embryo acute toxicity test. Chemosphere 164, 164-173.

König, M., Escher, B.I., Neale, P.A., Krauss, M., Hilscherová, K., Novák, J., Teodorović, I., Schulze, T., Seidensticker, S., Kamal Hashmi, M.A., Ahlheim, J., Brack, W., 2017. Impact of untreated wastewater on a major European river evaluated with a combination of *in vitro* bioassays and chemical analysis. Environ. Pollut. 220, 1220–1230.

Kunz, P.Y., Simon, E., Creusot, N., Jayasinghe, B.S., Kienle, C., Maletz, S., Schifferli, A., Schönlau, C., Aït-Aïssa, S., Denslow, N.D., Hollert, H., Werner, I., Vermeirssen, E.L.M., 2017. Effect-based tools for monitoring estrogenic mixtures: Evaluation of five *in vitro* bioassays. Water Res., In Press http://dx.doi.org/10.1016/j.watres.2016.1010.1062.

Le Fol, V., Aït-Aïssa, S., Cabaton, N., Dolo, L., Grimaldi, M., Balaguer, P., Perdu, E., Debrauwer, L., Brion, F., Zalko, D., 2015. Cell-specific biotransformation of benzophenone-2 and bisphenol-S in zebrafish and human *in vitro* models used for toxicity and estrogenicity screening. Environ. Sci. Technol. 49(6), 3860-3868.

Le Fol, V., Aït-Aïssa, S., Sonavane, M., Porcher, J.-M., Balaguer, P., Cravedi, J.-P., Zalko, D., Brion, F., 2017. *In vitro* and *in vivo* estrogenic activity of BPA, BPF and BPS in zebrafish-specific assays. Ecotoxicol. Environ. Saf. 142, 150-156.

1   Lemaire, G., Mnif, W., Pascussi, J.M., Pillon, A., Rabenoelina, F., Fenet, H., Gomez, E., Casellas,

2   C., Nicolas, J.C., Cavailles, V., Duchesne, M.J., Balaguer, P., 2006. Identification of new

3   human pregnane X receptor ligands among pesticides using a stable reporter cell system.

4   Toxicol. Sci. 91(2), 501-509.

5   Leusch, F.D.L., De Jager, C., Levi, Y., Lim, R., Puijker, L., Sacher, F., Tremblay, L.A., Wilson,

6   V.S., Chapman, H.F., 2010. Comparison of five *in vitro* bioassays to measure estrogenic

7   activity in environmental waters. Environ. Sci. Technol. 44(10), 3853-3860.

8   Leusch, F.D.L., Khan, S.J., Laingam, S., Prochazka, E., Froscio, S., Trang, T., Chapman, H.F.,

9   Humpage, A., 2014. Assessment of the application of bioanalytical tools as surrogate

10   measure of chemical contaminants in recycled water. Water Res. 49, 300-315.

11 Leusch, F.D.L., Snyder, S.A., 2015. Bioanalytical tools: half a century of application for potable

12   reuse. Environ. Sci. Water Res. Technol. 1, 606-621.

13 Malaj, E., von der Ohe, P.C., Grote, M., Kuhne, R., Mondy, C.P., Usseglio-Polatera, P., Brack, W.,

14   Schafer, R.B., 2014. Organic chemicals jeopardize the health of freshwater ecosystems on

15   the continental scale. Proc. Natl. Acad. Sci. U.S.A. 111(26), 9549-9554.

16 Martin, M.T., Dix, D.J., Judson, R.S., Kavlock, R.J., Reif, D.M., Richard, A.M., Rotroff, D.M.,

17   Romanov, S., Medvedev, A., Poltoratskaya, N., Gambarian, M., Moeser, M., Makarov, S.S.,

18   Houck, K.A., 2010. Impact of environmental chemicals on key transcription regulators and

19   correlation to toxicity end points within EPA's ToxCast program. Chem. Res. Toxicol.

20   23(3), 578-590.

21 Maruya, K.A., Dodder, N.G., Mehinto, A.C., Denslow, N.D., Schlenk, D., Snyder, S.A., Weisberg,

22   S.B., 2016. A tiered, integrated biological and chemical monitoring framework for

23   contaminants of emerging concern in aquatic ecosystems. Integr. Environ. Assess. Manag.

24   12(3), 540-547.

25 Mortelmans, K., Zeiger, E., 2000. The Ames Salmonella/microsome mutagenicity assay. Mutat.

26   Res. Fund. Mol. Mech. Mut. 455(1-2), 29-60.

1 Munz, N.A., Burdon, F.J., de Zwart, D., Junghans, M., Melo, L., Reyes, M., Schonenberger, U.,

2   Singer, H.P., Spycher, B., Hollender, J., Stamm, C., 2017. Pesticides drive risk of

3   micropollutants in wastewater-impacted streams during low flow conditions. Water Res.

4   110, 366-377.

5 Murk, A.J., Legler, J., van Lipzig, M.M.H., Meerman, J.H.N., Belfroid, A.C., Spenkelink, A., van

6   der Burg, B., Rijs, G.B.J., Vethaak, D., 2002. Detection of estrogenic potency in wastewater

7   and surface water with three *in vitro* bioassays. Environ. Toxicol. Chem. 21(1), 16-23.

8 Neale, P.A., Antony, A., Bartkow, M., Farre, M., Heitz, A., Kristiana, I., Tang, J.Y.M., Escher, B.I.,

9   2012. Bioanalytical assessment of the formation of disinfection by-products in a drinking

10   water treatment plant. Environ. Sci. Technol. 46(18), 10317–10325.

11 Neale, P.A., Aït-Aïssa, S., Brack, W., Creusot, N., Denison, M.S., Deutschmann, B., Hilscherová,

12   K., Hollert, H., Krauss, M., Novak, J., Schulze, T., Seiler, T.-B., Serra, H., Shao, Y., Escher,

13   B.I., 2015. Linking *in vitro* effects and detected organic micropollutants in surface water

14   using mixture-toxicity modeling. Environ. Sci. Technol. 49(24), 14614-14624.

15 Neale, P.A., Munz, N.A., Aït-Aïssa, S., Altenburger, R., Brion, F., Busch, W., Escher, B.I.,

16   Hilscherová, K., Kienle, C., Novak, J., Seiler, T.B., Shao, Y., Stamm, C., Hollender, J.,

17   2017. Integrating chemical analysis and bioanalysis to evaluate the contribution of

18   wastewater effluent on the micropollutant burden in small streams. Sci. Total Environ. 576,

19   785–795.

20 Neuwoehner, J., Junghans, M., Koller, M., Escher, B.I., 2008. QSAR analysis and specific

21   endpoints for classifying the physiological modes of action of biocides in synchronous green

22   algae. Aquat. Toxicol. 90(1), 8-18.

23 OECD, 2004. Test No. 202: Daphnia sp. Acute Immobilisation Test, OECD Publishing, Paris.

24 OECD, 2011. Test No. 201: Freshwater Alga and Cyanobacteria, Growth Inhibition Test, OECD

25   Publishing, Paris.

1 OECD, 2013. Test No. 236: Fish Embryo Acute Toxicity (FET) Test, OECD Guidelines for the

2   Testing of Chemicals, Section 2, OECD Publishing, Paris.

3 Oekotoxzentrum, 2017. Proposals for Acute and Chronic Quality Standards,

4   http://www.ecotoxcentre.ch/expert-service/quality-standards/proposals-for-acute-and-

5   chronic-quality-standards/, accessed 26 February 2017.

6 Pillon, A., Boussioux, A.M., Escande, A., Aït-Aïssa, S., Gomez, E., Fenet, H., Ruff, M., Moras, D.,

7   Vignon, F., Duchesne, M.J., Casellas, C., Nicolas, J.C., Balaguer, P., 2005. Binding of

8   estrogenic compounds to recombinant estrogen receptor alpha: Application to environmental

9   analysis. Environ. Health Perspect. 113(3), 278-284.

10 Prasse, C., Stalter, D., Schulte-Oehlmann, U., Oehlmann, J., 2015. Spoilt for choice: A critical

11   review on the chemical and biological assessment of current wastewater treatment

12   technologies. Water Res. 87, 237-270.

13 Reifferscheid, G., Maes, H.M., Allner, B., Badurova, J., Belkin, S., Bluhm, K., Brauer, F.,

14   Bressling, J., Domeneghetti, S., Elad, T., Fluckiger-Isler, S., Grummt, H.J., Guertler, R.,

15   Hecht, A., Heringa, M.B., Hollert, H., Huber, S., Kramer, M., Magdeburg, A., Ratte, H.T.,

16   Sauerborn-Klobucar, R., Sokolowski, A., Soldan, P., Smital, T., Stalter, D., Venier, P.,

17   Ziemann, C., Zipperle, J., Buchinger, S., 2012. International round-robin study on the Ames

18   fluctuation test. Environ. Mol. Mutagen. 53(3), 185-197.

19 Schmidt, S., Busch, W., Altenburger, R., Küster, E., 2016. Mixture toxicity of water contaminants-

20   effect analysis using the zebrafish embryo assay (*Danio rerio*). Chemosphere 152, 503-512.

21 Scholz, S., Kurauchi, K., Kinoshita, M., Oshima, Y., Ozato, K., Schirmer, K., Wakamatsu, Y.,

22   2005. Analysis of estrogenic effects by quantification of green fluorescent protein in

23   juvenile fish of a transgenic medaka. Environ. Toxicol. Chem. 24(10), 2553-2561.

24 Schroeder, A.L., Ankley, G.T., Houck, K.A., Villeneuve, D.L., 2016. Environmental surveillance

25   and monitoring—The next frontiers for high-throughput toxicology. Environ. Toxicol.

26   Chem. 35(3), 513-525.

1   Silva, E., Rajapakse, N., Kortenkamp, A., 2002. Something from "nothing" - Eight weak estrogenic
2   chemicals combined at concentrations below NOECs produce significant mixture effects.
3   Environ. Sci. Technol. 36(8), 1751-1756.

4   Simmons, S.O., Fan, C.-Y., Ramabhadran, R., 2009. Cellular Stress Response Pathway System as a
5   Sentinel Ensemble in Toxicological Screening. Toxicol. Sci. 111(2), 202-225.

6   Sonavane, M., Creusot, N., Maillot-Maréchal, E., Péry, A., Brion, F., Aït-Aïssa, S., 2016.
7   Zebrafish-based reporter gene assays reveal different estrogenic activities in river waters
8   compared to a conventional human-derived assay. Sci. Total Environ. 550, 934-939.

9   Spirhanzlova, P., Leleu, M., Sebillot, A., Lemkine, G.F., Iguchi, T., Demeneix, B.A., Tindall, A.J.,
10   2016. Oestrogen reporter transgenic medaka for non-invasive evaluation of aromatase
11   activity. Comp. Biochem. Physiol. C Toxicol. Pharmacol. 179, 64-71.

12   Spycher, S., Smejtek, P., Netzeva, T.I., Escher, B.I., 2008. Toward a class-independent quantitative
13   structure - Activity relationship model for uncouplers of oxidative phosphorylation. Chem.
14   Res. Toxicol. 21(4), 911-927.

15   Tang, J.Y.M., McCarty, S., Glenn, E., Neale, P.A., Warne, M.S.J., Escher, B.I., 2013. Mixture
16   effects of organic micropollutants present in water: Towards the development of effect-
17   based water quality trigger values for baseline toxicity. Water Res. 47(10), 3300-3314.

18   Tang, J.Y.M., Busetti, F., Charrois, J.W.A., Escher, B.I., 2014. Which chemicals drive biological
19   effects in wastewater and recycled water? Water Res. 60, 289-299.

20   Tice, R.R., Austin, C.P., Kavlock, R.J., Bucher, J.R., 2013. Improving the human hazard
21   characterization of chemicals: A Tox21 update. Environ. Health Perspect. 121(7), 756-765.

22   Umbuzeiro, G.D.A., Heringa, M., Zeiger, E., 2016. *In vitro* genotoxicity testing: Significance and
23   use in environmental monitoring. Adv. Biochem. Eng. Biotechnol., 1-22.

24   US EPA, 2008. Estimation Programs Interface Suite for Windows, v4.1., United States
25   Environmental Protection Agency, Washington, DC, USA.

US EPA, 2015. Interactive Chemical Safety for Sustainability (iCSS) Dashboard v2, http://actor.epa.gov/dashboard/, Accessed 5th Jan 2016.

Vaal, M.A., Van Leeuwen, C.J., Hoekstra, J.A., Hermens, J.L.M., 2000. Variation in sensitivity of aquatic species to toxicants: Practical consequences for effect assessment of chemical substances. Environ. Manage. 25(4), 415-423.

Van der Linden, S.C., Heringa, M.B., Man, H.Y., Sonneveld, E., Puijker, L.M., Brouwer, A., Van der Burg, B., 2008. Detection of multiple hormonal activities in wastewater effluents and surface water, using a panel of steroid receptor CALUX bioassays. Environ. Sci. Technol. 42(15), 5814-5820.

Vergauwen, L., Schmidt, S.N., Stinckens, E., Maho, W., Blust, R., Mayer, P., Covaci, A., Knapen, D., 2015. A high throughput passive dosing format for the Fish Embryo Acute Toxicity test. Chemosphere 139, 9-17.

Verhaar, H.J.M., Vanleeuwen, C.J., Hermens, J.L.M., 1992. Classifying environmental pollutants 1. Structure-activity-relationships for prediction of aquatic toxicity. Chemosphere 25(4), 471-491.

Villeneuve, D.L., Blankenship, A.L., Giesy, J.P., 2000. Derivation and application of relative potency estimates based on in vitro bioassay results. Environ. Toxicol. Chem. 19(11), 2835-2843.

Vörösmarty, C.J., McIntyre, P.B., Gessner, M.O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S.E., Sullivan, C.A., Liermann, C.R., Davies, P.M., 2010. Global threats to human water security and river biodiversity. Nature 467(7315), 555-561.

Wernersson, A.-S., Carere, M., Maggi, C., Tusil, P., Soldan, P., James, A., Sanchez, W., Dulio, V., Broeg, K., Reifferscheid, G., Buchinger, S., Maas, H., Van Der Grinten, E., O'Toole, S., Ausili, A., Manfra, L., Marziali, L., Polesello, S., Lacchetti, I., Mancini, L., Lilja, K., Linderoth, M., Lundeberg, T., Fjällborg, B., Porsbring, T., Larsson, D.J., Bengtsson-Palme, J., Förlin, L., Kienle, C., Kunz, P., Vermeirssen, E., Werner, I., Robinson, C.D., Lyons, B.,

1   Katsiadaki, I., Whalley, C., den Haan, K., Messiaen, M., Clayton, H., Lettieri, T., Carvalho,
2   R.N., Gawlik, B.M., Hollert, H., Di Paolo, C., Brack, W., Kammann, U., Kase, R., 2015.
3   The European technical report on aquatic effect-based monitoring tools under the water
4   framework directive. Env. Sci. Eur. 27(7), 1-11.
5   Wilson, V.S., Bobseine, K., Lambright, C.R., Gray, L.E., 2002. A novel cell line, MDA-kb2, that
6   stably expresses an androgen- and glucocorticoid-responsive reporter for the detection of
7   hormone receptor agonists and antagonists. Toxicol. Sci. 66(1), 69-81.
8   Wink, S., Hiemstra, S., Huppelschoten, S., Danen, E., Niemeijer, M., Hendriks, G., Vrieling, H.,
9   Herpers, B., van de Water, B., 2014. Quantitative High Content Imaging of Cellular
10   Adaptive Stress Response Pathways in Toxicity for Chemical Safety Assessment. Chem.
11   Res. Toxicol. 27(3), 338-355.
12   Wink, S., Hiemstra, S., Herpers, B., van de Water, B., 2017. High-content imaging-based BAC-
13   GFP toxicity pathway reporters to assess chemical adversity liabilities. Archives of
14   Toxicology 91(3), 1367-1383.
15   Yamarik, T.A., Andersen, F.A., Cosmetic Ingredient Review, E., 2004. Safety assessment of
16   dichlorophene and chlorophene. Int. J. Toxicol. 23, 1-27.
17   Yeh, R.Y.L., Farre, M.J., Stalter, D., Tang, J.Y.M., Molendijk, J., Escher, B.I., 2014. Bioanalytical
18   and chemical evaluation of disinfection by-products in swimming pool water. Water Res.
19   59, 172-184.
20

21 **Table 1**: The applied assays with the number of chemicals tested per assay and the number of chemicals active in each assay.

Assay ID	Assay name	Measured endpoint or molecular target	Method reference	No. of tested chemicals	No. of active chemicals
1	AhR CALUX	Activation of aryl hydrocarbon receptor (AhR)	Brennan et al. (2015)	18	7
2	HG5LN-hPXR	Activation of pregnane x receptor (PXR)	Lemaire et al. (2006)	34	20
3	PPARγ-bla	Binding of chemicals to peroxisome proliferator-activated receptor gamma (PPARγ)	Invitrogen (2010)	18	4
4	MDA-kb2	Activation of androgen receptor (AR)	Wilson et al. (2002)	24	3
5	Anti-MDA-kb2	Inhibition of androgen receptor (AR)	Wilson et al. (2002)	25	12
6	MELN	Activation of estrogen receptor (ER)	Balaguer et al. (1999)	34	9
7	ZELH-zfERalpha	Activation of estrogen receptor (ER)	Cosnefroy et al. (2012)	34	2
8	ZELH-zfERbeta2	Activation of estrogen receptor (ER)	Cosnefroy et al. (2012)	34	2
9	GR CALUX	Activation of glucocorticoid receptor (GR)	Van der Linden et al. (2008)	20	0
10	cyp19a1b-GFP	cyp19a1b gene expression	Brion et al. (2012)	26	2
11	ChgH-GFP	Estrogen receptor (ER) modulation, modulation of steroidogenesis	Spirhanzlova et al. (2016)	24	17
12	THbZIP-GFP (XETA)	Modulation of thyroid hormone signaling	Fini et al. (2007)	20	9
13	Ames fluctuation test	Mutagenicity (+/-S9)	Reifferscheid et al. (2012)	22	1

14	Ames microplate agar	Mutagenicity (+/-S9)	DeMarini et al. (1989), Mortelmans and Zeiger (2000)	12	1
15	AREc32	Induction of oxidative stress response	Escher et al. (2012)	18	9
16	Microtox	Inhibition of bioluminescence	(Escher et al. 2017)	34	16
17	Algal growth inhibition	Growth inhibition	OECD (2011)	17	12
18	*Daphnia* immobilization test	Immobilization	OECD (2004)	17	15
19	FET (96 well plate)	Mortality	OECD (2013)	20	13
20	FET (glass vial)	Mortality	OECD (2013)	29	17

Table 2: Summary of the selected test chemicals, their mode of action, proposed annual average concentration environmental quality standard (AA-EQS), maximum hazard quotient in Busch et al. (2016) and chemical score given in this study.

Chemical ID	Chemical	CAS No.	Molecular Weight (g/mol)	Chemical use group	Mode of action[a]	Annual Average Environmental Quality Standards (µg/L)	Maximum hazard quotient† (Busch et al. 2016)	Chemical score[*]
1	1,2-Benzisothiazolinone	2634-33-5	151.18	Biocide	Skin sensitization (soft electrophilic reactive toxicity)[b]		$7.88 \times 10^{-4\text{g}}$	6
2	2,4-Dichlorophenoxyacetic acid	94-75-7	221.04	Herbicide	Synthetic auxin	0.2^{e}	1.78×10^{-3}	4
3	4-Nonylphenol	104-40-5	220.35	Industrial Chemical	Endocrine disruption	0.043^{e}; 0.3^{f}	2.14×10^{-2}	5
4	Benzo(a)pyrene	50-32-8	252.32	Combustion by-product	Nucleic acid damage	$1.7 \times 10^{-4\text{f}}$	4.88×10^{-3}	4
5	Benzo(b)fluoranthene	205-99-2	252.31	Combustion	Nucleic acid damage		3.34×10^{-3}	4

					by-product			
6	Bezafibrate	41859-67-0	361.83	Pharmaceutical	Lipid metabolism	2.3^e	4.15×10^{-2}	4
7	Bisphenol A	80-05-7	228.29	Industrial Chemical	Endocrine disruption	0.24^e	2.10×10^{-2}	5
8	Carbamazepine	298-46-4	236.28	Pharmaceutical	Ion channel modulation	2.0^e	1.48×10^{-3}	4
9	Carbendazim	10605-21-7	191.19	Fungicide	Mitosis, cell cycle	0.44^e	2.65×10^{-2}	4
10	Chlorophene	120-32-1	218.68	Industrial Chemical	Skin sensitization (soft electrophilic reactive toxicity)^c		4.19×10^{-3}	5
11	Chlorpyrifos	2921-88-2	350.59	Insecticide	Neuroactive	$4.6\ 10^{-4e}$	1.07×10^{-1}	6
12	Clofibric acid	882-09-7	214.65	Herbicide	Synthetic auxin		5.64×10^{-3}	5
13	Cyprodinil	121552-61-2	225.29	Fungicide	Protein biosynthesis inhibition	0.33^e	5.34×10^{-3}	4
14	Diazinon	333-41-5	304.35	Insecticide	Neuroactive	0.012^e	4.67×10^{0}	5
15	Diclofenac	15307-86-5	296.15	Pharmaceutical	Antiinflammatory	0.05^e	6.77×10^{-2}	5
16	Dinoseb	88-85-7	240.21	Herbicide	Respiration inhibition, uncoupler of		1.65×10^{-3}	5

					oxidative phosphorylation[d]			
17	Diuron	330-54-1	233.10	Herbicide	Photosynthesis inhibition	0.07^{e}; 0.2^{f}	7.06×10^{-1}	6
18	Fipronil	120068-37-3	437.15	Insecticide	Neuroactive		1.44×10^{-1}	4
19	Flutamide	13311-84-7	276.21	Pharmaceutical	Endocrine disruption		-	2
20	Genistein	446-72-0	270.24	Phytoestrogen	Mitosis, cell cycle		6.13×10^{-4}	4
21	Hexadecylpyridinium	7773-52-6	304.54	Industrial Chemical	Cell membrane disruption		1.24×10^{2}	2
22	Mefenamic acid	61-68-7	241.29	Pharmaceutical	Antiinflammatory	1.0^{e}	8.70×10^{-2}	5
23	Metolachlor	51218-45-2	283.80	Herbicide	Mitosis, cell cycle		3.23×10^{-2}	4
24	Metoprolol	37350-58-6	267.37	Pharmaceutical	Beta blocker	8.6^{e}	3.51×10^{-4}	3
25	Perfluorooctanoic acid	335-67-1	414.07	Industrial Chemical	Carcinogen, Endocrine disruption, lipid metabolism		1.63×10^{-3}	3
26	Picoxystrobin	117428-22-5	367.32	Fungicide	Respiration inhibition		-	3
27	Propiconazole	60207-90-1	342.23	Fungicide	Sterol biosynthesis inhibition		3.49×10^{-2}	4

28	Sulfamethoxazole	723-46-6	253.28	Antibiotic	Antibiotic	0.6^e	9.81×10^{-4}	4
29	Tebuconazole	107534-96-3	307.82	Fungicide	Sterol biosynthesis inhibition	0.24^e	2.40×10^{-3}	4
30	Telmisartan	144701-48-4	514.62	Pharmaceutical	Angiotensin receptor or enzyme		3.09×10^{2}	4
31	Terbutylazine	5915-41-3	229.71	Herbicide	Photosynthesis inhibition	0.22^e	3.86×10^{-1}	4
32	Triclocarban	101-20-2	315.59	Biocide	Lipid metabolism		1.64×10^{-2}	3
33	Triclosan	3380-34-5	289.55	Biocide	Lipid metabolism, uncoupler of oxidative phosphorylation[d]	0.02^e	4.22×10^{-1}	4
34	Triphenylphosphate	115-86-6	326.29	Industrial Chemical	Neuroactive		2.01×10^{-2}	4

[a]Busch et al. (2016); [b]Basketter et al. (1999); [c]Yamarik et al. (2004); [d]Spycher et al. (2008); [e]Proposed by Oekotoxzentrum (2017); [f]European Commission (2013).

†maximum hazard quotient in fish, daphnia or algae is presented. *Chemical score based on rank in the hazard quotient list, the availability of an EQS, a unique mode of action that is not covered by a higher ranked chemical, diversity of chemical use group, specific mode of action that is covered by the test battery, specific mode of action that leads to expectation of enhanced toxicity in the whole organism bioassays and not readily biodegradable based on BIOWIN. Further information is provided in Table S1.

29 **Table 3**: Effect concentrations EC (M) for the selected water pollutants in the studied assays. Lowest observable effect concentration (LOEC) reported

30 for the ChgH-GFP (spiked) and THbZIP-GFP (XETA) assays. SE=standard error, CI=95% confidence intervals.

Chemical	AhR CALUX (1)	HG5LN-hPXR (2)	PPARγ-bla (3)	MDA-kb2 (4)	Anti-MDA-kb2 (5)	MELN (6)	ZELH-zfERalpha (7)	ZELH-zfERbeta2 (8)
	$EC_{10} \pm SE$	$EC_{10} \pm SE$	$EC_{10} \pm SE$	$EC_{10} \pm SE$	$EC_{SR0.2} \pm SE$	$EC_{10} \pm SE$	$EC_{10} \pm SE$	$EC_{10} \pm SE$
1,2-Benzisothiazolinone	-	$>3.00\times10^{-5}$	-	-	-	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$
2,4-Dichlorophenoxyacetic acid	-	$>3.00\times10^{-5}$	-	$>2.00\times10^{-4}$	$>2.00\times10^{-4}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$
4-Nonylphenol	$(1.36\pm0.06)\times10^{-5}$	$>3.00\times10^{-5}$	$>1.07\times10^{-5}$	$>2.00\times10^{-4}$	$>6.67\times10^{-5}$	$(9.69\pm0.61)\times10^{-7}$	$>1.00\times10^{-5}$	$>1.00\times10^{-5}$
Benzo(a)pyrene	$(8.38\pm0.55)\times10^{-10}$	$>1.00\times10^{-5}$	$>4.61\times10^{-6}$	$(5.20\pm0.40)\times10^{-6}$	$(1.29\pm0.14)\times10^{-6}$	$(5.30\pm0.69)\times10^{-7}$	$>1.00\times10^{-5}$	$>1.00\times10^{-5}$
Benzo(b)fluoranthene	$(8.23\pm0.94)\times10^{-10}$	$>3.00\times10^{-5}$	$>3.96\times10^{-6}$	$(5.08\pm0.69)\times10^{-6}$	$>2.00\times10^{-4}$	$(7.26\pm1.35)\times10^{-7}$	$>1.00\times10^{-5}$	$>1.00\times10^{-5}$
Bezafibrate	-	$>3.00\times10^{-5}$	-	-	-	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$
Bisphenol A	$>4.47\times10^{-5}$	$(4.75\pm0.48)\times10^{-6}$	$>2.51\times10^{-5}$	$>6.67\times10^{-5}$	$(1.72\pm0.24)\times10^{-7}$	$(6.91\pm0.21)\times10^{-8}$	$(1.41\pm0.19)\times10^{-6}$	$(2.86\pm0.27)\times10^{-6}$
Carbamazepine	-	$(3.63\pm0.32)\times10^{-5}$	-	$>2.00\times10^{-4}$	$(5.37\pm1.48)\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$
Carbendazim	$>2.00\times10^{-5}$	$>3.00\times10^{-6}$	$>5.44\times10^{-6}$	$>2.00\times10^{-4}$	$>2.00\times10^{-4}$	$>3.00\times10^{-6}$	$>3.00\times10^{-6}$	$>3.00\times10^{-6}$
Chlorophene	$>9.67\times10^{-5}$	$(1.02\pm0.06)\times10^{-5}$	$>1.01\times10^{-5}$	$>2.00\times10^{-4}$	$(2.10\pm0.24)\times10^{-7}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$
Chlorpyrifos	$(1.25\pm0.10)\times10^{-6}$	$(5.11\pm0.34)\times10^{-6}$	$>5.84\times10^{-5}$	$>2.00\times10^{-4}$	$(1.42\pm0.16)\times10^{-5}$	$(8.91\pm0.41)\times10^{-6}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$
Clofibric acid	-	$>3.00\times10^{-5}$	-	-	-	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$
Cyprodinil	$(6.36\pm0.88)\times10^{-6}$	$(5.26\pm0.23)\times10^{-6}$	$>4.21\times10^{-5}$	$>2.00\times10^{-4}$	$(1.34\pm0.36)\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$
Diazinon	$(1.35\pm0.06)\times10^{-5}$	$(1.44\pm0.11)\times10^{-6}$	$(5.30\pm0.46)\times10^{-5}$	$>2.00\times10^{-4}$	$(8.77\pm0.58)\times10^{-6}$	$(8.53\pm0.34)\times10^{-6}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$

Diclofenac	$>1.37\times10^{-4}$	$(3.78\pm0.32)\times10^{-5}$	$(1.21\pm0.09)\times10^{-6}$	$>2.00\times10^{-4}$	$>6.67\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$
Dinoseb	-	$>3.00\times10^{-5}$	-	-	-	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$
Diuron	-	$(3.64\pm0.34)\times10^{-5}$	-	$>2.00\times10^{-4}$	$>5.00\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$
Fipronil	$>3.32\times10^{-6}$	$(6.35\pm0.78)\times10^{-6}$	$>1.64\times10^{-6}$	$>2.22\times10^{-5}$	$(1.83\pm0.24)\times10^{-6}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$
Flutamide	-	$(2.17\pm0.10)\times10^{-6}$	-	-	$(2.07\pm0.08)\times10^{-7}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$
Genistein	$>1.30\times10^{-5}$	$(1.24\pm0.13)\times10^{-6}$	$>1.41\times10^{-5}$	$(4.14\pm0.42)\times10^{-5}$	$>3.33\times10^{-5}$	$(1.22\pm0.23)\times10^{-8}$	$(9.71\pm0.71)\times10^{-7}$	$(2.54\pm0.37)\times10^{-8}$
Hexadecylpyridinium	-	$(1.00\pm0.11)\times10^{-6}$	-	$>1.00\times10^{-5}$	$>5.56\times10^{-7}$	$>1.00\times10^{-5}$	$>1.00\times10^{-5}$	$>1.00\times10^{-5}$
Mefenamic acid	-	$(1.08\pm0.06)\times10^{-5}$	-	$>2.00\times10^{-4}$	$>6.67\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$
Metolachlor	-	$(2.68\pm0.11)\times10^{-7}$	-	-	-	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$
Metoprolol	-	$>3.00\times10^{-5}$	-	$>2.00\times10^{-4}$	$>2.00\times10^{-4}$	$>3.00\times10^{-5}$	$>3.00\times10\text{-}5$	$>3.00\times10^{-5}$
Perfluorooctanoic acid	$>1.18\times10^{-4}$	$>3.00\times10^{-5}$	$>2.19\times10^{-5}$	$>2.00\times10^{-4}$	$>6.67\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$
Picoxystrobin	-	$(1.73\pm0.21)\times10^{-5}$	-	-	-	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$
Propiconazole	$>2.20\times10^{-5}$	$(2.94\pm0.18)\times10^{-6}$	$>2.23\times10^{-5}$	$>2.00\times10^{-4}$	$(2.87\pm0.39)\times10^{-6}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$
Sulfamethoxazole	-	$>3.00\times10^{-5}$	-	-	-	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$
Tebuconazole	-	$(1.88\pm0.27)\times10^{-6}$	-	-	-	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$
Telmisartan	$(1.57\pm0.08)\times10^{-5}$	$>5.00\times10^{-6}$	$(1.43\pm0.10)\times10^{-7}$	$>2.00\times10^{-4}$	$>5.00\times10^{-5}$	$>5.00\times10^{-6}$	$>5.00\times10^{-6}$	$>5.00\times10^{-6}$
Terbutylazine	-	$(1.03\pm0.06)\times10^{-5}$	-	-	-	$(1.52\pm0.06)\times10^{-5}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$
Triclocarban	$>4.80\times10^{-6}$	$>3.00\times10^{-5}$	$>3.31\times10^{-7}$	$>3.33\times10^{-6}$	$>2.22\times10^{-6}$	$>1.00\times10^{-5}$	$>1.00\times10\text{-}5$	$>1.00\times10^{-5}$
Triclosan	$>8.50\times10^{-5}$	$(1.77\pm0.16)\times10^{-6}$	$>2.05\times10^{-6}$	$>7.41\times10^{-6}$	$(6.52\pm0.68)\times10^{-7}$	$>3.00\times10^{-5}$	$>1.00\times10\text{-}5$	$>1.00\times10^{-5}$
Triphenylphosphate	$>1.00\times10^{-6}$	$(9.10\pm0.26)\times10^{-7}$	$(4.24\pm0.34)\times10^{-6}$	$>2.00\times10^{-4}$	$(6.25\pm1.13)\times10^{-6}$	$(1.71\pm0.09)\times10^{-6}$	$>3.00\times10^{-5}$	$>3.00\times10^{-5}$

Chemical	GR CALUX (9) EC$_{10}$ ± SE	cyp19a1b-GFP (10) EC$_{10}$ ± SE (M)	ChgH-GFP (11) EC$_{10}$ (unspiked) (M)	ChgH-GFP (11) LOEC (spiked) (M)†	THbZIP-GFP (XETA) (12) LOEC (unspiked) (M)†	THbZIP-GFP (XETA) (12) LOEC (spiked) (M)†	Ames fluctuation test (13) EC$_{10}$ ± SE	Ames microplate agar (14) EC$_{IR1.5}$ ± SE (M)
1,2-Benzisothiazolinone	-	-	-	-	-	-	-	-
2,4-Dichlorophenoxyacetic acid	>4.52×10^{-4}	>3.00×10^{-5}	>4.52×10^{-4}	3.39×10^{-4}[b]	>4.52×10^{-4}	>4.52×10^{-4}	-	-
4-Nonylphenol	>2.84×10^{-5}	>2.50×10^{-6}	>2.27×10^{-6}	2.27×10^{-6}[b]	-	-	-	-
Benzo(a)pyrene	>9.91×10^{-5}	>3.00×10^{-6}	-	-	-	-	(1.70±0.11)×10^{-7}[c]	(6.90±0.79)×10^{-6}[c]
Benzo(b)fluoranthene	-	>3.00×10^{-6}	>3.96×10^{-8}	>3.96×10^{-8}	>3.96×10^{-8}	>3.96×10^{-8}	>9.91×10^{-8}	>3.96×10^{-3}
Bezafibrate	>2.76×10^{-4}	-	-	-	-	-	-	-
Bisphenol A	>1.10×10^{-4}	(1.58±0.09)×10^{-6}	(1.24±0.21)×10^{-5}	2.19×10^{-6}[a]	1.64×10^{-5}[a]	1.64×10^{-5}[a]	>1.10×10^{-5}	>1.31×10^{-2}
Carbamazepine	-	>3.00×10^{-5}	>1.27×10^{-4}	1.27×10^{-4}[a]	8.56×10^{-5}[a]	>1.27×10^{-4}	>4.23×10^{-4}	-
Carbendazim	>2.64×10^{-4}	>3.00×10^{-5}	>2.62×10^{-6}	>2.62×10^{-6}	>2.62×10^{-6}	>2.62×10^{-6}	-	-
Chlorophene	>5.72×10^{-5}	>3.00×10^{-6}	(5.27±0.70)×10^{-6}	4.57×10^{-6}[a]	>9.14×10^{-5}	>9.14×10^{-5}	>1.43×10^{-6}	>1.37×10^{-4}
Chlorpyrifos	>3.57×10^{-5}	>3.00×10^{-6}	>2.85×10^{-6}	2.85×10^{-6}[a]	1.43×10^{-6}[a]	7.13×10^{-6}[a]	>2.85×10^{-5}	-
Clofibric acid	-	-	-	-	-	-	-	-
Cyprodinil	>4.44×10^{-4}	>1.00×10^{-5}	>8.88×10^{-6}	4.44×10^{-6}[a]	-	-	>1.61×10^{-5}	>4.44×10^{-3}

Diazinon	$>1.64\times10^{-5}$	$>1.00\times10^{-4}$	$>6.16\times10^{-5}$	8.21×10^{-6a}	4.11×10^{-5a}	$>6.16\times10^{-5}$	$>1.64\times10^{-5}$	$>3.29\times10^{-2}$
Diclofenac	$>3.38\times10^{-5}$	$>1.00\times10^{-6}$	$>1.57\times10^{-4}$	1.57×10^{-4b}	$>1.59\times10^{-4}$	7.86×10^{-5a}	$>3.38\times10^{-5}$	$>3.38\times10^{-2}$
Dinoseb	-	-	-	-	-	-	$>4.16\times10^{-7}$	-
Diuron	$>2.15\times10^{-4}$	$>3.00\times10^{-5}$	$>8.58\times10^{-5}$	4.29×10^{-5a}	$>8.58\times10^{-5}$	$>8.58\times10^{-5}$	$>4.29\times10^{-5}$	$>4.29\times10^{-2}$
Fipronil	-	$>3.00\times10^{-5}$	$>2.29\times10^{-7}$	2.29×10^{-7b}	$>2.29\times10^{-6}$	$>2.29\times10^{-6}$	$>2.29\times10^{-7}$	-
Flutamide	-	-	-	-	-	-	$>3.62\times10^{-7}$	-
Genistein	$>2.31\times10^{-5}$	$(4.20\pm0.24)\times10^{-7}$	$>3.70\times10^{-6}$	$>3.70\times10^{-6}$	$>3.70\times10^{-6}$	$>3.70\times10^{-6}$	$>4.63\times10^{-6}$	$>1.11\times10^{-3}$
Hexadecylpyridinium	-	$>3.00\times10^{-7}$	$>2.79\times10^{-8}$	$>2.79\times10^{-8}$	$>2.79\times10^{-8}$	$>2.79\times10^{-8}$	-	-
Mefenamic acid	-	$>3.00\times10^{-6}$	$>4.14\times10^{-5}$	3.11×10^{-5a}	$>2.07\times10^{-6}$	2.07×10^{-6a}	-	-
Metolachlor	-	$>3.00\times10^{-5}$	3.52×10^{-5}	$>3.52\times10^{-5}$	-	-	$>2.22\times10^{-8}$	-
Metoprolol	$>3.74\times10^{-4}$	$>3.00\times10^{-5}$	$>1.02\times10^{-4}$	$>1.02\times10^{-4}$	5.11×10^{-5a}	5.11×10^{-5a}	$>3.74\times10^{-5}$	-
Perfluorooctanoic acid	$>1.12\times10^{-4}$	$>3.00\times10^{-5}$	$>2.42\times10^{-4}$	1.81×10^{-4a}	2.42×10^{-4a}	$>2.42\times10^{-4}$	-	-
Picoxystrobin	-	-	-	-	-	-	-	-
Propiconazole	$>7.31\times10^{-5}$	$>3.00\times10^{-5}$	$>2.92\times10^{-6}$	2.92×10^{-7a}	$>2.92\times10^{-6}$	$>2.92\times10^{-6}$	$>7.31\times10^{-6}$	$>8.77\times10^{-3}$
Sulfamethoxazole	-	-	-	-	-	-	$>3.95\times10^{-7}$	-
Tebuconazole	-	-	-	-	-	-	$>3.25\times10^{-7}$	-
Telmisartan	-	$>1.00\times10^{-6}$	$>9.72\times10^{-6}$	$>9.72\times10^{-6}$	-	-	-	-
Terbutylazine	$>4.35\times10^{-4}$	$>3.00\times10^{-5}$	-	-	-	-	$>4.37\times10^{-7}$	-
Triclocarban	$>3.17\times10^{-4}$	$>3.00\times10^{-7}$	$>3.17\times10^{-6}$	1.58×10^{-6a}	$>1.59\times10^{-7}$	$>1.59\times10^{-7}$	-	-
Triclosan	$>2.16\times10^{-5}$	$>1.00\times10^{-6}$	$>1.73\times10^{-6}$	$>1.73\times10^{-6}$	$>1.73\times10^{-6}$	$>1.73\times10^{-6}$	$>1.08\times10^{-6}$	$>1.04\times10^{-3}$
Triphenylphosphate	$>1.92\times10^{-5}$	$>3.00\times10^{-5}$	$(2.08\pm0.20)\times10^{-6}$	3.06×10^{-6a}	6.13×10^{-6a}	$>6.13\times10^{-6}$	$>1.92\times10^{-6}$	$>3.06\times10^{-2}$

31   [a]p-value <0.01; [b]p-value <0.05 [c]benzo(a)pyrene only active with S9

Chemical	AREc32 (15)	Microtox (16) Escher et al. (2017)	Algae growth inhibition (17)	Daphnia immobilization test (18)	FET (96 well plate) (19)	FET (Glass vial) (20)
	$EC_{IR1.5} \pm SE$	EC_{50} (95% CI)	EC_{50} (95% CI)	EC_{50} (95% CI)	EC_{50} (95% CI)	EC_{50} (95% CI)
1,2-Benzisothiazolinone	-	$1.62(1.44 \text{ to } 1.83) \times 10^{-5}$	-	-	-	$2.14(1.94 \text{ to } 2.37) \times 10^{-5}$
2,4-Dichlorophenoxyacetic acid	-	$2.34(2.03 \text{ to } 2.69) \times 10^{-3}$	-	-	-	$>7.71 \times 10^{-4}$
4-Nonylphenol	$(6.73 \pm 0.73) \times 10^{-5}$	$>2.50 \times 10^{-5*}$	$6.78(6.48 \text{ to } 7.08) \times 10^{-6}$	$7.21(6.80 \text{ to } 7.64) \times 10^{-8}$	$2.71(2.10 \text{ to } 3.51) \times 10^{-5}$	$5.09(4.54 \text{ to } 5.69) \times 10^{-6}$
Benzo(a)pyrene	$(2.22 \pm 0.10) \times 10^{-7}$	$>1.90 \times 10^{-5}$	$>3.00 \times 10^{-4}$	$5.12(4.46 \text{ to } 5.87) \times 10^{-9}$	$>2.48 \times 10^{-5}$	$>3.94 \times 10^{-5}$
Benzo(b)fluoranthene	$(2.33 \pm 0.16) \times 10^{-7}$	$>1.78 \times 10^{-5}$	$>3.00 \times 10^{-5}$	$>1.19 \times 10^{-5}$	$>2.48 \times 10^{-5}$	$>1.92 \times 10^{-5}$
Bezafibrate	-	$2.56(2.35 \text{ to } 2.79) \times 10^{-3}$	-	-	$>2.76 \times 10^{-4}$	$>3.21 \times 10^{-5}$
Bisphenol A	$(1.24 \pm 0.11) \times 10^{-4}$	$6.08(5.66 \text{ to } 6.52) \times 10^{-5}$	$1.12(1.03 \text{ to } 1.21) \times 10^{-4}$	$1.86(1.02 \text{ to } 3.40) \times 10^{-5}$	$6.67(5.94 \text{ to } 7.50) \times 10^{-5}$	$7.67(7.15 \text{ to } 8.24) \times 10^{-5}$
Carbamazepine	-	$7.41(6.29 \text{ to } 8.72) \times 10^{-4}$	-	-	$1.59(1.13 \text{ to } 2.25) \times 10^{-4}$	$>4.23 \times 10^{-4}$
Carbendazim	$>1.24 \times 10^{-4}$	$>1.20 \times 10^{-2}$	-	-	-	$1.94(1.50 \text{ to } 2.50) \times 10^{-6}$
Chlorophene	$(7.49 \pm 0.71) \times 10^{-5}$	$2.42(2.13 \text{ to } 2.75) \times 10^{-5}$	$1.26(1.18 \text{ to } 1.36) \times 10^{-5}$	$2.21(1.93 \text{ to } 2.53) \times 10^{-6}$	$2.07(1.74 \text{ to } 2.46) \times 10^{-5}$	$1.13(1.05 \text{ to } 1.23) \times 10^{-5}$
Chlorpyrifos	$>1.31 \times 10^{-4}$	$>2.82 \times 10^{-4}$	$1.85(1.76 \text{ to } 1.95) \times 10^{-5}$	$2.42(1.95 \text{ to } 3.00) \times 10^{-10}$	$>2.84 \times 10^{-4}$	$>2.45 \times 10^{-5}$
Clofibric acid	-	$2.07(1.98 \text{ to } 2.16) \times 10^{-3}$	-	-	-	$>2.39 \times 10^{-3}$
Cyprodinil	$(2.75 \pm 0.11) \times 10^{-5}$	$>4.57 \times 10^{-5*}$	$2.71(2.48 \text{ to } 2.97) \times 10^{-5}$	$1.49(1.10 \text{ to } 2.02) \times 10^{-7}$	$5.61(4.76 \text{ to } 6.61) \times 10^{-6}$	$2.78(2.56 \text{ to } 3.03) \times 10^{-6}$
Diazinon	$>1.30 \times 10^{-4}$	$>1.12 \times 10^{-3}$	$2.51(2.36 \text{ to } 2.66) \times 10^{-4}$	$4.86(4.29 \text{ to } 5.49) \times 10^{-10}$	$2.36(2.24 \text{ to } 2.50) \times 10^{-5}$	$1.84(1.49 \text{ to } 2.27) \times 10^{-5}$
Diclofenac	$>1.00 \times 10^{-3}$	$1.28(1.23 \text{ to } 1.33) \times 10^{-4}$	$8.66(7.56 \text{ to } 9.92) \times 10^{-4}$	$0.98(0.87 \text{ to } 1.10) \times 10^{-4}$	$4.19(3.68 \text{ to } 4.77) \times 10^{-5}$	$4.33(3.15 \text{ to } 5.96) \times 10^{-5}$

Dinoseb	-	$1.43(1.36 \text{ to } 1.50) \times 10^{-5}$	-	-	-	-
Diuron	-	$>1.54 \times 10^{-4}*$	$3.44(3.04 \text{ to } 3.89) \times 10^{-7}$	$0.98(0.91 \text{ to } 1.07) \times 10^{-4}$	$4.43(3.58 \text{ to } 5.48) \times 10^{-5}$	$1.30(1.19 \text{ to } 1.43) \times 10^{-5}$
Fipronil	$>4.00 \times 10^{-4}$	$>3.47 \times 10^{-6}*$	$>3.00 \times 10^{-5}$	$0.96(0.62 \text{ to } 1.49) \times 10^{-7}$	$>2.86 \times 10^{-5}$	$>4.61 \times 10^{-5}$
Flutamide	-	$>4.67 \times 10^{-5}*$	-	-	$1.98(1.62 \text{ to } 2.44) \times 10^{-5}$	-
Genistein	$(7.15 \pm 0.46) \times 10^{-5}$	$>6.61 \times 10^{-4}$	$>3.00 \times 10^{-5}$	$>3.70 \times 10^{-5}$	$1.10(0.72 \text{ to } 1.68) \times 10^{-5}$	$1.91(1.46 \text{ to } 2.50) \times 10^{-5}$
Hexadecylpyridinium	-	$8.34(7.72 \text{ to } 9.63) \times 10^{-6}$	-	-	-	-
Mefenamic acid	-	$3.55(3.29 \text{ to } 3.38) \times 10^{-4}$	-	-	-	$0.92(0.61 \text{ to } 1.39) \times 10^{-5}$
Metolachlor	-	$7.33(6.65 \text{ to } 8.09) \times 10^{-4}$	-	-	-	$2.51(1.56 \text{ to } 4.05) \times 10^{-4}$
Metoprolol	-	$>9.12 \times 10^{-3}$	$2.51(2.25 \text{ to } 2.81) \times 10^{-3}$	$1.54(1.29 \text{ to } 1.83) \times 10^{-3}$	$>3.74 \times 10^{-4}$	$>2.92 \times 10^{-3}$
Perfluorooctanoic acid	$>3.79 \times 10^{-4}$	$4.45(4.23 \text{ to } 4.69) \times 10^{-3}$	-	-	-	$2.16(1.98 \text{ to } 2.35) \times 10^{-3}$
Picoxystrobin	-	$>1.05 \times 10^{-3}$	-	-	-	-
Propiconazole	$>4.59 \times 10^{-5}$	$>6.92 \times 10^{-4}$	$4.00(3.48 \text{ to } 4.18) \times 10^{-5}$	$1.09(0.74 \text{ to } 1.59) \times 10^{-5}$	$6.48(5.03 \text{ to } 8.35) \times 10^{-5}$	$5.09(4.88 \text{ to } 5.31) \times 10^{-5}$
Sulfamethoxazole	-	$1.48(1.40 \text{ to } 1.56) \times 10^{-3}$	-	-	-	-
Tebuconazole	-	$2.67(2.49 \text{ to } 2.87) \times 10^{-4}$	-	-	-	$4.84(3.75 \text{ to } 6.25) \times 10^{-5}$
Telmisartan	$>4.04 \times 10^{-5}$	$>4.79 \times 10^{-3}$	-	-	-	$>8.01 \times 10^{-6}$
Terbutylazine	-	$>1.78 \times 10^{-3}$	-	-	-	$>2.46 \times 10^{-5}$
Triclocarban	$(4.81 \pm 0.40) \times 10^{-6}$	$>6.31 \times 10^{-4}$	$>3.00 \times 10^{-5}$	$4.06(3.20 \text{ to } 5.15) \times 10^{-8}$	$>1.98 \times 10^{-5}$	$>4.18 \times 10^{-6}$
Triclosan	$(1.71 \pm 0.14) \times 10^{-5}$	$3.02(2.90 \text{ to } 3.15) \times 10^{-5}$	$1.01(0.96 \text{ to } 1.07) \times 10^{-6}$	$1.05(0.86 \text{ to } 1.30) \times 10^{-6}$	$4.66(3.79 \text{ to } 5.72) \times 10^{-6}$	$1.47(1.25 \text{ to } 1.73) \times 10^{-6}$
Triphenylphosphate	$>3.81 \times 10^{-4}$	$>1.82 \times 10^{-3}$	$3.57(3.36 \text{ to } 3.80) \times 10^{-5}$	$1.81(1.28 \text{ to } 2.57) \times 10^{-6}$	$3.16(1.88 \text{ to } 5.31) \times 10^{-5}$	$6.30(5.58 \text{ to } 7.11) \times 10^{-6}$

33   *EC_{50} > water solubility C_w^{sat} (salt) (Escher et al. 2017).

34   **Figure 1:** Design of the panel of bioassays/biological endpoints (ovals) recommended for water

35   quality assessment and where they are situated along the adverse outcome pathway (boxes). The

36   cyp19a1b-GFP and ChgH-GFP assays were also included in the study, but are not shown in the

37   Figure.

38

39   **Figure 2:** A) Summary of experimental results of this study with the selected representative water

40   pollutants. Active chemicals are grouped according to their effect concentration (EC) with yellow

41   indicating least potent and red indicating most potent, inactive chemicals are shown in green, and

42   grey indicates no experimental data (all EC values can be found in Table 3) and B) comparison

43   between the experimental results from the current study and the expected activity based on the data
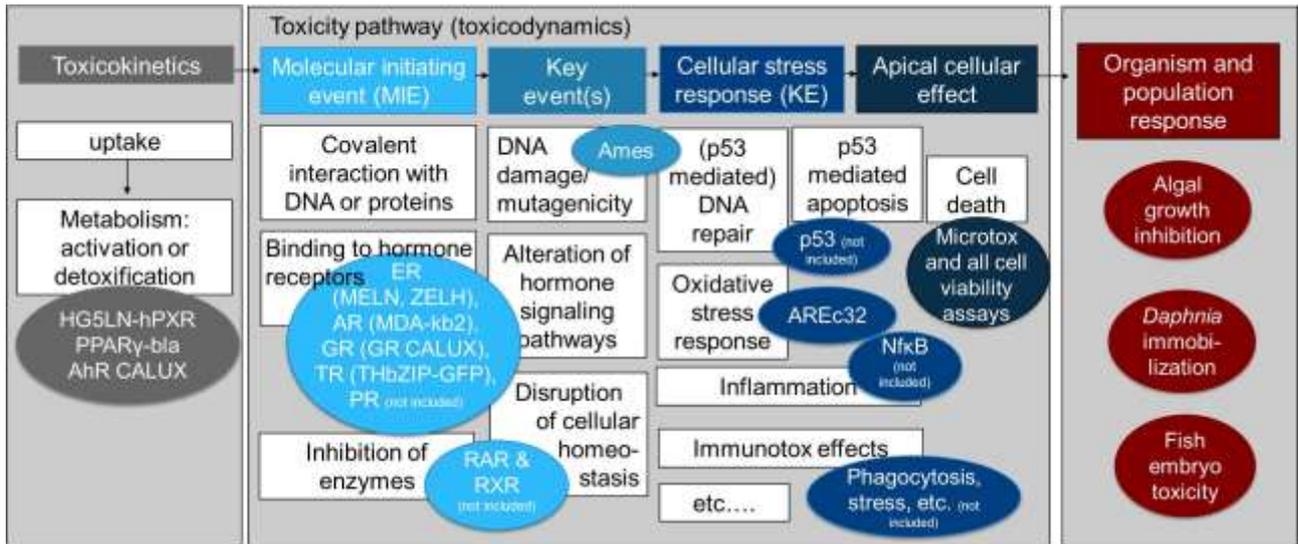
44   mining exercise.

45

46   **Figure 3:** Experimental $\log(1/EC_{50})$ values versus QSAR-predicted $\log(1/EC_{50})$ baseline toxicity

47   values for A) Microtox, B) algal growth inhibition, C) *Daphnia* immobilization test and D) 48 h

48   FET. Note 48 h FET experimental $\log(1/EC_{50})$ values are the average of two assays. TR=toxic ratio.

49

50   **Figure 4**: Percent effect elucidated by enriched water samples from different sites of the Danube

51   River explained by detected chemicals for HG5LN-hPXR based on A) literature EC values and B)

52   EC values measured in the present study, MELN based on C) literature EC values and D) EC values

53   measured in the present study and FET based on E) literature EC values and F) EC values measured

54   in the present study (Figures 4A, C and E are reprinted with permission from Neale et al. (2015).

55   Linking *in vitro* effects and detected organic micropollutants in surface water using mixture-toxicity

56   modeling. Environmental Science & Technology, 49(24): 14614-14624. Copyright 2015 American
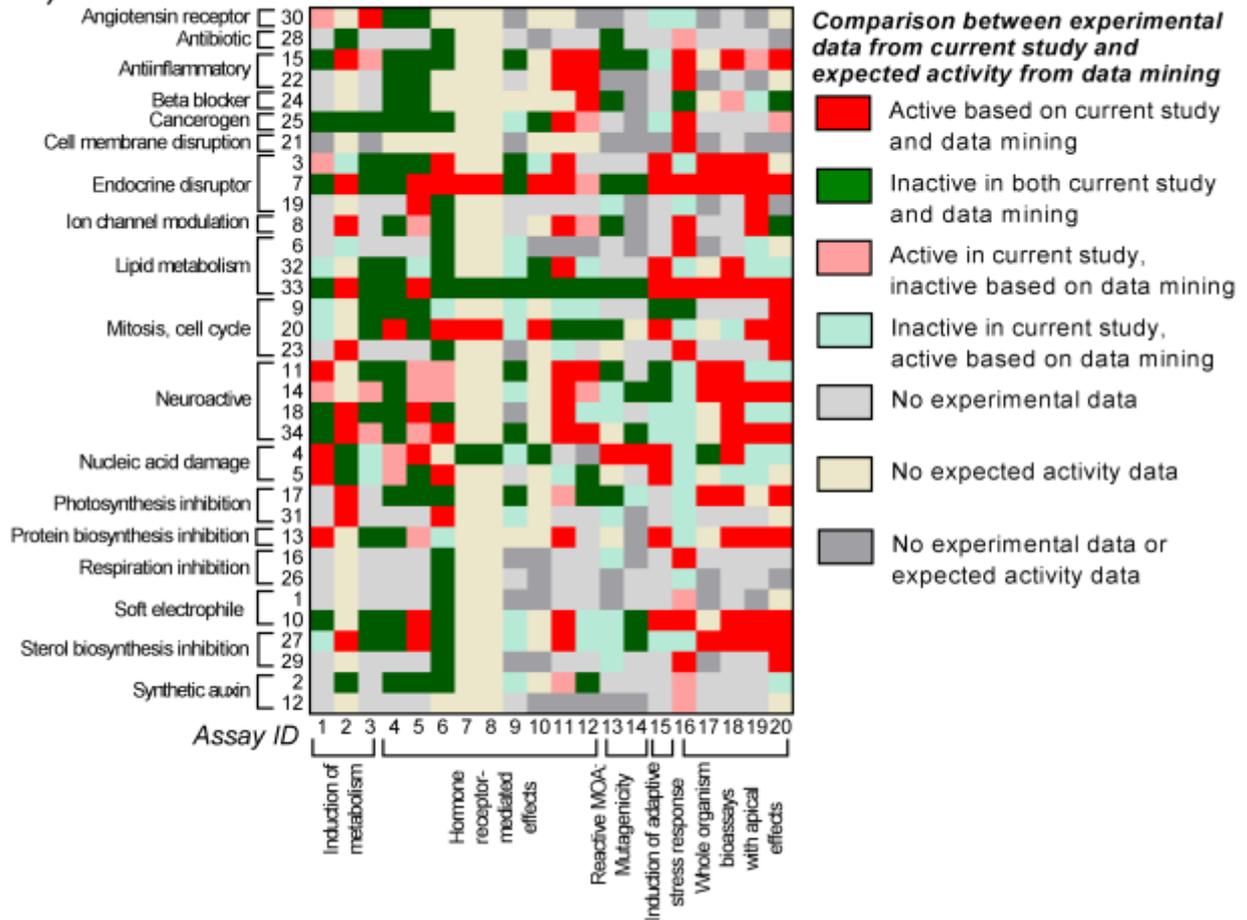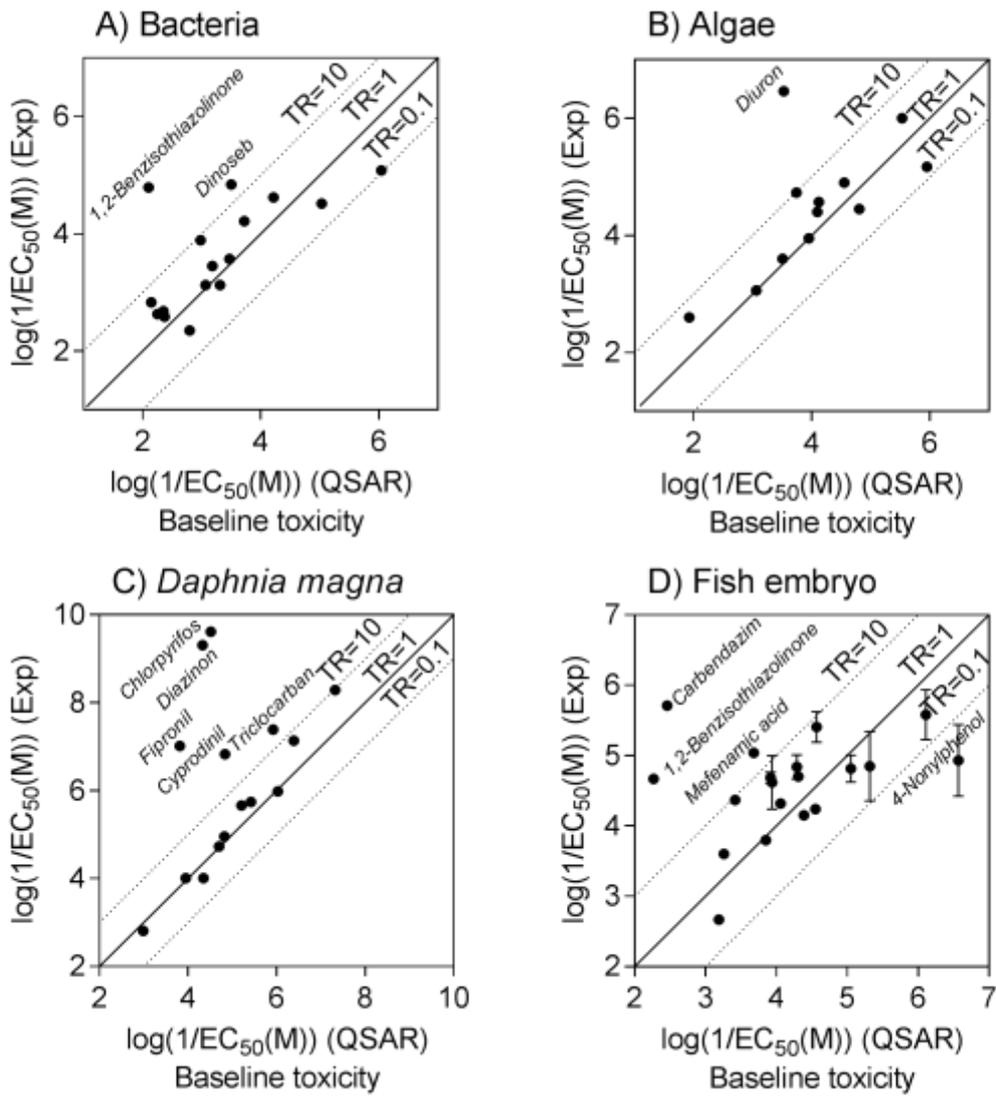
57   Chemical Society).

58

59 **Figure 1**

Figure 2

B) Mode of action   Chemical ID

Comparison between experimental data from current study and expected activity from data mining

Active based on current study and data mining

Inactive in both current study and data mining

Active in current study, inactive based on data mining

Inactive in current study, active based on data mining

No experimental data

No expected activity data

No experimental data or expected activity data

64

65

66

57

Figure 3

Figure 4