**An Integrated Approach for Identifying Priority Contaminant in the Great Lakes Basin –
Investigations in the Lower Green Bay/Fox River and Milwaukee Estuary Areas of
Concern**

Shibin Li^{12*}, Daniel L. Villeneuve¹, Jason P. Berninger^{12A}, Brett R. Blackwell¹, Jenna E. Cavallin¹, Megan N. Hughes¹, Kathleen M. Jensen¹, Zachary Jorgenson³, Michael D. Kahl¹, Kathy Lee⁴, Anthony L. Schroeder⁵, Kyle E. Stevens¹, Linnea M. Thomas¹, Matthew A. Weberg¹, Gerald T. Ankley¹.

¹ U.S. Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Mid-Continent Ecology Division, 6201 Congdon Boulevard, Duluth, MN, USA, 55804

² National Research Council, U.S. Environmental Protection Agency, 6201 Congdon Boulevard, Duluth, MN, USA, 55804

³ US Fish and Wildlife Service, Twin Cities Ecological Field Services Field Office, 4101 American Blvd East, Bloomington, MN, USA, 55425

⁴ U.S. Geological Survey, Toxic Substances Hydrology Program, 415 S. Pokegama Ave, Grand Rapids, MN, USA, 55744

⁵ University of Minnesota Crookston, Math, Science and Technology Department, 2900 University Ave., Crookston, MN, USA, 56716

* Address correspondence to lishibin1116@gmail.com

^A Current address: U.S. Geological Survey Columbia Environmental Research Center, 4200 E New Haven Rd, Columbia, MO, USA, 65201.

ABSTRACT

Environmental assessment of complex mixtures typically requires integration of chemical and biological measurements. This study demonstrates the use of a combination of instrumental chemical analyses, effects-based monitoring, and bio-effects prediction approaches to help identify potential hazards and priority contaminants in two Great Lakes Areas of Concern (AOCs), the Lower Green Bay/Fox River located near Green Bay, WI, USA and the Milwaukee Estuary, located near Milwaukee, WI, USA. Fathead minnows were caged at four sites within each AOC (eight sites total). Following 4 d of in situ exposure, tissues and biofluids were sampled and used for targeted biological effects analyses. Additionally, 4 d composite water samples were collected concurrently at each caged fish site and analyzed for 132 analytes as well as evaluated for total estrogenic and androgenic activity using cell-based bioassays. Of the analytes examined, 75 were detected in composite samples from at least one site. Based on multiple analyses, one site in the East River and another site near a paper mill discharge in the Lower Green Bay/Fox River AOC, were prioritized due to their estrogenic and androgenic activity, respectively. The water samples from other sites generally did not exhibit significant estrogenic or androgenic activity, nor was there evidence for endocrine disruption in the fish exposed at these sites as indicated by the lack of alterations in *ex vivo* steroid production, circulating steroid concentrations, or vitellogenin mRNA expression in males. Induction of hepatic *cyp1a* mRNA expression was detected at several sites, suggesting the presence of chemicals that activate the aryl hydrocarbon receptor. To expand the scope beyond targeted investigation of endpoints selected *a priori*, several bio-effects prediction approaches were employed to identify other potentially disturbed biological pathways and related chemical

constituents that may warrant future monitoring at these sites. For example, several chemicals such as diethylphthalate and naphthalene, and genes and related pathways, such as cholinergic receptor muscarinic 3 (CHRM3), estrogen receptor alpha1 (esr1), chemokine ligand 10 protein (CXCL10), tumor protein p53 (p53), and monoamine oxidase B (Maob), were identified as candidates for future assessments at these AOCs. Overall, this study demonstrates that a better prioritization of contaminants and associated hazards can be achieved through integrated evaluation of multiple lines of evidence. Such prioritization can guide more comprehensive follow-up risk assessment efforts.

KEYWORDS: Mixture, Screening, Chemical-Biomolecule Interactions, ToxCast™, Contaminants of Emerging Concern, Risk Assessment

1. INTRODUCTION

The Great Lakes Restoration Initiative (GLRI) was initiated in 2010 to accelerate efforts to protect and restore the largest system of fresh surface water in the world – the North American Great Lakes. The GLRI is addressing a variety of threats to the Great Lakes including nutrient runoff, introduction of invasive species, habitat depletion, and potential effects of contaminants, including chemicals of emerging concern (CECs). The present work is part of a multiagency GLRI effort aimed at development and application of novel methods for assessing complex mixtures of aquatic contaminants in Great Lakes tributaries, including multiple Areas of Concern (AOCs) (Ekman et al., 2013).

Assessing the potential for adverse biological effects of complex mixtures of contaminants in aquatic environments is challenging, and often best achieved through the integration of multiple lines of evidence. Traditionally, instrumental analysis has been the most widely used approach for surveillance and monitoring of mixtures of chemical contaminants in aquatic environments (e.g., Wagner et al. 1998; Stackelberg et al. 2001; Foster et al., 2005; Kolpin et al., 2002). While instrumental analysis can effectively characterize the occurrence of many contaminants, this approach alone has shortcomings relative to complex mixture assessment (Escher et al., 2011; Escher et al., 2013; Leusch et al., 2014). Some chemicals are not detected at biologically relevant concentrations using standard extraction and instrumental analyses, while in other instances the identity of chemicals present in a sample may not be known. When chemicals are

successfully detected, there often are insufficient toxicological data available to estimate possible hazards of detected chemicals. Lastly, there is often a limited understanding of chemical interactions among the complex chemical mixtures that exist in the environment.

In response to these recognized limitations, complementary biological effects-based monitoring approaches have been employed by a number of regulatory programs in the United States, such as the National Pollution Discharge Elimination System (NPDES) program for effluents (USEPA, 1984). Effects-based methods provide a direct measure of the integrated biological response to a complex mixture. Responses measured may be highly integrative, as in the case for testing associated with the NPDES program which utilizes apical endpoints in whole animals, or they may be specific to certain chemical classes and/or biological activities, as in the case of in vitro assays focused on a single signaling pathway (Macova et al., 2011; Tang et al., 2013).

Under the umbrella of effects-based monitoring, methods seek to balance the degree of control with environmental realism. For example, assays with fish may be conducted in the lab with field-collected water samples, which provides control of confounding environmental factors and low logistic cost, but fails to consider uncertainties due to fluctuating chemical exposures or the possible degradation of contaminants in collected/stored samples. At the other end of the spectrum, direct evaluation of feral fish considers chemical impacts in a realistic exposure scenario, but is only occasionally used due to high collection costs. Exposing fish in situ with caging systems can offer a cost-effective middle-ground between controlled laboratory exposure and field monitoring. Although the exposure duration of caged fish will typically be less than wild fish, in the short-term both can experience similar fluctuating chemical exposures. Further, since fish used in caging studies typically are from laboratory cultures, they have a known chemical exposure history and health status, which is not the case with field-collected animals.

Ideally, in caged fish studies chemical characterization of the exposure would closely match what the animal's experience. Grab samples may not reflect an exposure over the course of the study, as both chemical concentration and composition can fluctuate. Different types of passive samplers offer the potential to capture integrated contaminant samples (Miege et al., 2012; Vrana et al., 2001), but back-calculation to actual water concentrations is challenging (Yilmaz et al., 2014). As an alternative, Kahl et al. (2014) described a composite sampling system that provides a time-integrated sample of water directly matched to organism deployment.

Samples collected from the field can be assessed using targeted (or supervised) or untargeted (unsupervised) measurements (Ekman et al. 2013). Targeted assays/endpoints are employed when there are specific biological activities/pathways of concern. For example, a common *in vivo* measurement for detecting exogenous estrogens is vitellogenin (egg yolk precursor; vtg) protein or mRNA in male fish (Hutchinson et al., 2006; Sumpter and Jobling, 1995).

Analogously water samples can also be assessed using targeted *in vitro* systems. such as the estrogen-responsive T47D-KBluc cell bioassay (Cavallin et al. 2015; Wehmas et al., 2011; Wilson et al., 2004).

Untargeted approaches such as 'omics' measurements in fish tissues (e.g., Berninger et al., 2014; Garcia-Reyero et al., 2011; Martinovic-Weigelt et al., 2014; Skelton et al., 2014) and analysis of water samples or extracts using batteries of pathway-based *in vitro* assays (Escher et al., 2014; Schroeder et al., 2016) are an important complement to supervised measurements when, as typically is the case with environmental samples, unknown/unmeasured contaminants are of concern. Untargeted approaches ideally assess activities of contaminants without bias for any specific biological pathways, which is especially important for surveillance efforts (Ekman et al. 2013; Escher et al. 2014; Schroeder et al. 2016).

A significant challenge in utilizing the sometimes extensive amount of analytical chemistry data collected in conjunction with monitoring/surveillance efforts is identification of potential biological hazard. Comprehensive biological effects data typically are available only for a small number of the measured chemicals, and what is available often is widely distributed across the peer-reviewed literature. However, critical progress has been made to capture and centralize information concerning chemical effects on molecular/biochemical endpoints in both in vitro and in vivo systems. For example, the Comparative Toxicogenomics Database (CTD; ctdbase.org) includes more than 24 million gene-chemical, chemical-disease and gene-disease interactions collected from the open literature (Davis et al., 2015). Another increasingly important source of bioeffects data for a large number of chemicals is being generated through high-throughput toxicity testing (HTT). For example, the US Environmental Protection Agency (USEPA) Toxicity ForeCaster (ToxCast™) and associated Tox21 programs have generated dose-response data for more than 8500 unique substances using up to 821 assay endpoints (Dix et al., 2007; Houck et al., 2013; Judson et al., 2010). These data are publically available in a variety of formats (<https://www.epa.gov/chemical-research/toxicity-forecaster-toxcasttm-data>). Both curated literature sources like CTD, and HTT programs like ToxCast™, have created important opportunities to bridge analytical chemistry with known biological effects for a more comprehensive assessment of contaminant mixtures. These types of datasets can be used to develop hypotheses regarding pathways that may be impacted by chemicals detected at field sites, but not explicitly investigated by using supervised analyses.

The objective of the present study was to integrate a combination of instrumental analytical measurements, targeted effects-based monitoring, and bio-effects prediction approaches to help identify potential hazards and priority contaminants in two Great Lakes AOCs, the Lower Green

Bay/Fox River AOC located near Green Bay, WI, USA and the Milwaukee River Estuary, located in Milwaukee, WI, USA. This integrated, weight-of-evidence evaluation contributes to risk-based characterization of complex contaminant mixtures in the environment which can inform subsequent monitoring, remediation, and/or restoration activities aimed at maintaining or improving the quality of critical fresh water resources.

2. MATERIALS AND METHODS

2.1 Location and Sites

A total of eight sites within the Lower Green Bay/ Fox River and Milwaukee Estuary were selected to evaluate the application of an integrated approach for identifying potential hazards and priority contaminants (Figure 1). These locations were designated AOCs under the 1987 U.S.-Canada Great Lakes Water Quality Agreement, and are characterized by significant impairments of beneficial uses as a result of local human activities. The Lower Green Bay/ Fox River AOC (Green Bay AOC) consists of 11.3 km of the Fox River and 88.1 km² area of southern Green Bay. Various factors, such as upstream land use, high turbidity, sedimentation, fluctuating dissolved oxygen, and frequent algal blooms, among others, contribute to the impacts in this shallow basin (USEPA, accession date 02/10/2016). The Milwaukee Estuary AOC (Milwaukee AOC) consists of 13.8 km of rivers, a portion of the nearshore waters of Lake Michigan, and the inner and outer Milwaukee harbor. The legacy contaminants of concern at the two AOCs originate from both point and nonpoint sources, such as historic industrial discharges, municipal wastewater treatment plants (WWTPs), combined sewer overflows, and agricultural and urban runoff (USEPA, accession date 02/10/2016). While the presence of legacy

contaminants like polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs), and metals were primary drivers for designation of these locations as AOCs, unknown potential hazard in terms of biological responses associated with a wide-range of point and non-point source inputs makes these locations suitable for the application of integrated monitoring approaches. Given the amount of existing knowledge concerning the occurrence and effects of legacy contaminants in the two systems (e.g., PCBs in Green Bay), the focus of the current work was on potential hazards associated with CECs that may be entering the Great Lakes via these tributaries. In the present study, the term CECs is used to refer to the diversity of chemicals, present in the environment, for which exposure and/or effects characterization is quite limited. This could include pharmaceuticals, personal care products, flame retardants, current use pesticides, perfluorinated compounds and others. As such, the term encompasses a broad range of organic chemicals with a wide range of commercial uses, production volumes, and biological activities.

Four sites were selected at each location based on anticipated sources of CECs, and feasibility of water sampling and fish deployment (Figure 1). Many of these sites were located downstream of WWTPs. For the Green Bay AOC, DPERE (De Pere) and GRBAY (Green Bay) were near the respective WWTP outfalls, PRGAM was near a paper mill discharge, and the fourth site, EASTR, was in the East River, just upstream of its confluence with the Fox River (Figure 1A). For the Milwaukee AOC, JONESI was at the Jones Island WWTP outfall, and KINNI, MENMR, and MILWR were located on the Kinnickinick, Menominee, and Milwaukee rivers, respectively; each of the three tributaries that converge upstream of the Jones Island WWTP before flowing into Lake Michigan (Figure 1B). Water depths for the caged fish deployments ranged from 1.2 to 9.1

m. A detailed description of the location and water body of these sites are in Supplementary Tables S1, S2.

2.2 Fish Handling, Transport, Deployment and Processing (Figure S1)

All protocols involving fish handling, transport, deployment and sampling were reviewed and approved by the Animal Care and Use Committee, in accordance with the Animal Welfare Act and Interagency Research Animal Committee guidelines. Adult fathead minnows (*Pimephales promelas*), aged 5-6 months with well-defined secondary sex characteristics, were obtained from an on-site culture unit at the Duluth (MN) USEPA laboratory. Fish were from mass cultures held at 20 °C with a 16:8-h light:dark photoperiod, and fed a mixture of brine shrimp and ground trout chow twice daily. Fish for the field exposures were transported to designated sites along with time zero (T_0) reference fish. Specifically, for each deployment, groups of six males and six females from the mass culture were placed together in plastic bags containing approximately 4 L of oxygenated Lake Superior water, and subsequently put in coolers on ice. Each group of 12 animals was used as one replicate treatment at the study sites, with two replicates deployed at each site. Temperature of the water in the bags was monitored during shipping using automatic temperature loggers (HOBO Data Loggers, Onset, Bourne, MA). No food was provided in transit. The coolers were shipped overnight to the appropriate locations. Following arrival, fish were held overnight prior to deployment or sampling (in the case of the T_0 reference fish). For in situ deployment, fish were acclimated for five to ten minutes to the local site water temperature (± 3 °C), by placing bagged fish into buckets of water from the deployment site. Two replicate cages of fish, each containing six males and six females were attached to a buoy-cable system at a mid-water depth and deployed for 4 d at the study sites using the basic apparatus described elsewhere (Kahl et al., 2014).

On the same day each batch of fish was deployed in situ, T₀ reference fish that were handled, shipped, and held in the same manner as those deployed in the field were dissected and samples collected as described below. Additional controls consisting of fish (a) taken directly from mass culture before shipping, (b) held in oxygen saturated bags on ice for 2 d but not transported, or (c) held in bags on ice for 2 d, then in the laboratory for an additional 4 d (in a constant flow of Lake Superior water), were also sampled to assess the potential impact of variables associated with shipping, temperature change, altered photoperiod, and other handling-related stress (Kahl et al., 2014). After 4 d of in situ exposure, fish were retrieved from the cages, securely placed in buckets of site water, and transported to local laboratories for processing. All fish were processed within 6 h of being retrieved from the field. At the processing site, fish were anesthetized with 100 mg/L tricaine methanesulfonate (MS-222, Finquel, Argent, Redmond, WA) buffered with 200 mg/L sodium bicarbonate and weighed. Obvious external lesions and status of secondary sexual characteristics were evaluated. Blood, liver, gonads, and brain from both sexes, and urine from males were collected and stored following procedures described elsewhere (Ekman et al., 2013). Due to logistic constraints, only two sites were deployed/sampled per day. Thus, sites were divided into four groups: (a) DPERE and EASTR, (b) GRBAY and PRGAM, (c) MENMR and MILWR, and (d) JONESI and KINNI.

Dissolved oxygen, conductivity, and temperature were measured at the study sites on Day 1 (deployment) and Day 4 (retrieval). Hardness, and alkalinity measurements were also obtained on several sub-surface grab water samples.

2.3 Water Sampling and Chemical Analysis

A critical aspect of the aforementioned test system involved a composite water autosampler, which was simultaneously deployed with the caged fish at the same depth, thereby enabling the temporally-integrated characterization of the exposure experienced by the animals (Kahl et al., 2014). Details concerning the design, construction and programming of the autosamplers are provided elsewhere (Kahl et al., 2014). The sampler was programmed to collect approximately (26 mL) of water at 15 min intervals for the entire exposure period. Grab water samples were also collected from a depth of approximately 1 m from the bottom using a peristaltic pump (model 410; Solinst) at the beginning and/or end of the exposures. Two grab and one composite sample were collected at each site in the Green Bay AOC, while one grab and one composite sample were collected at each site in the Milwaukee AOC (Table S3). Subsamples of the composite and grab water samples were placed on ice and shipped overnight to the US Geological Survey National Water Quality Laboratory in Denver, Colorado, USA. Upon arrival, water samples were extracted and analyzed for 132 organic compounds including pesticides, industrial chemicals, and pharmaceuticals among others. The complete list of target analytes, specific extraction, quality assurance/quality control, and quantification methods, and actual chemical concentrations at the two AOCs in the present study are detailed elsewhere (Lee et al., 2012).

Estimated values were assigned for chemicals with concentrations below the laboratory reporting limit (LRL) based on previously discussed justification and rationales (Lee et al., 2012; Oblinger et al., 1999). Specifically, a value of 1/10 of the LRL was assigned when no measurement in the entire dataset for that particular analyte exceeded the LRL. A value of 1/4 of the lowest measured concentration was assigned when there was at least one measurement in the whole dataset that exceeded the LRL for that analyte. These estimated values were incorporated in

order to avoid zeros in the dataset, which are problematic for a number of the computational analyses employed in our study. For the purpose of visualization and cluster analysis, chemistry data were median centered across the entire dataset for Great Lakes AOCs as first reported elsewhere (Lee et al., 2012). Based on this median-centered dataset, heat maps and clustering diagrams were generated using MeV version 4.9 (Dana-Farber Cancer Institute, 2013).

2.4 Targeted Bio-effects Measurements

A series of targeted biochemical and molecular endpoints were measured which can indicate chemical impacts on xenobiotic metabolism and/or aspects of the HPG axis function in fish, as reported in previous investigations (Adedeji et al., 2012; Ankley et al., 2010; Cavallin et al., 2015; Villeneuve et al., 2008).

Total estrogenic and androgenic activity of the composite water samples was measured using the T47D-KBluc and MDA-kb2 cell lines, which have been stably transfected with human estrogen receptor α (ER α)- or androgen receptor (AR)-luciferase reporter gene constructs, respectively (Wilson et al., 2004; Wilson et al., 2002). Composite water samples were directly used to prepare culture media and tested at two concentrations, 100% and 25% of ambient. Media prepared using high-quality deionized water was used for controls and for sample dilution as described elsewhere (Cavallin et al., 2014; Wehmas et al., 2011). Activity was considered significant when relative luminescence activity exceeded baseline values by more than three standard deviations (Kahl et al., 2014). Sample estrogenicity and androgenicity were expressed relative to 17 α -ethinyl estradiol (EE2) and 17 β -trenbolone (TRB) standards, as EE2 or TRB equivalents (-EQ), following procedures detailed elsewhere (Cavallin et al., 2015). Standards were dosed into control media as concentrated stock solutions dissolved in ethanol (<0.1% per well) as a carrier

solvent, and a 0.1% ethanol control was included on each plate to test for potential solvent effects.

Ex vivo production of sex steroids (estradiol, E2; testosterone, T) by ovaries, and *ex vivo* production of T by testes, were determined using methods previously detailed elsewhere (Ankley et al., 2007; Jensen et al., 2001; McMaster et al., 1995). Briefly, subsamples of testes (4.91 ± 0.22 mg wet wt) and ovaries (11.63 ± 0.51 mg wet wt) were transferred to 500 μ l of Medium 199 supplemented with 1 μ M of 25-OH-cholesterol and 1 mM IBMX on ice. Once all fish were sampled, tubes were transferred to a 25°C water bath and incubated for 12 h. Incubations were terminated by placing the tubes back on ice and media was collected from each tube, frozen on dry ice, and stored at -20°C until steroids were extracted and analyzed by radioimmunoassay. Vitellogenin protein concentrations were measured in plasma of males using enzyme-linked immunosorbent assays (ELISAs) with a fathead minnow polyclonal antibody and a purified fathead minnow vitellogenin standard (Korte et al., 2000; Parks et al., 1999).

Relative abundance of select mRNA transcripts in the male liver was quantified using real-time quantitative polymerase chain reaction (QPCR). The transcripts examined included an estrogen-inducible gene, *vtg*, and two genes involved in oxidative xenobiotic metabolism, cytochrome P450 (CYP)1a1 mRNA (*cyp1a1*) and CYP3A mRNA (*cyp3a*). Total RNA was extracted from the liver using RNeasy Plus Mini kits (Valencia, CA) according to the manufacturer's protocol. Quantity and quality of total RNA was then evaluated using a Nanodrop ND 1000 spectrophotometer (Nanodrop Technologies, Wilmington, DE). Purity of the samples was assessed according to the optical density at 260/280 nm. Total RNA samples were diluted to 10 ng/ μ l for use in QPCR assays. Relative abundance of *vtg*, *cyp1a1*, and *cyp3a* transcripts were determined using Taqman RNA to CT 1 step kits (Applied Biosystems, Foster City, CA). Each

20 μ L QPCR reaction contained 20 ng total RNA, 300 nM forward primer, 300 nM reverse primer, and 150 nM probe. The thermocycling program was set to 48°C for 15 min, 95°C for 10 min, and finally 40 cycles of 95°C for 15 s and 60°C for 1 min. For each gene, relative transcript abundance was quantified relative to a gene-specific mRNA standard curve containing six concentrations, following a 10-fold dilution series starting at 2×10^7 copies/ μ L. The mRNA standards were prepared by cDNA amplification and in vitro transcription reactions (MEGAscript, Ambion, Carlsbad, CA) as described previously (Villeneuve et al., 2007). The primers, probes, and protocol have been described elsewhere (Ankley et al., 2007; Cavallin et al., 2015) (Table S4).

2.5 Bio-effects Prediction

In conjunction with the targeted biological effects measurements, we employed two prior knowledge approaches for bio-effect predictions (Schroeder et al., 2016) based on site-specific chemistry data. One approach uses a publicly available *in vivo* chemical interaction source, the CTD. This database was used to identify genes and associated pathways potentially impacted by chemicals detected in the environmental samples. Specifically, the list of chemicals detected at each site was uploaded to the CTD, which was followed by a batch query, extracting all gene targets with reported interactions with those chemicals from the database (Schroeder et al. 2016). The resulting data files were then used to build chemical-gene interaction networks. Site specific chemical-gene interaction networks were derived from a knowledge assembly model (KAM) and visualized using Cytoscape v3.0.2 software as described elsewhere (Martinovic-Weigelt et al., 2014). Briefly, nodes representing either chemicals (rectangles) or genes (ovals) are connected by edges (lines) indicating the existence of an interaction (Figure S2). The number of edges connected to each gene (degree) denotes the total number of chemicals detected at the site

that have been reported to interact with that gene target. Genes were then ranked based on the number of chemicals at each site that could interact with that gene based on literature curated in the CTD. The gene lists were further analyzed using Set Analyzer Tool (ctdbase.org) to identify biological processes and pathways potentially impacted by the chemical mixtures at each location (University). Enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were considered significant at $p < 0.01$.

A more quantitative bio-effects prediction approach that involved comparing site-specific chemical concentrations to the biological effect levels measured in ToxCast™ high throughput screening assays was also employed. The ToxCast™ database (see iCSS dashboard; <http://actor.epa.gov/dashboard/>) contains information from 13 different assay platforms, comprising over 700 high-throughput assays that have been used to test over 2000 chemicals, including industrial products, pharmaceuticals, pesticides and food additives (Kavlock et al., 2012). Unlike the CTD, the ToxCast™ data were generated in a standardized manner where the chemicals were prepared, delivered, tested at multiple concentrations, and analyzed according to the same protocol. An advantage arising from such a procedure is that these data are indicative of both the chemical-pathway interaction and relative potency of chemicals and related effect concentrations (Schroeder et al., 2016). Specifically, the ToxCast™ database provides a half-maximal activity concentration (AC50) for each chemical-assay pair, which is the concentration at which a chemical induces 50% of maximum effect on activity. With the AC50 and chemical concentration information, an exposure-activity ratio (EAR) can be calculated. The EAR is defined as the measured chemical concentration in the water sample, divided by the in vitro AC50 values (Becker et al., 2015):

$$(1) \text{ EAR} = \text{Measured concentration in the environment } (\mu\text{M}) / \text{AC50}(\mu\text{M})$$

Simplistically, an EAR value for a given chemical equal to or greater than 1, indicates that the measured concentration in the water sample would, theoretically, be high enough to elicit the observed bioactivity. Although this does not necessarily translate directly to exposure and effects *in vivo*, the EAR calculation does provide a risk-based relative estimate of pathway-specific bioactivity in the context of chemical prioritization. A summed EAR value for each site can be calculated by adding EAR values across all assays for that site.

An important feature of ToxCast™ database is its incorporation of dose response information but, unfortunately, not every chemical has a reliable dose-response relationship in terms of an AC50 derivation. Hence, EARs for some chemicals were not calculated due to a lack of reliable AC50 values in the database (Table S5). For example, for assays measuring peroxisome proliferator-activated receptor gamma activity, EARs of DEET (N, N-diethyl-meta-toluamide) were excluded due to an inaccurate dose-response curve estimated by the customized models (hit call negative; flagged as biochemical assay with <50% efficacy; Figure S3).

2.6 Data analysis

Analysis of variance was performed for targeted biological measurements. Data were tested for normality using a Kolmogorov-Smirnov test, and homogeneity of variance using Levene's test. If data failed either test, logarithmic transformation was performed and normality and homogeneity of variances were checked again. If the data failed assumption after transformation, a non-parametric Kruskal–Wallis ANOVA test was performed. Data from T47D-KBluc and MDA-kb2 assays were analyzed using a one-way ANOVA followed by Tukey's multiple comparisons for differences among sites. Gene expression and plasma vitellogenin data were analyzed using

mixed ANOVA model (Bates and DebRoy, 2004) with ‘site’ as fixed effect and ‘group’ as a random effect. If random effects were not significant, a one-way ANOVA followed by Tukey’s test was applied. If the data did not pass the check of normality and homogeneity of variance (*vtg* mRNA expression, Figure 5A), ANOVA followed by Tukey’s multiple comparisons was applied to each group separately to examine within group differences. All statistical analyses were conducted using R version 3.2.3 (R Core Team, 2015) and JMP SAS version 11.2.1 (SAS Institute 2014). All differences were considered significant at $p < 0.05$.

3. RESULTS

3.1. In Situ Exposures

There was 100% survival and recovery of 192 fish shipped and deployed in cages in situ (two cages per site, 6 fish of each sex per cage) at the Green Bay and Milwaukee sites. Water quality varied across the study sites. Briefly, mean water temperature, pH, conductivity, alkalinity, and hardness at deployment of the cages ranged from 15.9 °C to 24.9 °C, 7.81 to 8.61, 313 to 943 $\mu\text{S}/\text{cm}$, 137 to 259 mg/L CaCO_3 , and 158 to 302 CaCO_3 , respectively (Table S2). The successful retrieval of all deployed fish in a healthy state shows that overall water quality was at least sufficient to allow fish survival over the 4-d exposure period.

3.2. Characterization of Contaminant Mixture

Of the 132 chemicals analyzed in the present study, 75 were detected at one or more sites; detected chemicals included steroid hormones, pesticides, plasticizers and derivatives, surfactants, polyaromatic hydrocarbons (PAHs), and pharmaceuticals and personal care products

(PPCPs) (Ekman et al., 2013) (Figure 2). Out of the 75 chemicals detected, the majority were present at two or more sites, suggesting ubiquitous background contamination by some chemicals. For example, atrazine and bisphenol-A, were consistently detected across all sites with concentrations ranging from 0.043 to 0.074 µg/L, and 0.02 to 0.092 µg/L, respectively. Likewise, PAHs such as benzo[a]pyrene (0.0015 ~ 0.048 µg/L), phenanthrene (0.0021 ~ 0.044µg/L), and pyrene (0.0022 ~ 0.104 µg/L) were frequently detected across all sites.

In general, for the Green Bay AOC, there were no distinct patterns of chemical concentration relative to the sampling time and spatial orientation of the sites along the upstream to downstream gradients (Figure 2). For example, bisphenol-A was detected at relatively similar concentrations across all the sites, despite the fact it has applications that might be associated more with one site than another (e.g., coating thermal paper). The general similarity in patterns of chemicals across the DPERE, PRGAM, and GRBAY sites are likely due to their shared upstream watershed area which encompasses a wide range of agricultural, urban, and industrial sources that may be influencing chemical composition. In contrast, the watershed area upstream of the EASTR is much smaller, predominantly urban, and not subject to the same upstream influences as the other sites. Indeed, based on cluster analyses of the chemistry profile, samples from EASTR were most distinct from those collected from other sites at the Green Bay AOC. A closer examination indicated that several potential estrogenic chemicals were present at this site, including 17β-estradiol, 4-nonylphenol, and estrone, with highest concentrations of 0.159 ng/L, 0.028 µg/L, and 0.777 ng/L, respectively. Thus, not surprisingly, EASTR, which is upstream of the confluence of the East River with the Fox River (Figure 1), reflects a different set of inputs than the sites along the Fox River.

For the Milwaukee AOC, there was a clear pattern of chemical occurrence related to sampling location. Specifically, the JONESI site clustered most distinctly from the other sites. Although all four sites are located in an urbanized area, JONESI is the only one close to a WWTP facility. Multiple pharmaceuticals were detected at their highest concentrations at this site; these included citalopram (0.028 µg/L), lidocaine (0.0461 µg/L), tramadol (0.104 µg/L), methocarbamol (0.338 µg/L), diltiazem (0.0297 µg/L), carbamazepine (0.0634 µg/L), venlafaxine (0.0288 µg/L), and diphenhydramine (0.0297 µg/L). Other routine wastewater indicators such as 3-beta-coprostanol (0.255 µg/L), caffeine (0.0915 µg/L), and cotinine (0.0386 µg/L), were also detected at their highest concentrations at this site. Chemicals not used as indicators of WWTP effluent generally had similar concentrations across all sites in this AOC, suggesting a relatively homogeneous background contamination in this predominantly urban watershed.

3.3. Targeted Bio-effects

Estrogenic activity was observed in the T47D cell assay at several sites, such as JONESI, EASTR, GRBAY, and DPERE (Figure 3A). Similarly, androgenic activity was observed in the MDA-kb2 assay at PRGAM, MENMR, DPERE, GRBAY, and JONESI, with highest activity (12.3 ± 2.7 ng TRB-equivalents/L) associated with a paper mill site (PRGAM) (Figure 3B). No quantifiable levels of EE2 and TRB EQs were detected in Milli-Q water and media controls (data not shown).

In general, no statistically significant differences were observed for *ex vivo* T production, with the exception of a small increase ($p < 0.05$) by ovary tissue harvested from adult female fathead minnows caged at the EASTR site (Figure 4A, C). Relative to group specific day 0 controls, *ex vivo* E2 synthesis was significantly elevated in ovary tissue from females exposed at three sites

in the Lower Fox River/Green Bay AOC; the EASTR, GRBAY, and PRGAM (Figure 4B).

Plasma vitellogenin concentrations in male fish caged were not significantly increased compared to corresponding d 0 controls (Figure 4D).

Expression of the estrogen-inducible gene, *vtg* (Figure 5A), and two genes involved in oxidative xenobiotic metabolism, *cyp1a1* (Figure 5B) and *cyp3a* (Figure 5C) in the livers of adult male fathead minnows showed variable effects. Compared with relevant group-specific day 0 controls, an increase ($p < 0.05$) in *vtg* expression was observed in males caged at the EASTR site. However, as noted earlier, there was no evidence of a significant increase in plasma vitellogenin concentrations in males caged at the EASTR site (Figure 4D). Due to variation in the baseline expression of *vtg* mRNA among the four groups of fish shipped, differences among sample groups were not compared (Figure 5A). The greatest impact on gene expression was observed for *cyp1a1*, whose expression in liver tissue increased ($p < 0.05$) in fish caged at the EASTR, DPERE, MENMR, JONESI, and KINNI sites (only males were examined). Compared with relevant controls, no significant differences were observed in *cyp3a* expression across all sites.

3.4. Bio-effects Prediction

To complement insights provided by the targeted endpoints, bio-effects prediction approaches were employed to generate hypotheses and identify priority biological pathways and/or chemicals for possible follow-up investigation. Information from CTD was used to identify gene targets reported to be modulated by chemicals detected in water from the study sites (Figure 6). Based on this approach, several genes involved in hepatic xenobiotic metabolism (e.g., *cyp1a2*, *cyp1a1*, *cyp2b6*, *cyp3a4*) were associated with chemicals detected in the present study. Based on chemicals detected in the water samples (benzo[a] pyrene, bisphenol-A, EE2, 17 α -estradiol, 17 β -

estradiol, 4-nonylphenol, and testosterone etc.), there was also evidence based on CTD to expect possible interactions with estrogen receptor isoforms *esr1* and *esr2* (Figure 6). Applying functional enrichment analysis ($P < 0.01$) to the list of genes for which modulation by chemicals present in the site waters was plausible based on CTD, we identified a number of pathways that could be impacted by exposure to site water, which included: steroid hormone biosynthesis, apoptosis, and pathways in cancer (Figure 6).

EAR values were calculated based on the chemical concentration data and AC50 values reported for ToxCast™ assays (Figure 7). Out of 821 ToxCast™ assays, six assays had a summed EAR value greater than 0.1 from at least one sample (Figure 7; Table S6). A variety of chemicals were associated with these EAR values, including 17 β -estradiol, 4-nonylphenol, estrone, diethylphthalate, triclosan, naphthalene, benzo[a]pyrene, phenanthrene, pyrene, phenobarbital, carbamazepine, phenytoin, diltiazem and others.

4. DISCUSSION

In this study, eight sites within two Great Lakes AOCs were investigated relative to CEC contaminants and related hazards using an integrated methodology, which incorporated the use of chemical analyses, biological effect monitoring, and bio-effect prediction. Despite the general similarity of the composition of chemical profiles across different sites in these two urbanized AOCs, several sites, such as EASTR and PRGAM, could be prioritized for future investigations based on their distinct chemical concentrations and/or relevant biological effects identified by multiple approaches. Furthermore, bio-effect prediction methodologies demonstrated their

capability of identifying additional pathways and priority chemicals for a more comprehensive monitoring or risk assessment.

Of particular interest to our studies was possible endocrine activity of CECs at these sites. Municipal discharges are known sources of estrogenic and occasionally androgenic CECs (Sheehan et al. 2002; Braga et al. 2005; Sun et al. 2008; Cavallin et al. 2015), and many of the sites in this study were either located near a WWTP (e.g., Green Bay AOC), or along a gradient in relation to a WWTP discharge (e.g., Milwaukee AOC). We used both targeted in vitro and in vivo measures of potential endocrine activity. Elevated *vtg* mRNA expression in males suggested the presence of exogenous estrogens at one particular site, EASTR, which is consistent with the existence of known estrogens at the site and their in vitro potency (0.13 ± 0.01 ng EE2-equivalents/L) measured by T47D-KBluc bioassay. This relatively low in vitro potency also agreed well with the concentrations of estrogens therein. For example, estrone was detected highest at EASTR of 0.777 ng/L, which was still lower than the lowest effect concentration of 0.307 μ g/L for plasma vitellogenin induction in a 21-d fathead minnow reproduction test (Thorpe et al., 2007). Similarly, 17 β -estradiol had a concentration of 0.159 ng/L, lower than the effective concentration of 18 ng/L for plasma vitellogenin induction in male fathead minnows following a 21 d exposure (Shappell et al., 2010). Some uncertainties of these comparisons may exist for sites in Green bay AOC because, grab rather than composite water samples, were analyzed for chemical profiles due to limited volumes of composite samples retrieved from the field. Overall, based on these in vitro and in vivo observations, EASTR could be prioritized among these sites for future monitoring efforts focused on potential endocrine effects. Regarding androgenic activity, attention could be focused on PRGAM, MENMR, GRBAY, JONESI, and DPERE. PRGAM, a paper mill effluent discharge site, has the highest androgenic activity (12.3

± 2.7 ng β -TRB-equivalents/L, MDA-kb2 assay), which was higher than a previously reported reproductive effect concentration of 11 ng α -TRB /L in fathead minnows (Jensen et al., 2006). Notably, none of the androgenic chemicals measured, such as testosterone, cis-androsterone, dihydrotestosterone, androstenedione, and epitestosterone were detected in the current study. This suggests the presence of novel androgenic chemicals at the PRGAM site, an observation consistent with that made at other sites impacted by pulp and paper mills (Durhan et al., 2002). This observation serves to highlight the utility of effects-based approaches as a complement for instrumental analyses where the bulk of biologically active chemicals may be functionally “invisible” (Tang et al., 2013). In any instance, based on this observation, PRGAM could be prioritized for further investigations as to the occurrence and consequences of androgenic activity.

Two genes (*cyp1a1* and *cyp3*) coding for enzymes involved in hepatic xenobiotic metabolism were also evaluated since their expression can be altered by a number of chemicals frequently found in WWTP effluents as well as other non-point sources. Further, their induction can lead to possible adverse effects through oxidative stress pathways (Denison, 1995). Catalytic measurement of *cyp1a* induction has been frequently used as a biomarker in fish for exposure to chemicals that bind the aryl hydrocarbon (Ah) receptor (Denison, 1995; Petrulis et al., 2001), and elevated *cyp1a1* was observed at several of the study sites. At some sites, the chemistry profile agreed with *cyp1a1* gene analyses. For example, increased *cyp1a1* expression at EASTR was in line with the presence of various PAHs, such as benzo[a]pyrene (0.046 $\mu\text{g/L}$), phenanthrene (0.0444 $\mu\text{g/L}$), and pyrene (0.0818 $\mu\text{g/L}$), which can activate the Ah receptor. It should be noted that many of the chemicals known to induce *cyp1a1*, such as PAHs and PCBs, would be more likely associated with sediments than the water column due to their hydrophobicity. A relevant

follow-up monitoring question in this regard would be to determine the degree to which chemicals causing the observed *cyp1a1* induction are, in fact, derived from sediments versus ongoing point-source inputs.

Interestingly, none of the sites altered *cyp3* expression in the caged fish, which was in agreement with the low concentrations of relevant pharmaceuticals at these sites since *cyp3a4* expression is frequently induced/inhibited by a number of pharmaceuticals, such as phenobarbital, carbamazepine, phenytoin, and diltiazem (Luo et al., 2002; Sutton et al., 1997).

In order to examine potential chemical disturbances of other biological pathways, two bio-effects prediction methodologies were used to link chemical occurrence data to potential hazards as a means to discover other impacted pathways, or at least generate hypotheses that might guide further investigation. According to CTD-derived predictions, major pathways that could plausibly be impacted by chemicals detected at the sites were related to metabolism, steroid hormone biosynthesis, apoptosis, and cancer (Figure 6). Among these, metabolism and apoptosis are likely too general to suggest specific endpoints to examine in subsequent monitoring. Further, cancer is generally not of concern from an ecotoxicological perspective, although in the case of AOCs, tumor incidence in fish has been cited as evidence for on-going beneficial use impairments (Blazer et al., 2009). Thus, follow-up monitoring might consider histological evaluation of resident fish. Alternatively a number of "hallmarks of cancer" have been identified (Goodson et al., 2015; Smith et al., 2016), some of which could be used as targeted endpoints for in vitro and/or short term in vivo monitoring (e.g., using caged fish), as a preliminary step before more resource intensive sampling of native fish was undertaken. Notably, disruption of apoptosis-related pathways is among the hallmarks of cancer (Smith et al. 2016).

In the case of CTD-predicted potential effects on steroid hormone biosynthesis, we had included targeted endpoints that could detect potential effects on those pathways. Some impacts on *ex vivo* T and E2 production were observed (Figure 4). However, CTD approach has a limitation in that it only considers chemical incidence and previous association with various biological effects. That is, the CTD approach does not explicitly account for either the relative concentrations detected in the sample, nor the relative potency of the chemicals. Hence, the approach is best employed for first tier hypothesis generation, as opposed to any sort of quantitative risk assessment. EAR analyses provide a more quantitative approach that can account for both relative environmental concentrations and pathway-based potency (Schroeder et al. 2016). (Figure 7). Assay targets with the highest EAR values across all sites included the cholinergic receptor muscarinic 3 (CHRM3), *esr1*, chemokine ligand 10 protein (CXCL10), tumor protein p53 (p53), and monoamine oxidase B (Maob) (Figure 7 and Table S6). These elevated EARs were associated with one or more detected chemicals. For example, summed EARs greater than 0.1 for *esr1* were observed across all sites due to the presence of various estrogenic chemicals such as 17 β -estradiol, 4-nonylphenol, and estrone. The elevated EAR for CHRM3, which upregulates intracellular calcium as a signaling pathway, was related to the presence of naphthalene in the current study. The MILWR site with a naphthalene concentration of 0.0197 $\mu\text{g/L}$, had the highest EAR (0.66) associated with CHRM3. However, the concentrations detected are much lower than USEPA's ambient water quality criteria for naphthalene (acute exposure, 2300 $\mu\text{g/L}$; chronic exposure 620 $\mu\text{g/L}$) (USEPA, 1980). Therefore, KINNI and MILWR, with the highest summed EAR, were not prioritized since the summed EARs were mainly driven by the naphthalene response in the CHRM3-related assay. In contrast, MENMR, with the highest concentration of diethylphthalate (1.25 $\mu\text{g/L}$) together with the existence of

triclosan (6 ng/L), had the highest EAR (0.11) associated with Maob, a key player in dopamine metabolism. Water quality criteria have not been established for these compounds. Consequently, association of their ambient concentrations with toxicologically-relevant bioactivity may be sufficient to stimulate follow up monitoring concerning effects of these compounds in the Menominee River watershed, potentially with a particular focus on dopamine-signaling related effects. These examples highlight how bio-prediction methodologies can help scientists and regulators identify pathways/bioactivities, and groups of chemicals for additional analytical characterizations for a particular site for a more comprehensive assessment, particularly in circumstances where established criteria or toxicological benchmarks are lacking.

5. CONCLUSIONS

Across a wide range of stakeholders, there is a need to identify and prioritize potential biological effects of chemicals present in the environment in order to direct environmental risk assessment and/or management activities. By an integrated and interactive evaluation of multiple lines of evidence, this study identified EASTR and PRGAM as priority sites for future investigations, due to the presence of estrogenic and androgenic activities, respectively. This study also demonstrated the use of bio-effects prediction methodologies in identifying additional priority chemicals, pathways and bioactivities not captured in targeted analyses. Based on this analysis, we concluded that potential monitoring of diethylphthalate and triclosan concentrations in the Menominee River, particularly in relation to potential behavioral or health consequences of disrupted dopamine signaling, may be a priority for further investigation. Overall, a prioritization of contaminants and associated hazards with limited resources can be better achieved through an

integrated and interactive evaluation of multiple lines of evidence. And such a prioritization, utilizing the tools discussed in this study, could help in a more comprehensive follow-up risk assessment effort.

ACKNOWLEDGEMENT

The present study was performed while two of the authors, Shibin Li and Jason P. Berninger, held National Research Council Research Associateship Awards at the U.S. Environmental Protection Agency, Mid-Continent Ecology Division. The authors also thank the following team members who have been critical in the implementation and success of the program that this study is a part of: JoAnn Banda (USFWS), Steve Choy (USFWS), Dan Gefell (USFWS), and Jeremy Moore (USFWS). This document has been subjected to review by the National Health and Environmental Effects Research Laboratory and approved for publication. Approval does not signify that the contents reflect the views of the Agency nor does mention of trade names or commercial products constitute endorsement or recommendation for use. The findings and conclusions in this article are those of the authors and do not necessarily represent the views or policies of the corresponding agencies.

REFERENCES

- Adedeji OB, Durhan EJ, Garcia-Reyero N, Kahl MD, Jensen KM, Lalone CA, Makynen EA, Perkins EJ, Thomas L, Villeneuve DL, Ankley GT. Short-term study investigating the estrogenic potency of diethylstilbesterol in the fathead minnow (*Pimephales promelas*). *Environmetnal Toxicology and Chemistry* 2012; 46: 7826-35.
- Ankley GT, Jensen KM, Kahl MD, Durhan EJ, Makynen EA, Cavallin JE. Use of chemical mixtures to differentiate mechanisms of endocrine action in a small fish model. *Aquatic toxicology* 2010; 99: 389-96.

- Ankley GT, Jensen KM, Kahl MD, Makynen EA, Blake LS, Greene KJ. Ketoconazole in the fathead minnow (*Pimephales promelas*): Reproductive toxicity and biological compensation. *Environmental Toxicology and Chemistry* 2007; 26: 1214-1223.
- Bates DM and DebRoy S. Linear mixed models and penalized least squares. *Journal of Multivariate Analysis* 2004; 91: 1-17.
- Becker RA, Friedman KP, Simon TW, Marty MS, Patlewicz G, Rowlands JC. An exposure:activity profiling method for interpreting high-throughput screening data for estrogenic activity—Proof of concept. *Regulatory Toxicology and Pharmacology* 2015; 71: 398-408.
- Berninger JP, Martinovic-Weigelt D, Garcia-Reyero N, Escalon L, Perkins EJ, Ankley GT. Using transcriptomic tools to evaluate biological effects across effluent gradients at a diverse set of study sites in Minnesota, USA. *Environmental science and technology* 2014; 48: 2404-12.
- Blazer VS, Rafferty SD, Baumann PC, Smith SB, Obert EC. Assessment of the "fish tumors or other deformities" beneficial use impairment in brown bullhead (*Ameiurus nebulosus*): I. Orocuteaneous tumors. *Journal of Great Lakes Research* 2009; 35: 517-526.
- Cavallin JE, Durhan EJ, Evans N, Jensen KM, Kahl MD, Kolpin DW. Integrated assessment of runoff from livestock farming operations: Analytical chemistry, in vitro bioassays, and in vivo fish exposures. *Environmental toxicology and chemistry* 2014; 33: 1849-57.
- Cavallin JE, Jensen KM, Kahl MD, Villeneuve DL, Lee K, Schroeder A, et al. Pathway-based approaches for assessment of real-time exposure to an estrogenic wastewater treatment plant effluent on fathead minnow reproduction. *Environmental Toxicology and Chemistry* 2015; DOI: 10.1002/etc.3228
- Davis AP, Grondin CJ, Lennon-Hopkins K, Saraceni-Richards C, Sciaky D, King BL, et al. The Comparative Toxicogenomics Database's 10th year anniversary: update 2015. *Nucleic Acids Research Nucleic Acids Research* 2015; 43: D914-D920.
- Denison M. Development of toxic equivalency factors for PCB congeners and the assessment of TCDD and PCB mixtures in rainbow trout. *Environmental Toxicology and Chemistry* 1995; 14: 861-871.
- Dix DJ, Houck KA, Martin MT, Richard AM, Setzer RW, Kavlock RJ. The ToxCast program for prioritizing toxicity testing of environmental chemicals. *Toxicological sciences* 2007; 95: 5-12.
- Durhan EJ, Lambricht C, Wilson V, Butterworth BC, Kuehl OW, Orlando EF, et al. Evaluation of androstenedione as an androgenic component of river water downstream of a pulp and paper mill effluent. *Environtal Toxicology Chemistry* 2002; 21: 1973-6.
- Ekman DR, Ankley GT, Blazer VS, Collette TW, Garcia-Reyero NI, Iwanowicz LR, et al. Biological Effects Based Tools for Monitoring Impacted Surface Waters in the Great Lakes: A Multiagency Program in Support of the Great Lakes Restoration Initiative. *Environmental Practice* 2013; 15: 409-426.
- Escher BI, Allinson M, Altenburger R, Bain PA, Balaguer P, Busch W, et al. Benchmarking organic micropollutants in wastewater, recycled water and drinking water with in vitro bioassays. *Environmental science and technology* 2014; 48: 1940-56.
- Escher BI, Lawrence M, Macova M, Mueller JF, Poussade Y, Robillot C, et al. Evaluation of Contaminant Removal of Reverse Osmosis and Advanced Oxidation in Full-Scale Operation by Combining Passive Sampling with Chemical Analysis and Bioanalytical Tools. *Environmental Science and Technology* 2011; 45: 5387-5394.
- Escher BI, van Daele C, Dutt M, Tang JYM, Altenburger R. Most Oxidative Stress Response In Water Samples Comes From Unknown Chemicals: The Need For Effect-Based Water Quality Trigger Values. *Environmental Science and Technology* 2013; 47: 7002-7011.
- Foster KL, Mackay D, Parkerton TF, Webster E, Milford L. Five-stage environmental exposure assessment strategy for mixtures: gasoline as a case study. *Environmental science and technology* 2005; 39: 2711-8.

- Garcia Reyero N, Lavelle CM, Escalon BL, Martinovic D, Kroll KJ, Sorensen PW, et al. Behavioral and genomic impacts of a wastewater effluent on the fathead minnow. *Aquatic Toxicology* 2011; 101: 38-48.
- Goodson WH, Lowe L, Carpenter DO, Gilbertson M, Manaf Ali A, Lopez de Cerain Salsamendi A, et al. Assessing the carcinogenic potential of low-dose exposures to chemical mixtures in the environment: the challenge ahead. *Carcinogenesis* 2015; 36: S254-S296.
- Houck KA, Richard AM, Judson RS, Martin MT, Reif DM, Shah I. ToxCast: Predicting Toxicity Potential Through High-Throughput Bioactivity Profiling. *High-Throughput Screening Methods in Toxicity Testing* 2013: 1-31.
- Hutchinson TH, Ankley GT, Segner H, Tyler CR. Screening and Testing for Endocrine Disruption in Fish-Biomarkers As "Signposts," Not "Traffic Lights," in Risk Assessment. *Environmental Health Perspectives* 2006; 114: 106-114.
- Jensen KM, Korte JJ, Kahl MD, Pasha MS, Ankley GT. Aspects of basic reproductive biology and endocrinology in the fathead minnow (*Pimephales promelas*). *Comparative Biochemistry and Physiology Part C: Toxicology and Pharmacology* 2001; 128: 127-141.
- Jensen KM, Makynen EA, Kahl MD, Ankley GT. Effects of the feedlot contaminant 17alpha-trenbolone on reproductive endocrinology of the fathead minnow. *Environmental Science and Technology* 2006; 40: 3112-7.
- Judson RS, Houck KA, Kavlock RJ, Knudsen TB, Martin MT, Mortensen HM, et al. In vitro screening of environmental chemicals for targeted testing prioritization: the ToxCast project. *Environmental health perspectives* 2010; 118: 485-92.
- Kahl MD, Villeneuve DL, Stevens K, Schroeder A, Makynen EA, LaLone CA, et al. An inexpensive, temporally integrated system for monitoring occurrence and biological effects of aquatic contaminants in the field. *Environmental toxicology and chemistry* 2014; 33: 1584-95.
- Kavlock R, Chandler K, Houck K, Hunter S, Judson R, Kleinstreuer N, et al. Update on EPA's ToxCast Program: Providing High Throughput Decision Support Tools for Chemical Risk Management. *Chemical Research in Toxicology* 2012; 25: 1287-1302.
- Kolpin DW, Furlong ET, Meyer MT, Thurman EM, Zaugg SD, Barber LB, et al. Pharmaceuticals, hormones, and other organic wastewater contaminants in U.S. streams, 1999-2000: a national reconnaissance. *Environmental science and technology* 2002; 36: 1202-11.
- Korte JJ, Kahl MD, Jensen KM, Pasha MS, Parks LG, LeBlanc GA, et al. Fathead minnow vitellogenin: Complementary DNA sequence and messenger RNA and protein expression after 17b-estradiol treatment. *Environmental Toxicology and Chemistry* 2000; 19: 972-981.
- Lee KE, Langer SK, Menheer MA, Foreman WT, Furlong ET, and Smith S, 2012, Characterization of chemicals of emerging concern in Great Lakes Areas of Concern 2012-2012, collection methods, analyses methods, quality assurance, and data, U.S. Geological Survey Data Series 575, 86
- Leusch FD, Khan SJ, Laingam S, Prochazka E, Froschio S, Trinh T, et al. Assessment of the application of bioanalytical tools as surrogate measure of chemical contaminants in recycled water. *Water Research* 2014; 49: 300-15.
- Luo G, Cunningham M, Kim S, Burn T, Lin J, Sinz M, et al. CYP3A4 induction by drugs: correlation between a pregnane X receptor reporter gene assay and CYP3A4 expression in human hepatocytes. *Drug Metabolism and Disposition* 2002; 30: 795-804.
- Macova M, Toze S, Hodgers L, Mueller JF, Bartkow M, Escher BI. Bioanalytical tools for the evaluation of organic micropollutants during sewage treatment, water recycling and drinking water generation. *Water Research* 2011; 45: 4238-47.
- Martinovic-Weigelt D, Mehinto AC, Ankley GT, Denslow ND, Barber LB, Lee KE, et al. Transcriptomic effects-based monitoring for endocrine active chemicals: assessing relative contribution of

- treated wastewater to downstream pollution. *Environmental science and technology* 2014; 48: 2385-94.
- McMaster ME, Munkittrick KR, Jardine JJ, Robinson RD and Van Der Kraak GJ. Canadian Technical Report of Fisheries and Aquatic Sciences. Protocol for measuring in vitro steroid production by fish gonadal tissue. 1995;1961: 78
- MDI Biological Laboratory and North Carolina State University. Comparative Toxicogenomics Database (CTD). <http://ctdbase.org/> Accession date: 7/10/2016.
- Miege C, Budzinski H, Jacquet R, Soulier C, Pelte T, Coquery M. Polar organic chemical integrative sampler (POCIS): application for monitoring organic micropollutants in wastewater effluent and surface water. *Journal of Environmental Monitoring* 2012; 14: 626-635.
- Oblinger-Childress CJ, Foreman WT, Connor BF, Maloney TJ. New reporting procedures based on long-term method detection levels and some considerations for interpretations of water-quality data provided by the U.S. Geological Survey National Water Quality Laboratory, 1999. <http://nwql.usgs.gov/rpt.shtml?OFR-99-193>
- Parks LG, Cheek AO, Denslow ND, Heppell SA, McLachlan JA, LeBlanc GA, et al. Fathead minnow (*Pimephales promelas*) vitellogenin: purification, characterization and quantitative immunoassay for the detection of estrogenic compounds. *Comparative biochemistry and physiology. Part C, Pharmacology, toxicology and endocrinology* 1999; 123: 113-25.
- Petrulis JR, Chen G, Benn S, LaMarre J, Bunce NJ. Application of the ethoxyresorufin-O-deethylase (EROD) assay to mixtures of halogenated aromatic compounds. *Environmental Toxicology* 2001; 16: 177-84.
- Schroeder AL, Ankley GT, Houck K, Villeneuve DL. Environmental surveillance and monitoring-The next frontiers for high-throughput toxicology. *Environmental Toxicology and Chemistry* 2016; 35: 513-525
- Shappell NW, Hyndman KM, Bartell SE, Schoenfuss HL. Comparative biological effects and potency of 17 α - and 17 β -estradiol in fathead minnows. *Aquatic Toxicology* 2010; 100: 1-8.
- Skelton DM, Ekman DR, Martinovic-Weigelt D, Ankley GT, Villeneuve DL, Teng Q, et al. Metabolomics for in situ environmental monitoring of surface waters impacted by contaminants from both point and nonpoint sources. *Environmental science and technology* 2014; 48: 2395-403.
- Smith MT, Guyton KZ, Gibbons CF, Fritz JM, Portier CJ, Rusyn I, et al. Key Characteristics of Carcinogens as a Basis for Organizing Data on Mechanisms of Carcinogenesis. *Environmental health perspectives* 2016; 124: 713-21.
- Stackelberg PE, Kauffman LJ, Ayers MA, Baehr AL. Frequently co-occurring pesticides and volatile organic compounds in public supply and monitoring wells, southern New Jersey, USA. *Environmental Toxicology and Chemistry* 2001; 20: 853-865.
- Sumpter JP, Jobling S. Vitellogenesis as a biomarker for estrogenic contamination of the aquatic environment. *Environmental Health Perspectives* 1995; 103: 173-178.
- Sutton D, Butler AM, Nadin L, Murray M. Role of CYP3A4 in human hepatic diltiazem N-demethylation: inhibition of CYP3A4 activity by oxidized diltiazem metabolites. *Journal of Pharmacology and Experimental Therapeutics* 1997; 282: 294-300.
- Tang JY, Aryal R, Deletic A, Gernjak W, Glenn E, McCarthy D, et al. Toxicity characterization of urban stormwater with bioanalytical tools. *Water Research* 2013; 47: 5594-606.
- Thorpe KL, Benstead R, Hutchinson TH, Tyler CR. Associations between altered vitellogenin concentrations and adverse health effects in fathead minnow (*Pimephales promelas*). *Aquatic Toxicology* 2007; 85: 176-183.
- United States Army Corps of Engineers. Evaluation of dredged material proposed for discharge in waters of the US-Testing manual. 1998; 823/B-98/004.
- United States Environmental Protection Agency. Development of water quality-based permit limitations for toxic pollutants. *Federal Register* 1984; 49:9016-9019.

- United States Environmental Protection Agency. Great Lakes Restoration Initiative (GLRI).
<https://www.epa.gov/green-bay-fox-river-aoc/about-lower-green-bayfox-river-aoc>. Accession date: 02/10/2016.
- United States Environmental Protection Agency. iCSS ToxCast Dashboard.
<http://actor.epa.gov/dashboard/>. Accession date: 02/06/2016.
- United States Environmental Protection Agency. Ambient water quality criteria for naphthalene.
<https://nepis.epa.gov/Exe/ZyNET.exe/2000LN3Z.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1976+Thru+1980&Docs=&Query=&Time=&EndTime=&SearchMethod=1&TocRestrict=n&Toc=&TocEntry=&QField=&QFieldYear=&QFieldMonth=&QFieldDay=&IntQFieldOp=0&ExtQFieldOp=0&XmlQuery=&File=D%3A%5Czyfiles%5CIndex%20Data%5C76thru80%5CTxt%5C00000003%5C2000LN3Z.txt&User=ANONYMOUS&Password=anonymous&SortMethod=h%7C-&MaximumDocuments=1&FuzzyDegree=0&ImageQuality=r75g8/r75g8/x150y150g16/i425&Display=hpfr&DefSeekPage=x&SearchBack=ZyActionL&Back=ZyActionS&BackDesc=Results%20page&MaximumPages=1&ZyEntry=1&SeekPage=x&ZyPURL> 1980.
- Villeneuve DL, Blake LS, Brodin JD, Cavallin JE, Durhan EJ, Jensen KM, et al. Effects of a 3beta-hydroxysteroid dehydrogenase inhibitor, trilostane, on the fathead minnow reproductive axis. *Toxicological sciences* 2008; 104: 113-23.
- Villeneuve DL, Blake LS, Brodin JD, Greene KJ, Knoebel I, Miracle AL, et al. Transcription of key genes regulating gonadal steroidogenesis in control and ketoconazole- or vinclozolin-exposed fathead minnows. *Toxicological sciences* 2007; 98: 395-407.
- Vrana B, Paschke A, Popp P, Schuurmann G. Use of semipermeable membrane devices (SPMDs). Determination of bioavailable, organic, waterborne contaminants in the industrial region of Bitterfeld, Saxony-Anhalt, Germany. *Environmental Science and Pollution Research International* 2001; 8: 27-34.
- Wagner RJ, Roberts LM, Geological S. Pesticides and volatile organic compounds in surface and ground water of the Palouse subunit, Central Columbia Plateau, Washington and Idaho, 1993-95. U.S. Geological Survey, 1998: 97-4285
- Wehmas LC, Cavallin JE, Durhan EJ, Kahl MD, Martinovic D, Mayasich J, et al. Screening complex effluents for estrogenic activity with the T47D-KBluc cell bioassay: Assay optimization and comparison with in vivo responses in fish. *Environmental Toxicology and Chemistry* 2011; 30: 439-445.
- Wilson VS, Bobseine K, Gray LE, Jr. Development and characterization of a cell line that stably expresses an estrogen-responsive luciferase reporter for the detection of estrogen receptor agonist and antagonists. *Toxicological sciences* 2004; 81: 69-77.
- Wilson VS, Bobseine K, Lambright CR, Gray LE, Jr. A novel cell line, MDA-kb2, that stably expresses an androgen- and glucocorticoid-responsive reporter for the detection of hormone receptor agonists and antagonists. *Toxicological sciences* 2002; 66: 69-81.
- Yilmaz A, Karacık B, Henkelmann B, Pfister G, Schramm KW, Yakan SD, et al. Use of passive samplers in pollution monitoring: A numerical approach for marinas. *Environment International* 2014; 73: 85-93.

1 Figure Legends

2 Figure 1. Maps of study sites at (A) Lower Green Bay/ Fox River and (B) Milwaukee
3 Estuary. In Figure 1A, DPERE and GRBAY are near municipal wastewater treatment
4 plants (WWTP) outfalls, PRGAM are near Proctor and Gamble Discharge a paper mill
5 discharge, and EASTR is in the East River. In Figure 1B, JONESI is at Jones Island
6 WWTP outfall, and KINNI, MENMR, and MILWR are located on Kinnickinic, Menominee,
7 and Milwaukee rivers, respectively. Maps were generated from Google Maps
8 (<https://www.google.com/maps/>).

9 Figure 2. Heat maps of 134 analyte measurements for each site, which includes all grab
10 (G) and composite (C) samples in Green Bay Area of concerns (AOC) (A) and
11 Milwaukee AOC (B). Samples are arranged left to right, and chemicals are arranged top
12 to bottom. The color bar at the top indicates the medium centered values ranging
13 from 0 to 83.75 (yellow). Note that for Green Bay AOC, no composite water sample was
14 analyzed due to volume restrictions. Also, chemicals were not shown with a medium-
15 centered value of 1 across all sites.

16 Figure 3. Bar plots of *in vitro* (A) 17 α -ethinylestradiol (EE2) equivalents and (B) 17 β -
17 trenbolone (TRB) equivalents of composite water samples (25% and 100%) from 8 sites
18 using the T47D-KBluc and MDA-kb2 assays, respectively. Horizontal dotted red line
19 indicates the significant response level of 3 standard deviations above the assay media
20 control mean. Bars represent mean ($\pm$ standard error), with $n = 3$ assay replicates per
21 sample. No statistically significant differences were observed ($p > 0.05$) for EE2 or TRB
22 equivalents among groups.

Figure 4. Boxplot of (A) adult female fathead minnow *ex vivo* gonad (or ovary) testosterone (T) synthesis, (B) adult female fathead minnow gonad *ex vivo* estradiol (E2) synthesis, (C) adult male fathead minnow testis or gonad *ex vivo* T synthesis, and (D) adult male fathead minnow plasma VTG protein concentrations (A, B, C: n=11-12; D: n=6-12;). For deployment c reasons, sites were grouped as follows: (G1) DPERE and EASTER, (G2) GRBAY and PRGAM, (G3) MENMR and MILWR, and (G4) JONESI and KINNI. Time zero (or however they were referred to in text) control fish (G0) were also investigated for each group. Gray bars are assigned to fish deployed in the field, and white bars are designated to time zero control fish. Significant ($p < 0.05$) differences across groups are indicated with letters. Different letters indicate statistically significant differences.

Figure 5. Expressions of hepatic transcripts for (A) vitellogenin (*vtg*) messenger RNA (mRNA), (B) cytochrome P450 (*CYP1a1 mRNA*), and (C) *CYP3A mRNA* in adult male fathead minnow (n=6-12; n represents sample size). For logistic reasons, sites are grouped as follows: (G1) DPERE-EASTER, (G2) GRBAY – PRGAM, (G3) MENMR–MILWR, and (G4) JONESI – KINNI. Control fish (G0) is investigated for each group. Gray bars are assigned to fish deployed in the field, and white bars are denoted to control fish. In Figure A, significant ($p < 0.05$) differences within G1 group are indicated with asterisk. In Figure B and C, significant ($p < 0.05$) differences across groups are indicated with letters. Different letters indicate statistically significant differences.

Figure 6. Heat map of the gene list with the corresponding number of chemicals detected in water samples that are known to interact with each gene, and pathway enrichment analysis of those genes for each site. Gene list was generated by compiling

genes that interacted with at least one chemical, together with the total number of chemicals interacting with that gene according to bio-effect predictions based on the comparative toxicogenomics database. Sites are arranged top to bottom on the left panel. Gene names and associated pathways are listed left to right on the bottom panel. The color bar on the left indicates the number of chemicals ranging from 0 (blue) to 19 (red) times. The 'Entity color by Pathway' panel on the right denotes critical enriched pathways with IDs from Kyoto Encyclopedia of Genes and Genomes (KEGG) and Reactome Pathway Database (REACT).

Figure 7. Barplot of exposure activity ratios (EAR). EAR values are presented on the y-axis. The color bar on the right indicates assay names. Site names are listed on the x-axis. Detailed descriptions of the assays are provided in Table S5. Note that EAR predictions based on the NovaScreen assay (NVS_NR_hPPARg) are not accurate and not included in the current barplot due to an inaccurate AC50 in the database as detailed in the manuscript.

61 Figures

62 Figure 1.

63
64

65 Figure 2.

66 Figure 2A.

67

68 Figure 2B.

69
70

71 Figure 3.

75 Figure 4.

76

77

78 Figure 5.

79

80

82

Figure 7.