Developing of a Quantitative Toxicogenomics-based Approach for Water Quality Monitoring and Toxicity Evaluation

A Dissertation Presented

By

Na Gou

to

Department of Civil and Environmental Engineering

in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

in the field of

Civil Engineering

Northeastern University
Boston, Massachusetts

(Dissertation completed in December 2015)
Abstract

The recognized and unknown risks associated with the ever-increasing number of pollutants in our environment presents a serious threat to us all. This poses a pressing need for a breakthrough in toxicity-assessment technology because the currently available methods are neither feasible nor sufficient to provide the timely information needed for regulatory decision-making and technology development to eliminate these threats. This study developed a novel, feasible and cost-effective quantitative toxicogenomics-based toxicity assessment platform for high-throughput and effective chemical hazardous identification and environmental toxicity monitoring. We systematically optimized the assay platform, evaluated its robustness and performance, validated the assay output and demonstrated its wide applications. Compared with other main stream ‘omics’ technologies, the proposed method greatly improves the feasibility and cost effectiveness as a result of its much simpler, faster, and more reliable assay procedures. Furthermore, it provides multi-dimensional transcriptional level effect information with a temporal dimension and therefore can more accurately reflect the chemical-induced time-dependent cell responses with higher sensitivity and specificity. We demonstrated that information-rich toxicogenomics data are powerful for evaluating toxic effects, understanding toxicity mechanisms, and obtaining pollutant-specific molecular fingerprints for compound/sample classification and identification. One of the main challenges in applying toxicogenomics for environmental monitoring is the lack of a quantitative method to convert the toxicogenomic information into a readily usable and transferable format that can be incorporated into ecological risk assessment and regulatory framework. We proposed a new transcriptional effect level index (TELI) that exhibited a dose-response relationship and allowed for linking the transcriptional level effects to conventional toxicity endpoints. In addition, we pioneered quantitative molecule toxicity modeling within the context of toxicogenomics and paved the road for further mixture toxicity identification and prediction. Cross-species comparison and extrapolation is another key aspect related to predictive and mechanistic toxicity assessment to overcome the limitation of data generation ability. We have compared three different species for variety of compounds and demonstrated the possibility of cross-species extrapolation with stress-response pathway ensemble based toxicity assessment. Finally, we demonstrated successful application of the novel assay for mechanistic CECs toxicity assessment, whole effluent toxicity monitoring and risk-based water treatment technologies efficacy evaluation.
Acknowledgements

The entire journey of PhD. for me has been a long one and during these seven years course of PhD., I have come across excitements, frustrations, stresses and enjoyment as well. Progressing towards the completion makes me put all those circumstances in retrospect when I look back and make me realize that I could not have reached here without going through any of those difficult or enjoyable moments. I would like to express my deepest and sincerest gratitude to my advisor Professor April Z. Gu who has also been my friend, my guide and the sole responsible person for my enhancement as a researcher and a learner. Each and every moment I have spent with her during this course of time, used to be a moment of encouragement and a positive perspective towards life and knowledge. Her continuous support for me and the trust that she bestowed on me gave me the courage to achieve my goals and discover my true potentials. I am really grateful to my co-advisor Professor Akram Alshawabkeh for being patient and supportive towards me and providing me with their valuable advices whenever I was seeking for him and this dissertation would not have been complete without his suggestions. I would like to thank to the others of my committee members, Professor Ferdinand L Hellweger, Professor Philip Larese-Casanova, and Professor Loretta Fernandez for guiding my research for the past several years and helping me to develop my background environmental hydrology, environmental chemistry, and biology.
I acknowledge financial support from National Science Foundation (NSF, EEC-0926284 and CAREER CBET-0953633), and National Institute of Environmental Health Sciences (NIEHS, Grant No.P42ES017198) that made this study possible. I am sincerely grateful to the Department of Civil & Environmental Engineering at Northeastern University for providing me with the supports needed through my PhD.

I would like to extend my gratitude towards Dr. Annalisa Onnis-Hayden, and Dr Jiaqi Lan, post doctoral research associates in our group who guided me immensely with her experience on the environmental engineering and biology laboratory methods for my research.

I am grateful to my dearest group colleagues at lab: Ce Gao, Sheikh Mokhlesur Rahman, Xin Wen, Yueyue Li, Nick Tooker and Man Hu who stood by me whenever I needed their help and support and made my PhD experience a memorable one. They helped me whenever I needed them in the lab. I am also grateful to all other fellow graduate students in our department who helped in releasing stresses off of the mind from time to time while chatting during working in the lab.

Finally, I would like to acknowledge my family members without whom the completion of this dissertation would be impossible.

Na Gou
Northeastern University
2015
# TABLE OF CONTENTS

ABSTRACT ................................................................................................................ii  
ACKNOWLEDGEMENTS ............................................................................................iii  
TABLE OF CONTENTS .............................................................................................v  
LIST OF FIGURES ..................................................................................................xi  
LIST OF TABLES .....................................................................................................xv

**Chapter 1. Introduction and Objectives** .................................................................1  
1.1 Background and motivations  
   1.1.1 Environmental Pollutions Threaten Human Health and Ecological integrity.....1  
   1.1.2 Large and Ever-Increasing Number of Unregulated Chemicals .................1  
   1.1.3 Lack of Toxicological Information Hampers Chemical Management and Risk Mitigations.................................................................5  
   1.1.4 Challenges in Environmental Risk Assessment ...........................................6  
   1.1.5 Current toxicity assessment approach and limitations .............................9  
   1.1.6 Urgent Need in Risk-based Approach for Environmental Monitoring and Technology Assessment .................................................................10  
   1.1.7 Vision of Tox21 Program - Toxicity Testing in 21st Century ...................12  
   1.1.8 Toxicogenomics is a Promising Tool for Risk Assessment .....................14  
1.2 Research Objectives and Overview ..................................................................18  
1.3 Organization of Thesis .......................................................................................30

**Chapter 2. Advances in Quantitative Toxicogenomics for Environmental Applications on Water Quality Monitoring** .........................................................34  
2.1 Abstract  
2.2 Introduction ........................................................................................................35  
   2.2.1 The Emerging Field of Toxicogenomics ....................................................37  
      2.2.1.1 Common Toxicogenomics Technologies ..........................................38  
      2.2.1.2 Key Challenges on Applying Toxicogenomics Approach for Regulatory Risk Assessment .................................................................42
2.3 The Concept of Adverse Outcome Pathway (AOP) in Toxicogenomics

2.3.1 AOP Concept

2.3.2 AOP Development Status

2.3.3 Limitation of Current AOP for Ecotoxicity Assessment

2.3.4. Stress Response Pathways Ensemble-Based Toxicogenomics Assay

2.3.4.1 Highly conserved Stress Response Pathways Across Species

2.3.4.2 Cellular Stress Pathways Ensemble-based Toxicogenomics Application

2.3.4.3 Anchoring of Stress Response via Toxicogenomics Approach to Adverse Outcome

2.4. Quantitative Analysis of Toxicogenomics Data

2.4.1 Cellular Response Signal as Quantifier

2.4.2 Derivation of Molecular Toxicity Endpoints

2.4.3. Mixture Toxicity Assessment

2.5. Environmental Application of Toxicogenomics for Water Toxicity Assessment and Quality Monitoring

Chapter 3. Toxicogenomics approach - Assay development of Whole-cell-array with stress gene ensemble

3.1 Abstract

3.2 Introduction and Objectives
3.3 Materials and Methods ................................................................. 81
  3.3.1 Nanomaterials tested for toxicity ........................................... 81
  3.3.2 Measuring the temporal gene transcriptional activity in E.coli. ....... 82
  3.3.3 Concept and the determination of TELI .................................... 85
  3.3.4 Determination of EC50 and NOTEL ........................................ 86
3.4 Results and Discussion ............................................................... 87
  3.4.1 TELI as a quantitative toxicogenomic-based toxicity assessment
        endpoint .................................................................................... 87
  3.4.2 The correlation between TELI with other established toxicity
        endpoints ................................................................................... 96
  3.4.3 TELI for effects associated with different stress functional categories
        ....................................................... 99
3.5 Conclusions ................................................................................ 102

Chapter 4. Mechanistic Toxicity Assessment and Classification of Environmental
Contaminants of Emerging Concern Via Stress Response Ensemble Based
Toxicogenomics Assay ................................................................. 104
4.1 Abstract ...................................................................................... 106
4.2 Introduction and Objectives .......................................................... 109
4.3 Materials and Methods ............................................................... 109
  4.3.1 Chemical selection and evaluation ............................................ 109
  4.3.2 Toxicogenomics assay ............................................................ 109
  4.3.3 Dose-response and toxicogenomics endpoints derivation ......... 111
  4.3.4 Stress gene set enrichment analyses (Stress-GSEA) ............... 112
  4.3.5 Gene ontology analyses ......................................................... 112
  4.3.6 Clustering analysis ................................................................. 113
4.4 Results and Discussion ............................................................... 114
  4.4.1 Toxicogenomics assay reveals Toxicity mechanisms of CECs ....... 114
  4.4.2 Molecular Toxicity Endpoints Derivation and Comparison among CECs. 134
  4.4.3 Consensus Cluster ................................................................. 143
4.5 Conclusions ................................................................................ 146
Chapter 5. Mixture Toxicity Prediction via Quantitative Toxicogenomics-based Approach

5.1 Abstract
5.2 Introduction and Objectives
5.3 Materials and Methods
  5.3.1 Chemical selection and mixture combination design relational
  5.3.2 Toxicogenomics assay with stress response pathway ensemble
  5.3.3 Dose-response and toxicogenomics endpoints derivation
  5.3.4 Gene set enrichment analysis
  5.3.5 Principal components analysis
  5.3.6 Molecular mixture toxicity prediction model (CA and IA model)
5.4 Results and Discussion
  5.4.1 Comparison of toxicity profiles of individual components with their mixture
  5.4.2 Comparison of mixture with their components via Principal Components Analysis
  5.4.3 Dose-response patterns of mixture and its chemical components
  5.4.4 Quantitative prediction of combined effect at molecular level and comparison with observations
  5.4.6 Mixture prediction on specific stress pathways
5.5 Conclusion

Chapter 6. Link quantitative toxicogenomics endpoints to conventional toxicity endpoints - phenotypic anchoring and cross species extrapolation

6.1 Abstract
6.2 Introduction and Objectives
6.3 Materials and Methods
  6.3.1 Chemical selection and evaluation
  6.3.2 Toxicogenomics assay
  6.3.3 Dose-response and toxicogenomics endpoints derivation
6.3.5 Phenotypic assays

6.4 Results and Discussion

6.4.1 Integrated molecular endpoints based on cellular stress response biomarkers ensemble relate to cellular adverse outcomes EC50

6.4.2 Cross species extrapolation- correlation of molecular endpoints based on stress response ensemble in E coli to EC50 of prokaryotic bacterial V.fisheri luminescence inhibition (Microtox)

6.4.3 Cross species extrapolation - correlation of molecular endpoints based on stress response ensemble in E coli to EC50 with lower eukaryotic, green alga, S.capricornutum 96 hours growth inhibition

6.4.4 Cross species extrapolation - correlation of molecular endpoints based on stress response ensemble in E coli to EC50 with higher eukaryotic fish cell line RTgill W-1 cytotoxicity

6.5 Conclusion

Chapter 7. Mechanistic Toxicity Assessment of Nanomaterials by Whole-cell-array

Stress Genes Expression Analysis

7.1 Abstract

7.2 Introduction and Objectives

7.3 Materials and Methods

7.3.1 Nanomaterials

7.3.2 Prokaryotic stress genes cell-library with transcriptional GFP fusions

7.3.3 Measuring the temporal gene expression upon chemical exposure

7.3.4 Data processing and analysis

7.4 Results and Discussion

7.4.1 Distinctive and Complex Real-time gene-expression profiles for nAg and nTiO2_a

7.4.2 Concentration-dependent expression profiles

7.4.3 Determination of No Observed Transcriptional Effect Level (NOTEL)

7.4.4 Toxicity mechanism of nAg

7.4.5 Toxicity mechanism of nTiO2_a
7.4.6 Candidate biomarkers for nAg and nTiO2-a exposure..................229
7.5 Conclusion..................................................................................230

Chapter 8. Toxicity reduction evaluation for municipal wastewater treatment

plants.................................................................232
8.1 Introduction .................................................................232
8.2 Materials and Methods......................................................245
8.3 Results and Discussions ......................................................239
  8.3.1 Excellent plant performance for both plants .........................239
  8.3.2 Temporal Toxicity Evolution and Profile Changes during the CECs
  Degradation .................................................................240
  8.3.3 Toxicogenomics Assay Reveals the Nature and Dynamics of Toxicity
  Evolution during the Wastewater Treatment Process .........................242
8.4 Conclusion ...............................................................................244

Chapter 9. Quantitative Toxicogenomics Assay Reveals the Evolution and Nature of
Toxicity during Environmental Pollutants Transformation ......................246

9.1 Abstract
9.2 Introduction and Objectives ....................................................248
9.3 Materials and Methods..........................................................252
  9.3.1 Chemicals .................................................................252
  9.3.2 Electro-Fenton Degradation of CECs ..................................252
  9.3.3 Transformation Products Identification and Chemical Analysis .......253
  9.3.4 Sample Preparation for Toxicity Assessments .........................253
  9.3.5 Microtox Assay ................................................................254
  9.3.6 Toxicogenomics Assay and Data Computation .........................254
  9.3.7 Gene Enrichment Analysis ...............................................256
9.4 Results and Discussion ..........................................................257
  9.4.1 Temporal Toxicity Evolution and Profile Changes during the CECs
  Degradation. .............................................................................257
Chapter 10. Quantitative Toxicogenomics Assay Revealed the Impact of Disinfection Technologies On Effluent Toxicity .............................................273

10.1 Abstract
10.2 Introduction and Objectives ..........................................................274
10.3 Materials and Methods.................................................................279
   10.3.1 Wastewater Samples Collection and Preparation .................279
   10.3.2 Microtox Assay .................................................................280
   10.3.3 Toxicogenomics assay with stress response pathway ensemble ...281
10.4 Results and Discussion..................................................................283
   10.4.1 Removal of detected CECs through different disinfection processes.....283
   10.4.2 Baseline toxicity revealed by Microtox assay..........................286
   10.4.3 Overall effluent toxicity revealed by quantitative toxicogenomics assay288
   10.4.4 Quantitative toxicogenomics assay revealed distinct toxicity profile and
       nature among effluents of different disinfection processes......................292
10.5 Conclusion..................................................................................298

Chapter 11. Conclusions and Future Recommendations .......................299

11.1 Conclusions...............................................................................299
11.2 Recommendations for Future Studies.............................................303

REFERENCE

APPENDIENCE
LIST OF FIGURES

Chapter 1. Introduction and Objectives

Figure 1.1: Overview of the Objectives and Components of this study ………..20
Figure 1.2: Overview of the approaches and methodologies of this study………21

Chapter 2. Advances in Quantitative Toxicogenomics for Environmental Applications on Water Quality Monitoring

Figure 2.1: Examples of a biomarker star plot for four biological markers on mussel (Mytilus edulis) collected in the Baltic sea ……………………………..65

Chapter 3. Toxicogenomics approach - Assay development of Whole-cell-array with stress gene ensemble

Figure 3.1: Determination of Transcriptional Effect Level Index (TELI), as an endpoint for toxicogenomic based toxicity assessment. …………………………87
Figure 3.2: 3-D altered gene expression profile compiled from the altered gene expression level for all the genes in the test library ……………………………..89
Figure 3.3: Dose response curves based on TELI versus dose concentration for nAg, nTiO2_a, nTiO2_r, and fullerene, fitted by Four-parameter Logistic Equation ……………………………………………………………91
Figure 3.4 Transcriptional effect Level Index (TELI) values determined for toxicity assay for nAg, nTiO2_a, nTiO2_r, and fullerene at various dose concentrations with exposure time of 2 hrs………………………………………92
Figure 3.5 Comparison of TELI values for different stress gene functional categories among the four NMs studies……………………………………100
Chapter 4. Mechanistic Toxicity Assessment and Classification of Environmental Contaminants of Emerging Concern Via Stress Response Ensemble Based Toxicogenomics Assay

Figure 4.1: Toxicity profile based on the quantitative endpoints of different stress response categories of tested chemicals at various concentrations. ..................117
Figure 4.2: Toxicity profiles at selected concentration for tested chemicals to represent the variation among different categories. .........................119
Figure 4.3: Dose-response curves based on TELI values versus dose concentration for tested chemicals, fitted by the Four-parameter Logistic Equation. ..........137
Figure 4.4: Consensus clustering matrix of tested chemicals with different concentrations.................................................................146

Chapter 5. Mixture Toxicity Prediction via Quantitative Toxicogenomics-based Approach

Figure 5.1: Comparison of temporal altered stress response gene expression profiles of mixture with that of each individual chemical component. ............162
Figure 5.2: PCA Graphical representations of mixture and their single components, based on Induction factor (I) during two hour exposure with six concentrations. Samples are color coded according to chemicals and each spot represents one treatment (a chemical at a given concentration) with bigger symbol indicating relative higher level of exposed concentration........................................................................167
Figure 5.3: Dose-response analysis of mixture with their components............172
Figure 5.4: Mixture toxicity values indicated by probit unit. The open hole diamond indicates the observed values from toxicogenomics results, the red dashed line means the predicated values from CA model while blue are from IA model. .................................................................176

Chapter 6. Link quantitative toxicogenomics endpoints to conventional toxicity endpoints - phenotypic anchoring and cross species extrapolation
Figure 6.1: Correlation between TELI derived endpoints with EC50 based in E.coli 24 hours growth inhibition, log scale are used for all concentration-based endpoints………………………………………………………………………..198
Figure 6.2: Correlation between TELI derived endpoints with EC50 based in Microtox, log scale are used for all concentration-based endpoints. …………202
Figure 6.3: Correlation between TELI derived endpoints with EC50 based on Selenastrum 96 hours growth inhibition, log scale are used for all concentration-based endpoints ……………………………………………………………………205
Figure 6.4: Correlation between TELI derived endpoints with EC50 based on Fish Cell RTgill-W1 24hours cytotoxicity, log scale are used for all concentration-based endpoints. ……………………………………………………………………207

Chapter 7. Mechanistic Toxicity Assessment of Nanomaterials by Whole-cell-array
Stress Genes Expression Analysis
Figure 7.1: Real-time (temporal) gene expression profiles of 91 stress genes in E.coli in exposure to nAg 10mg/L (left) 50mg/L ………………………………..218
Figure 7.2: Concentration-dependent responses to NMs. (A) Number of differential expressed genes at each exposure concentration. (B) Dose-response curve based on the portion of differentially expressed genes upon exposure to NMs in our stress library as a function of NMs concentrations. ……………….221
Figure 7.3: Toxicity mechanisms revealed by the altered gene expression level for genes involved in different stress functional categories in exposure to nAg (10 mg/L)………………………………………………………………………………..225
Figure 7.4: Oxidative stress response regulator network upon exposure to nAg based on temporally dynamic gene expressions obtained in our study. ……..226

Chapter 8. Toxicity reduction evaluation for municipal wastewater treatment plants
Figure 8.1: Performance data for two wastewater treatment plants. Dashed lines indicate data missing for corresponding processes. …………………………..239
Figure 8.2: Toxicity changes along the WWTP processes for both plants. Y axis left: the chronic toxic unit as 1/EC10 with green alga selenastrum 96 hours growth inhibition test. .................................................................240

Figure 8.3: Toxicity changes along the WWTP processes for both plants. …242

Figure 8.4: Exemplary temporal toxicity evolution and profile changes along the WWTP processes for both plants as well as the enrichment blank control (at REF=1, the original concentration).........................................................243

Chapter 9. Quantitative Toxicogenomics Assay Reveals the Evolution and Nature of Toxicity during Environmental Pollutants Transformation

Figure 9.1: Toxicity changes during the electro-Fenton transformation of (a) bisphenol A, (b) triclosan and (c) ibuprofen. The transformation conditions are based on 5 mg/L of initial concentration for each CEC, 6.9 mg/L Fe(II), pH 3, 40 mA, and 2 mM Na2SO4 background electrolyte. ....................................................258

Figure 9.2: Exemplary temporal toxicity evolution and profile changes during the electro-Fenton oxidation transformation of the three CECs (at REF=1, the original concentration). Profiles for other concentrations (REF) are not shown. ………261

Figure 9.3: Toxicity nature profiles based on the quantitative endpoints of different stress response categories for the three CECs, (a) bisphenol A, (b) triclosan and (c) ibuprofen, during the electro-Fenton transformation process...263

Figure 9.4: Stress response pathways and biomarkers genes that showed altered expression (TELI_gene > 1.5) during the electro-Fenton transformation of the three CECs. (a) oxidative stress pathways, (b) DNA damage and repair pathways, and (c) membrane stress response pathways. .................................................268

Chapter 10. Quantitative Toxicogenomics Assay Revealed the Impact of Disinfection Technologies On Effluent Toxicity

Figure 10.1. Removal efficiency of 23 detected chemicals by 8 disinfection technologies: UV, UV/H2O2, O3 and Cl2.................................................................285

Figure 10.2 Comparison of toxicity level among effluents treated with different disinfection processes.................................................................288
Figure 10.3 Total TELI values for wastewater secondary effluent samples treated by various disinfection processes when REF=1 and REF=10..........................291
Figure 10.4 Hierarchical clustering of toxicogenomics test results expressed as TELI for all water samples at REF=1.................................................................295
Figure 10.4 TELI values of each stress category for wastewater secondary effluent samples treated by various disinfection processes............................................297

LIST OF TABLE

Chapter 2. Advances in Quantitative Toxicogenomics for Environmental Applications on Water Quality Monitoring
Table 2.1: List of Studies Using Stress Response and Pathways-Related Biomarkers for Ecotoxicity Assessment .................................................................49
Table 2.2: Applications of Toxicogenomics Technologies In Water Toxicity Assessment .................................................................................................74

Chapter 3. Toxicogenomics approach - Assay development of Whole-cell-array with stress gene ensemble
Table 3.1: Physical and chemical Characterization of the nanomaterials tested in this study ............................................................................................82
Table 3.2: Genes included in the stress cell-library for chemical-induced real time gene expression profiling analysis ..................................................83
Table 3.3: Summary of TELI-based toxicity characterization parameters and endpoints..................................................................................................94
Table 3.4 Correlation coefficients between TELI-based toxicity endpoints with other toxicity endpoints .................................................................96
Chapter 4. Mechanistic Toxicity Assessment and Classification of Environmental Contaminants of Emerging Concern Via Stress Response Ensemble Based Toxicogenomics Assay

Table 4.1: Summary of toxicity mechanism (MOA) from literature and our toxicogenomics assay.................................................................121
Table 4.2 Summary of TELI-dose response curve derived toxicogenomics endpoints..................................................................................140
Table 4.3 Correlation coefficient between TELI-based toxicogenomics endpoints with conventional toxicity endpoint....................................................142

Chapter 5. Mixture Toxicity Prediction via Quantitative Toxicogenomics-based Approach

Table 5.1: Overrepresented (p-value < 0.05) stress categories by stress gene set enrichment analysis.................................................................164
Table 5.2: The predication r square for measure of regression models..........177
Table 5.3 The prediction r square for measure of regression models for specific stress pathways........................................................................178

Chapter 6. Link quantitative toxicogenomics endpoints to conventional toxicity endpoints - phenotypic anchoring and cross species

Table 6.1: Chapter 5. Mixture Toxicity Prediction via Quantitative Toxicogenomics-based Approach.................................................................209

Chapter 8. Toxicity reduction evaluation for municipal wastewater treatment plants

Table 8.1 Information of Wastewater Treatment Plant........................................235

Chapter 10. Quantitative Toxicogenomics Assay Revealed the Impact of Disinfection Technologies On Effluent Toxicity

Table 10.1 Removal efficiency of 23 detected chemicals by 8 disinfection technologies: UV, UV/H₂O₂, O₃ and Cl₂......................................................285
Chapter 1

Introduction and Objectives

1.1 Background and motivations

1.1.1. Environmental Pollutions Threaten Human Health and Ecological integrity

Ecosystem and human health are the most important indicators of sustainable development (Gleick et al., 2014). The World Health Organization (WHO) estimates that thirteen million deaths annually are attributable to preventable environmental causes (Kimani, 2007; Global Ob-Gyn Group, 2015; Benetti, 2009). The report also estimates that 24% of the global disease burden (healthy life years lost) and 23% of all deaths (premature mortality) are attributable to environmental factors, with the environmental burden of diseases being 15 times higher in developing countries than in developed countries, due to differences in exposure to environmental risks and access to health care (United Nations Environmental Programme, 2015, Global Ob-Gyn Group, 2015, WHO, 2009). Environmental pollution is causing a lot of distress not only to humans but also ecosystem, driving many species to endangerment and even extinction (Kimani, 2007). Scientists estimate we’re now losing species at 1,000 to 10,000 times the background rate, with literally dozens going extinct every day (Chivian et al., 2008). It could be a scary future indeed, with as many as 30 to 50 percent of all species possibly heading toward extinction by mid-century (Chris et al., 2004).

1.1.2 Large and Ever-Increasing Number of Unregulated Chemicals

Environmental pollution includes potentially harmful chemicals, physical, and biological agents that pose harmful impacts on life and the environment. In the last 50
years, chemicals have become ubiquitous integrated into nearly all industrial processes and commercial products and made people unavoidable to contact with them in the workplace, in homes, through the use of products, and via air, water, food, and waste streams (Schwarzenbach et al., 2006; Sedlak et al., 2005; Snyder et al., 2008). Now there are more than 50 million different chemicals on earth, ever increasing large number of new chemicals are on their way, a novel substance is either isolated or synthesized every 2.6 seconds (Madrigal, 2009).

Current chemical control acts or laws, such as US-Toxic Substances Control Act (TSCA) (USEPA, 2002), EU- Registration, Evaluation and Authorization of Chemicals (REACH) (“Regulation (EC) No 1907/2006”, 2006), CHINA Inventory of Existing Chemical Substances (IECSC) (“Chinese Chemical Inventory”, 2013), Korea Toxic Chemicals Control Act (TCCA) (“Korea Toxic Chemicals”, 1991), etc., aim to regulated new and existing chemicals’ distribution and use. However, most of these acts or laws make guidelines and decisions based on modeling (structure-activity relationships) (Applegate, 2008) and/or very limited acute toxicity information voluntarily provided by manufacturers if available. For example, less than 40% new chemicals have information on acute toxicity with rat, mice, rabbit, etc., and mutagenicity information, even less data available on long-term effects or specific endpoints, such as sub-chronic, chronic neurotoxicological, developmental and reproductive effects. And less than 5% of available toxicity data are on aquatic organisms (Bergeson et al., 2000). Furthermore, only a small portion of chemicals are regulated or registered by these acts or laws, for example, about ~84,000 chemicals in TSCA, ~150,000 by EU-REACH, and ~45,000 in CHINA IECSC which account for less than 3% of the total chemicals.
There has been an increasing concern raised from the recognition that a large number (e.g., more than 87,000 EDC candidates are identified by the EPA (Hong et al., 2002; Patlak et al., 1996)) of not commonly regulated yet widely presented chemicals or microorganisms pose the potential to cause known or suspected adverse ecological and/or human health effects; such chemicals are referred to as “contaminants of emerging concern” (CECs) (Brooks et al., 2002; Lapworth et al. 2012; Jones-Lepp, 2012). CECs encompass a wide variety of chemicals including pharmaceuticals and personal care products, endocrine disrupting compounds, disinfection byproduct, nanomaterials, etc. CECs are ubiquitous detected worldwide in environment, as water, soil and air. Different countries have initiated variety of interim programs for monitoring of CECs. In the United States, the EPA’s Contaminant Candidate List (CCL) (USEPA, 2015) and Unregulated Contaminants Monitoring Rule (UCMR) (USEPA, 2012) allow the agency to monitor a number of CECs, as well as determine appropriate regulations. In EU, a watch list of with 33 CECs is elaborated from national monitoring programs for their high frequency of occurrence, the expected risk for human health and/or aquatic life, and/or for a lack of analytical techniques (“Directive 2000/60/EC”, 2000). However, only dozens or hundreds of CECs are included in these lists, large number of CECs is still waiting for actions. A recent national-scale examination of CECs by the Toxic Substances Hydrology Program of the U.S. Geological Survey (USGS) showed that one or more of CECs were found in 80 percent of the streams sampled at low concentrations (in most cases, less than 1 part per billion) (Kolpin et al., 2002). Half of the streams contained 7 or more of CECs, and about one-third of the streams contained 10 or more of CECs (Kolpin et al., 2002).
The conventional wastewater and water treatment is not designed to effectively remove these CECs, therefore providing limited reduction on these CECs, leading to discharge of these CECs into receiving streams and water bodies, some of which maybe drinking water sources (Petrović et al., 2003). Their harmful effects on human, animal, and aquatic lives have already become manifest (Lei et al., 2015; Diamond et al., 2011; Pal et al., 2010; Xi et al., 2009; Sanderson et al., 2009). Antibiotic resistance is another aspect that has received attention recently suggesting the possible role of CECs in widespread resistance phenomena (Li et al., 2016).

Although great progress has been made toward better understanding of the occurrence, sources, and potential environmental and health effects from these CECs, one main challenge that remains is the lack of a feasible and accepted method for assessing and quantifying the toxicity exerted by these pollutants in water. This lack greatly hampers the development and implementation of effective regulations, strategies, and technologies to control and eliminate the harmful effects from these CECs. The challenge arises from the following aspects unique to the CECs. First, the toxicity of most CECs is unknown, and their harmful effects often exist at very low concentrations, sometimes lower than the current detection limit (Snyder et al., 2004; Kidd et al., 2004; Aardema et al., 2002; Krewski et al., 2010). Second, some CECs are not responsive to traditional toxicity assays, and their dose-response curves do not necessarily follow the classic assumption that an increase in dose increases the effect (Vandenberg et al., 2012): their effects may disappear or become qualitatively different at higher levels. Third, these chemicals often occur in mixtures in the environment and they can interact additively, synergistically, or not at all (Trine et al, 2010; Murray et al., 2010; Buxton et al., 2007;
Sedlak et al., 2005; Snyder et al., 2008). Finally, the large and increasing number of CECs (e.g., more than 87,000 EDC candidates are identified by the EPA (USEPA, 1998; Chang et al., 2009) makes the conventional labor- and resource-intensive method of monitoring water quality neither feasible nor economical. All these issues point to the pressing need for more sophisticated and informative, yet feasible and reliable assessment methods to detect and evaluate the toxicity effects of these CECs so that their risk to the public and environment can be understood and eliminated.

1.1.3 Lack of Toxicological Information Hampers Chemical Management and Risk Mitigations

Currently, for most chemicals (>90%), little or marginally amount of toxicological information, particularly long-term chronic effect of low level exposure, is available (USEPA, 1998). According to National Academy of Sciences (NAS), no toxicity information was available for more than 80% of all toxic substances in commerce, and only 22% of high production volume (HPV) chemicals had the minimal toxicological data available while most of them based on only acute toxicity tests. Only less than 7% of the HPV chemicals registered today have relatively more complete safety information (USEPA, 1998). Based on U.S. Food and Drug Administration's (FDA) database, 78.4% of additives directly added to food do not have sufficient toxicological data for even estimation of safety exposure level and, 93% of those additives are lacking more long-term and systematic toxicity information, such as reproductive or development toxicity testing (Neltner et al., 2013).

Many of the chemical safety decisions were made decades ago often based on extrapolations from limited data (US Congress, 1958) or on toxicology information
submitted by industry (Sylvain, 2002) rather than with independent scientific views, and only focused on massive produced chemicals. One of the main reasons for the lack of toxicological information for most chemicals is due to the limitations associated with current toxicological approach that rely on resource-intensive and laborious toxicity testing using laboratory animals, such as those described in the standard guidelines for the testing of chemicals from the Organisation for Economic Co-operation and Development (OECD) (FDA, 2002; Hallagan and Hall, 2009; Maffini et al., 2013, OECD, 1991). The limited resources and time do not allow high-throughput assessment of large number of chemicals and, in addition, the species- and test-specific endpoints are difficult to be extrapolated and applied to a broader universe of chemicals, life stages, and health effects. Additional reasons for the lack of toxicological information for chemicals in commerce include: 1) government agencies lack the enforcement powers and resources to fill the information gaps; 2) manufacturers usually have no incentives to add additional toxicology information once a chemical is approved. The lack of toxicology information and feasible assessment tools make chemicals management and risk mitigation, which are responsible for reducing, eliminating or preventing risks to the environment and human health, an very challenging and near unrealistic goal. Therefore, urgent actions are needed to effectively and efficiently fill these significant information gaps to ensure that public health and ecosystem security are protected (Gilbert, 2011).

1.1.4 Challenges in Environmental Monitoring

Pollutants in the environmental pose potential risks to both human and ecological system. Environmental risk assessment based on toxicity characterization (hazard assessment) and exposure assessment is of great importance for protecting human health
as well as ecosystem security from chemical contaminants that may be present in the environment (Schwarzenbach et al., 2006; Sedlak et al., 2005; Snyder et al., 2008). Exposure assessment is the quantitative or qualitative evaluation of chemicals contacts with human or other species. It describes the intensity, frequency, and duration of contact, and often evaluates the chemical intake or uptake rates, the exposure route (e.g., dermal, oral, or respiratory), and the resulting amount of the chemical that actually crosses the boundary and the amount absorbed (internal dose) (USEPA, 1992; Paustenbach et al., 2006). Exposure assessment of external concentrations via different route is generally determined using various chemical and physical analysis (NRC, 1997). Exposure assessments of internal doses of human body or other species could to be determined either by directed analysis of tissue/organ concentration (USEPA, 1992; Paustenbach et al., 2006), or by modeling tools, such as pharmacokinetic (PK) and pharmacodynamic (PD), which are powerful in quantitative risk assessment for the target tissue dose through integration of information on the external dose, the physiological structure of the animal, and the biochemical properties of organic acid toxicants (Kim et al., 2001).

Environmental monitoring of the contaminants in different media (water, soil, air) is essential for risk assessment. Environmental monitoring traditionally focuses only on chemicals that are considered responsible for the most significant human and environmental risks and for which regulatory benchmarks exist (“Public Law 92-500”, 1972; USEPA, 2015b; USEPA, 2015d). For water quality monitoring, for example, National Primary Drinking Water Regulations (NPDWRs) by USEPA are organized in 6 groups only with a total of 81 chemicals (3 disinfectants, 5 disinfectants by-products, 16
inorganic chemicals, 53 organic chemicals and 5 radionuclides) (USEPA, 2015e). If following the similar approach for regulation, it is foreseeable that the resources and efforts needed to monitor the ever-increasing number of identified harmful pollutant will continue to increase, and eventually will become un-feasible and unrealistic to implement.

It has been increasingly recognized that the current environmental monitoring approaches that rely on chemical and physical monitoring of targeted chemicals along are insufficient because they neither discerns the differences in exposure and effects nor considers the complex and broader risks that mixtures of contaminants and their transformation products, pose to the environment and human health (NRC, 2011). The evaluation of exposure need to be combined with toxicity information (that is, the expected response to a given level of exposure) as a “bottom-up” (from individual components to overall effect prediction) approach to predict the probability, nature, and magnitude of the adverse health effects that may occur. However, the presence of untargeted or non-detectable yet influential chemicals, lack of chemicals toxicology information and knowledge gaps on mixture effects (i.e. synergistic effects) hinders this “bottom-up” (from known individual components to overall effect prediction) approach for environmental risk assessment prediction with known chemicals compositions.

As an alternative, an overall and more realistic environmental toxicity assessment has been proposed for environmental samples referred as a “top-down” (from overall effect to hazard prioritization) approach. This approach employed effect-based analysis that often involving bioassays as an established alternative for hazard assessment of complex environmental mixtures. This approach provides an integrative parameter for the
presence of compounds affecting the applied test system without prior knowledge of chemical composition and concentrations, and avoids the need to evaluate the potential synergistic, and antagonistic effect.

Federal agencies and international organizations—including EPA (USEPA, 2002; USEPA, 1991), the National Toxicology Program (NTP) (Chhabra et al., 1990), the Food and Drug Administration (FDA, 1997; FDA, 2004), and the Organization for Economic Co-operation and Development (OECD, 2005) —developed documents that provide guidance on effect-based analysis for toxicity assessment and environmental monitoring.

### 1.1.5 Current toxicity assessment approach and limitations

Current toxicity assessment for chemical assessment and water quality monitoring are usually following the international standard guideline from OECD (Section 4: Guidelines for the Testing of Chemicals) (OECD, 1991), and International Organization for Standardization (ISO)’s water standards (ISO, 2012). According to these standard guidelines, toxicity testing on human and ecosystem health often involve studies of whole animals, typically rats, mice, dogs, and rabbits, or environmental relevant species such as fish, shrimp, daphnia for ecotoxicity assessment. Exposures can range from short-term acute toxicity (for example, an hour) to long-term chronic toxicity (for example, 2 years) and be continuous or episodic or consist of a single event. Tests may focus on a particular life stage, sex, or condition of exposure. The effects evaluated can be numerous and can include such diverse outcomes as subtle behavioral changes, impairment of reproduction, abnormal development, organ toxicity, cancer, and death (NRC, 2011).

The anatomical similarities of whole animal to humans and their susceptible to many common health problems make the whole organism integration approach more reliable,
more relevant and higher certainty (NRC, 2011). However, using whole animal test as routine application for chemicals assessment and environmental monitoring is limited by high cost, intensive labor, sophisticated analytical procedures, and long test durations (up to weeks or months) (Krewski et al., 2010). It could not meet the demand for risk assessment of large and ever-increasing number of chemicals as well as environmental samples in various environmental matrixes. For example, presently, the Whole Effluent Toxicity (WET) test is used for evaluating the impact of a discharge stream on the water environment by assessment of the lethal, reproductive, and either acute or chronic effects on selected indigenous organisms in the water (USEPA, 2002). However, application of the WET test has been limited because of its labor intensiveness, high cost, and long turn-around time.

In addition, tested animal are usually exposed to much higher doses than would be expected for typical environmental exposures, requiring assumptions about effects at lower doses or exposures (Krewski et al., 2010; NRC, 2007). Moreover, current toxicity tests provide little information on modes and mechanisms of action, which are critical for understanding interspecies differences in toxicity, and limit the ability to predict toxicity about chemicals that have not been tested (Krewski et al., 2010; NRC, 2007).

1.1.6 Urgent Need in Risk-based Approach for Environmental Monitoring and Technology Assessment

Since the ultimate goal of any remediation or water treatment technology is to eliminate the toxicity threats and risks associated with the contaminants, monitoring and understanding the toxicity evolution during remediation and treatment process is of great importance. Treatment processes not only transform the targeted chemical but also
impact the properties of the water matrix. Current water and wastewater regulations and accepted standard remediation and treatment effectiveness assessment procedures are based on concentration measurements of only targeted contaminants in reference to the regulation limits (Schwarzenbach et al., 2006; Sedlak et al., 2005; Snyder et al., 2008). This approach is insufficient because the large numbers of emerging contaminants are present as mixtures in the environment and they pose threat beyond the impacts of those targeted chemicals (Krewski et al., 2010; NRC, 2007). The presence of transformation products (either naturally via photo-oxidation, biotransformation or through engineered process, such as wastewater treatment, and remediation process) makes the problem even worse (Cwiertny et al., 2014). While the majority of generated transformation products are likely benign, substantial conservation of structure in transformation products can imply conservation or even creation of bioactivity across multiple biological endpoints and thus incomplete mitigation of ecological risk (Cwiertny et al., 2014). The potential toxic evolution and elevated toxicity effects of mixtures of transformation products during water treatment processes such as advanced oxidantion processes (AOPs), for example, have been reported (Cwiertny et al., 2014; Heringa et al., 2011; Stalter et al., 2010, Gou et al., 2013). There is an urgent need for a risk-based, practical and cost-effective treatment efficacy assessment method to allow comprehensive evaluation of the consequent impact on the overall water toxicity, directly or indirectly, resulting from various remediation technologies, directly enabling real time monitoring to achieve the goal of risk reduction.
1.1.7 Vision of Tox21 Program - Toxicity Testing in 21st Century

The growing evidence and severely rising public concern of the far-reaching global environmental health problem associated with harmful effects and risks posed by the ever-increasing number of contaminants in our environment demand a paradigm shift (proposed by National research council (NRC) and USEPA) in our toxicological assessment approach (Krewski et al., 2010; NRC, 2007; Cwiertny et al., 2014; Ankley et al., 2006; Aardema et al., 2002). There are urgent needs for development of toxicity mechanism and pathways-based cost-effective, reliable and high-throughput in vitro screening and tired testing as complementary and prescreening to the conventional resource-intensive and lengthy in vivo whole animal-based testing, then in time we hope to eventually build predictive models and tools to provide toxicological information that transcends the limits of data generation. Long-standing problems, such as the backlog of untested or insufficiently tested chemicals, could be addressed while reducing the time-, resource- and animal-intensive nature of the current system that sometimes leaves decision-makers without clear guidance concerning the potential risks they must address.

The long-term vision makes the development of predictive toxicity-pathway-based assays the central component of a broad toxicity-testing strategy for assessing biologic activity of new or existing compounds. The major components of new vision include: chemical characterization, toxicity testing, and dose-response and extrapolation modeling (Krewski et al., 2010; NRC, 2007).

Toxicity testing in the Tox21 vision is reconfigured through the development of in vitro medium- and high-throughput assays, to identify the perturbations in toxicity pathway that are expected to lead to adverse effects for evaluation of the host
susceptibility and understanding of the effects on human populations. The revolution, poised by the advantage of system biology, bioinformatics and biotechnology, is making it increasingly possible to study the effects of chemicals using cells, cellular components, and tissues-preferably of human origin-rather than whole animals. These powerful new approaches help to address a number of challenges facing the field of toxicity testing. They enable rapid screening of chemicals, which could reduce the backlog of the large number of industrial chemicals that have not yet been evaluated under current testing system. They should also reduce animal use and suffering (Krewski et al., 2010; NRC, 2007).

These *in vitro* bioassays are more sensitive towards early warning signs, since cellular stress responses are activated before actual harm occurs, and they indicate the presence of associated stressors. Implementation of human and other mammalian cell lines in environmental toxicity testing has facilitated evaluation of toxicological endpoints relevant for assessing the potential for deleterious effects on human health. Since no single bioassay is enough to describe the complex potential effects of the pollution phenomena, the use of battery of bioassays involving different toxicological endpoints as well as different species from different trophic levels is an efficient and essential tool for predicting environmental hazards to the aquatic ecosystem. (Zimmer & Ahlf, 1993; SETAC-Europe, 1993). Recently, single or battery of bioassays including many tests have been validated by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) in the US, European Centre for the Validation of Alternative Methods (ECVAM), etc. Intensely applications with bioassays have been attempted for environmental toxicity monitoring and risk assessment (Escher
et al. 2013; Macova et al., 2011).

However, there are still limitations associated with these single or battery in vitro assays that limit their wide applications. First, most of these in vitro assays only provide information for specific mode of action toxicity endpoints without more comprehensive and detailed information of the toxicity profiles and mechanisms. To perform a comprehensive study, battery of a large number of bioassays still require a substantial amount of time and efforts. Secondly, most in vitro bioassay, especially biomarker-based assays, yield only qualitative (i.e. positive or negative) rather than quantitative endpoints, which make it difficult to be incorporated into regulation, or other decision making processes. Thirdly, these isolated assays that cover only one or a few biomarkers or specific toxicity mechanisms effects cannot be directly anchored or translated to an integrated adverse outcome or phenotypic endpoints that can be further evaluated against established regulations or standards. Finally, there is a great lack of consented procedure to facilitate the standardization of assay protocols, data generation and analysis approach, and the resulting variability can be larger than that of chemical analyses.

1.1.8 Toxicogenomics is a Promising Tool for Toxicity and Hazards Assessment

Large-scale evaluations of the status of gene expression and protein concentrations in cells allow understanding of the integrated biologic activities in tissues and can be used to catalog changes after in vivo or in vitro treatment with environmental agents. So the advanced genomics technologies, which allow for high-throughput concurrently monitoring of the status of cellular response pathways globally upon exposure to chemical toxicants, can be used to expand the coverage of the universe of new and existing chemicals that need to be evaluated for human health risk assessment (Waters et
When applied to the study of large classes of chemicals or environmental samples, use of high-throughput toxicogenomics methods will allow economical screening in a short period and globally define modes or mechanisms of toxic action (Simmons and Portier, 2002). High-throughput toxicogenomics (functional genomics) assays working with pathway mapping tools are useful for predicting important characteristics related to the absorption, distribution, metabolism, excretion, toxicity and assist the human exposome study (Gombar et al. 2003; Masimirembwa et al., 2001; Wild, 2005). Several ongoing public toxicogenomics projects including Tox21 (a program collaborated by NIH, EPA, FDA and aimed at developing better toxicity assessment methods) (NIH, 2014), ToxCast (a environmental chemicals toxicity screening program by EPA) (Dix et al., 2007), Chemical Effects in Biological Systems (knowledgebase of toxicology studies by NIEHS) (Collins et al., 2007), Open TG-GATEs (a Japanese Toxicogenomics Project-Genomics Assisted Toxicity Evaluation) (Mattes et al., 2008), etc. (Abdo et al., 2015; NRC 2006).

Toxicogenomic approaches are distinguished by their ability to reveal patterns of change involving many individual molecules (Ankley et al., 2006). The resolving power of such patterns, when they can be recognized, will likely be much greater than that provided by individual molecules. However, the benefits of toxicogenomics information and tools in regulatory ecotoxicology are only starting to be elucidated (NRC, 2006; NRC, 2007). Existing challenges limited their application for environmental regulatory and risk assessment.

First, cDNA (complementary DNA) microarrays is the current central technology for gene profiling for transcriptomics studies and it allows examination of hundreds to
thousands genes simultaneously (Nuwaysir et al., 1999). However, several drawbacks associated with the microarray technology limit its application in environmental monitoring, such as advanced expertise, complex protocol, costly, lack of standardized protocol, platforms and data analysis making troubles in comparing and transferring results among different sources, and lacks of temporal resolution in that it can only produce limited snapshot profiles at arbitrarily selected time points and therefore may lead to irreproducible and biased results not reflective of the full picture of the genetic response to toxins.

Secondary, the biological system is a rather complex system, and most cell-based assays can only target limit cellular targets. Comprehensive risk assessment thus requires a selection of appropriate cellular targets to cover a range of MOAs and/or recipients relevant for the environmental samples/chemicals to be tested (Ankley et al., 2006; Simmons et al., 2009). However, toxicogenomics studies with an integrative approach that rely on the monitoring of indicative key gene (biomarker) modules in limited and rather conserved pathways (for all cell types of metazoans) linked to adverse outcome are still rare.

Thirdly, the key scientific challenge is to identify reliable patterns (signatures) that report specific exposures and their intensities. There will be significant statistical challenges in establishing criteria for recognizing transcriptomics, proteomics, and metabolomics signatures of exposure (NRC, 2006). The lack of quantifiers, that also considers temporal molecular response pattern and accumulative effects, for convert toxicogenomics information into quantitative toxicity endpoints limits its application for regulatory decision-making (Ankley et al., 2006; NRC, 2006). Currently, there are a few
established approaches for quantifying molecular perturbation NOTEL, benchmark dose (BMD) and $EC_{IR1.5}$ (Barnes et al., 1995; Lobenhofer et al., 2004; Macova et al., 2011). All these molecular quantifiers are derived from either the response intensity, or on the range of response, bias results could be derived. Approaches that can more comprehensively capture the information-rich toxicogenomics data and transform it into integrated and quantitative endpoints are in great need.

Furthermore, there is a great difficulty in extrapolating data from the simple biologic system of single cells to the complex interactions in whole animals (Krewski et al., 2010; NRC, 2007). In addition, to further incorporate these *in vitro* bioassays into current environmental regulations and decision-making framework, comparison and correlation with conventional whole-animal tests are unavoidable. Although there are many researchers who investigated the relationship between toxicogenomics information and organism level phenotypic endpoints, such as organisms weight, cytotoxicity assay and survival rate etc., most of them only performed qualitative analysis rather than quantitative analysis (Paules, 2003; Daston, 2008).

Finally, mixture toxicity evaluation is essential since current threshold regulatory doses/concentrations based on points of departure may not be sufficient against combined exposures. Unfortunately, many of the available mixture studies provide only anecdotal evidence and lack an explicit mechanistic understanding (Altenburger et al., 2012). Current evidence to support the mixture assessments from published toxicogenomics studies is yet mainly observational and needs further improvement. Few situations are considered, no explicit hypothesis on expected combined effects, and there is no quantitative endpoint for mixture toxicity prediction (Altenburger et al., 2012; Finne et
al., 2007; Spurgeon et al., 2010). Mixture toxicity prediction at molecular level have hardly been explored since there is no established framework or theory, it remains unclear if molecular mixture toxicity effect is quantifiable and predictable.

1.2 Research Objectives and Overview

To contribute to the efforts in filling in the knowledge gap and addressing some key aspects of the discussed challenges in toxicity and hazards assessment, as well as in environmental monitoring, this study aimed to develop a novel, fast, cost-effective yet powerful quantitative toxicogenomics-based toxicity assessment platform for high-throughput and effective chemical hazardous identification and environmental toxicity monitoring. Figure 1.1 summarizes the study overview that depicts a comprehensive and systematic design and approaches (Figure 1.2) for method development, optimization, validation, and application demonstration components. First, we developed a high-throughput, low-cost quantitative toxicogenomics assay using GFP (Green Fluorescent Protein)-fused *E coli* whole cell array that covers all known conserved cellular stress response pathway biomarkers. The sensitivity, reproducibility and robustness of the assay were then evaluated and optimized using combined experimental methods and bioinformatics tools. Most importantly, for the first time, a quantitative approach and computation methods were developed that allowed quantification and characterization of dose-response relationship on molecular level toxicity response level at both individual gene or pathway levels. The breakthrough in the molecular effects quantification further enabled us to propose and establish a new molecular mixture toxicity prediction model that had not yet been explored in literature. To validate the newly established assay platform, a number of performance evaluation and comparison were performed and they
include: 1) testing and comparing the assay outcome using well-studied reference model chemicals that have extensive toxicological information available; 2) correlating the molecular endpoints derived from our assays with those originated from established apical toxicity assay; and 3) linking the molecular endpoints derived from our assays to phenotypic and regulatory-relevant endpoints derived from a number of standard or accepted assays in multiple species, so called “phenotypic anchoring”. At last, we conducted a number of cases studies to demonstrate the potential wide application of this developed toxicity assay for 1) toxicity assessment, screening and classification of chemicals of emerging concern; 2) whole effluent water quality monitoring; and 3) risk-based treatment technology efficacy assessment for overall risk reduction rather than “biased” targeted chemical elimination alone. More detailed description of the proposed study aims and approach are as following.
Goal: Developing of a quantitative toxicogenomics platform for toxicity identification and environmental monitoring

**Method Development**
- Assay development
  - GFP-infused *E.coli* whole cell array
  - Stress response ensemble as potential AOP
- Quantitative approach
  - Quantifier development
  - Dose-response
  - Endpoints derivation
- Mixture toxicity prediction at molecular level
  - Mixture toxicity assessment
  - Mixture toxicity prediction

**Method Validation**
- Mechanistic validation
  - Reference compounds with variety of MOAs
- Phenotypic anchoring
- Cytotoxicity within same species
- Cross-species extrapolation
- Bacterial
- Lower eukaryotic
- Higher eukaryotic (cell)
- Correlation with standard water toxicity endpoints
  - ISO, WET, OECD

**Method Application**
- Mechanistic toxicity evaluation of CECs
- 30 emerging contaminants with different function
- Water toxicity monitoring
- Effluents from WWTP
- Technology effectiveness assessment
- Electron-fenton remediation
- Disinfection technologies

Figure 1.1 Overview of the Objectives and Components of this study
Figure 1.2 Overview of the approaches and methodologies of this study.

Toxicants
Specific Gene Promoter
GFP-infused e.coli whole cell array
Selected biomarkers
384 microplate
Microplate reader
Applied real-time gene expression for toxicity assessment.

Generated 3-D chemical-specific real-time gene expression profiles

Developed novel quantitative omics- approach and applied variety bioinformatics tools

MOAs
Toxicity mechanism identification for reference compounds and CECs
Correlate with standard water toxicity endpoints and anchor to phenotypic endpoints
Molecular mixture toxicity assessment and prediction
Technology effectiveness assessment

Bioinformatics analysis
- GSEA, GO, Clustering, PCA...

Technology effectiveness assessment

WWTP effluent toxicity assessment
The specific objectives, approaches, and outcomes are:

1. Develop a high-throughput, low-cost and feasible toxicogenomics-based toxicity assessment technology platform
   a. Assay development:
      Current available and widely used toxicogenomics technologies such as microarrays, proteomics, metabolomics are still quite resource-intensive, complex, costly, lack of standardized results interpretation approach and temporal resolution. To address this, we proposed to develop a toxicogenomics assay using GFP-fused recombinant *E. coli* strains that enables real-time gene expression monitoring with easy and simple protocols, reduced complexity, increased reliability. The cellular response network is complex, examine all potential toxicological target and pathway is not practical. An integrative approach that relies on the monitoring of indicative key gene (biomarker) modules in limited and rather conserved pathways (for all cell types of metazoans) linked to adverse outcome seems promising. We selected indicative key stress response biomarkers that are highly conserved among various species, and cover all those known to be involved in detecting stress and repairing damage caused by toxicants on primary cellular infrastructure. The uniqueness and advantages of our approach are: (1) higher sensitivity and specificity; (2) more informative toxicity-assessment results; (3) improved feasibility and cost effectiveness reflected in the easier, faster and reliable assay procedures, reusability and flexible customization of the cell-array
library. These advantages and features make prokaryotic real-time gene-expression profiling a promising method for toxicity assessment and for practical application in water monitoring.

b. Molecular effects quantification

One of the most challenging aspects of implementing toxicogenomics-based toxicity assessment in regulation and in risk assessment involves establishing the theoretical framework and comprehensive knowledge base required to obtain quantitative toxicogenomics assay endpoints and correlate them with phenotypic biological adverse outcome at regulatory relevant level. This study, for the first time, proposed a new the concept and computational method for converting the information-rich toxicogenomics data into integrated and quantitative endpoints-TELI (transcriptional effect level index) with prokaryotic *E coli* reporter cells. We developed mathematical logarithm and computational methods for data analysis. This study also explored and modeled the dose-response relationship using the developed molecular level perturbation quantifier -TELI. Multiple molecular response endpoints were derived from the dose-response relationship to completely and holistically describe the dose response behavior and characteristics. This study pioneered and demonstrated the validity and potential application of this concept for quantifying molecular perturbation cells and linking it with phenotypic endpoints.
c. Molecular mixture toxicity prediction

Mixture toxicity prediction at molecular level has hardly been explored since there is no established framework or theory, and it still remains unclear whether molecular mixture toxicity is quantifiable and predictable. The quantification of molecular response enables us to explore the mixture effect at molecular level. We applied the developed quantitative toxicogenomics approach for investigating the combined effects of various binary mixtures with a range from environmental relevant concentration to benchmark level. Quantitative mixture prediction on molecular level was proposed and developed based on the conventional phenotypic mixture toxicity prediction model framework. The model outputs were compared with experimental data for various binary mixtures. Our results demonstrated that the proposed molecular mixture model can explicitly and quantitatively predict the combined effects at the molecular level, and the pathway-based quantitative toxicogenomics approach is applicable for mixture toxicity studies. The developed toxicogenomics approach has been illustrated as a rapid, sensitive, and informative approach for mixture toxicity studies.

2. Validation of the developed quantitative toxicogenomics-based technology

a. Mechanistic validation with reference model chemicals

The toxicogenomics results revealed more detailed transcriptional information on the toxic mechanism and led to a better understanding of the mode of action (MOA). This study validated the developed
approach via comparing of mechanistic toxicity profiles with known MOAs for reference compounds. The consistent MOA obtained via toxicogenomics approach for all the tested reference compounds, in comparison with their previously known mechanisms, demonstrated that real-time gene expression profiling, which yields compound-specific and concentration sensitive molecular toxicity response “fingerprints” specific to each compound, therefore can be applied as a feasible method for toxicity assessment and screening of a large number of emerging contaminants.

b. Correlating molecular response endpoints from our toxicogenomics assay to apical endpoints from conventional toxicity test

Great challenges and knowledge gap exists in the emerging field of toxicogenomics with regard to defining the relationships between chemically induced changes in gene expression and alterations in conventional toxicological parameters such as conventional toxicity endpoints, clinical chemistry and histopathology. This study quantitatively evaluated the correlation of the gene expression result expressed by TELI from our cellular stress response ensemble toxicogenomics approach with conventional apical endpoints, EC50, based on growth inhibition on *E. coli*. The great correlations imply that “phenotype anchoring” is possible with our TELI derived endpoints TELImax and EC-TELI1.5, therefore allowing for the linking of phenotype changes with cellular response. These “phenotypic
anchoring” to cytotoxicity indicate that the stress response perturbation we measured in the toxicogenomics assay was able to capture and quantify the major cellular toxicity impacts that lead and contribute to cytotoxicity.

c. Link molecular endpoints to regulatory relevant phenotypic endpoints on multiple species

To further incorporate the developed toxicogenomics-based approach into current environmental regulations and decision-making framework, comparison and linking with conventional whole-animal tests are unavoidable. This study explored the linking between molecular endpoints to regulatory relevant phenotypic endpoints on multiple regulatory relevant species, including as V.fisheri, single cell green algal S.capricornutum and fish cell line RTgillW-1. These species are all widely used model organisms for water quality monitoring standard methods by ISO and USEAP. Our results showed that our toxicogenomics data can be quantitatively link to conventional toxicity endpoints from standard methods for water toxicity assessment, thus indicates the possibility of utilizing our quantitatively toxicogenomics approach with cellular stress ensemble for water toxicity assessment. This exploration provided further validation of molecular toxicity assay, and helped to distinguish more defined and focused molecular toxicity assays, entail possible interspecies
extrapolation and to gain greater confidence in animal models for the possible realization of the vision of mechanistic predictive toxicity.

3. Demonstration of the wide application of the quantitative toxicogenomics technology for toxicity screening, environmental monitoring and technology efficacy assessment

   a. Mechanistic toxicity assessment of CECs

Currently, most of the CECs have not been adequately tested for their potential impacts to human and wildlife due to the labor and resources intensiveness of conventional animal-based toxicity assessment as well as the complexity and challenges in toxicity characterization of CECs. Chemical toxicity screening and prioritization is in urgent needed. This study utilized the developed quantitative toxicogenomics-based approach for mechanistic toxicity evaluation and classification of 30 CECs, covering various group, such as pharmaceuticals, antibiotics, drinking water disinfection byproducts, food, industrial additive, pesticide, herbicide, nanomaterials etc. The results provided an initial database of comprehensive molecular toxicity information for a large number and diverse categories of CECs, which can assist future QSAR and predictive toxicity model development. The toxicity results not only confirmed known toxicity mechanism for some CECs, but also revealed new toxicological insights and information for many other CECs. The proposed method may serve as an alternative or complementary approach to the current toxicity assays for
environmental pollutants toxicity screening, evaluations and classification.

b. Water quality monitoring

Current available water toxicity evaluation method, such as WET, which provides very limited information on endpoint lethal or sub-lethal effects of toxicants. It neither identifies the toxicants nor reveals the mode of action underlying the toxicity effects. This study applied the developed quantitative toxicogenomics-based approach for monitoring the toxicity level and profiles changes during the wastewater treatment process to demonstrate the effectiveness of the assay for more sensitive detection of toxicity, revealing the toxicity nature and mechanism, and discern the difference among different WWTPs due to the high-resolution and information rich output serving as fingerprints. This work, for the first time, utilized a quantitative toxicogenomics approach as an integrated toxicity reduction evaluation (TRE) monitoring approach that identifies the specific toxic substances and toxicity, the pass-through variability of the overall toxicity in the effluent wastewaters, and finally insights into the probable solution for controlling the toxicity discharge.

c. Risk-based technology efficacy assessment for risk reduction and minimization

Current treatment effectiveness assessments and water quality monitoring that rely on measuring the reduction of target contaminants
alone are insufficient because they do not consider the complex and broader risks that specific or mixtures of contaminants and transformation products pose to the environment and human health. To demonstrate the advantages of using toxicogenomics based approach for risk-based technology efficacy evaluation, we performed two case studies. The first is for evaluation of the toxicity evolution and nature along the electro-Fenton oxidative degradation of three representative CECs whose oxidative degradation pathways have been relatively well studied. And the second case study was to evaluate the impact of different disinfection processes on wastewater effluent toxicity. The demonstration and verification of the proposed cost-effective and quantitative molecular toxicity assay will contribute to the paradigm shift in the assessment of remediation efficacy from those that suffers from “biased and limited” chemical information to a more reliable risk-based one that identifies realistic endpoints better reflecting the actual risks to receptors. This shift in water toxicity assessment approach could result in the changes in the remediation strategy development, remediation technology design, optimization, monitoring and, consequently impact risk management and environmental and human health protection.
1.3 Organization of the thesis

The dissertation consists of 11 chapters and they are briefly described as following:

**Chapter 1 Introduction and Objective** This chapter provided the background and motivations, and describes the overall objectives, approaches as well as outcomes of this study.

**Chapter 2 Advances in Quantitative Toxicogenomics for Environmental Monitoring**
This chapter reviewed the recent advances in various aspects related to the toxicogenomics field with particular focuses on the environmental applications and water quality monitoring.

**Chapter 3 A New Transcriptional Effect Level Index (TELI) for Toxicogenomics-based Toxicity Assessment**
This chapter proposed and demonstrated the potential application of a new Transcriptional Effect Level Index (TELI) to convert the information-rich temporal toxicogenomics data into integrated and quantitative endpoints. The results demonstrated that the methodology for converting the rich toxicogenomic information into a readily usable and transferable format that can be potentially linked to regulation endpoints and incorporated into a decision-making framework.

**Chapter 4 Mixture Toxicity Prediction via Quantitative Toxicogenomics-based Approach**
proposed and developed a Quantitative mixture prediction model on molecular level based on the conventional phenotypic mixture toxicity prediction model framework. Our results demonstrated that the proposed molecular mixture model can explicitly and quantitatively predict the combined effects at the
molecular level, and the pathway-based quantitative toxicogenomics approach is applicable for mixture toxicity studies.

**Chapter 5 Mechanistic Toxicity Assessment and Classification of Environmental Contaminants of Emerging Concern Via Stress Response Ensemble Based Toxicogenomics Assay**

This chapter applied our newly developed quantitative toxicogenomics method based on cellular stress response pathways biomarkers ensemble, for a fast, cost-effective, yet informative and mechanistic toxicity evaluation and classification of 30 CECs. The resulted demonstrated that the proposed technology may serve as an alternative or complementary approach to the current toxicity assays for environmental pollutants toxicity screening, evaluations and classification.

**Chapter 6 Link Quantitative Toxicogenomics Endpoints to Conventional Toxicity Endpoints - Phenotypic Anchoring and Cross Species Extrapolation**

This chapter explored and demonstrated the quantitative correlation between endpoint-TELI (Transcriptomics Effect Level Index) derived from a stress response pathway ensemble-based toxicogenomics assays in E coli with conventional apical endpoints in three regulatory relevant species (V.fisher, single cell green algal S.capricornutum and fish cell line RTgillW-1), for over 30 environmental chemicals and their mixtures. The toxicogenomics approach not only yielded several endpoints and but also much more detailed toxicity mechanisms and profiling fingerprints beyond single endpoints (i.e. EC50), therefore, in combination with statistical and bioinformatics tools, can be potentially applied for environmental monitoring.
Chapter 7 Mechanistic Toxicity Assessment of Nanomaterials by Whole-cell-array

**Stress Genes Expression Analysis** This chapter performed mechanistic toxicity assessment of nano-silver (nAg) and nano-titanium dioxide anatase (nTiO$_2$-a) via toxicogenomic approach, employing a whole-cell-array library consisting of 91 recombinated *Escherichia coli* (*E.coli*) K12 strains with transcriptional GFP-fusions covering most known stress response genes. The results, for the first time, revealed more detailed transcriptional information on the toxic mechanism of nAg and nTiO$_2$-a, and led to a better understanding of the mode of action (MOA) of metal and metal oxide nanomaterials (NMs).

Chapter 8 Application of Quantitative Toxicogenomics Assay for Wastewater Water Quality Monitoring and Treatment Efficacy Assessment In this chapter, we, for the first time, demonstrated the application of a quantitative toxicogenomics-based approach for evaluation of the toxicity evolution and nature along each treatment unit process in two wastewater treatment plants. Toxicity level and profile evolution along each treatment processes were monitored to reveal the dynamic toxicity changes and mechanisms, as well as their association with treatment effectiveness assessment.

Chapter 9 Quantitative Toxicogenomics Assay Reveals the Evolution and Nature of Toxicity during Environmental Pollutants Transformation This chapter demonstrated the application of a novel, fast and cost-effective quantitative toxicogenomics-based approach for evaluation of the evolution and nature along the electro-Fenton oxidative degradation of three representative CECs whose oxidative degradation pathways have been relatively well studied, and they were
bisphenol A, triclosan and ibuprofen. The results demonstrated that the quantitative toxicogenomics assay might serve as a useful tool for remediation technology efficacy assessment, and provides guidance on process design and optimization for desired toxicity elimination and risk reduction.

**Chapter 10 Quantitative Toxicogenomics Assay Revealed the Impact of Disinfection Technologies On Effluent Toxicity**

In this chapter, we applied the developed quantitative toxicogenomics-based approach to elucidate the toxicity level and profiles in the municipal WWTP secondary effluent that were disinfected by different technologies, including chlorination (Cl₂), chloramination (NH₂Cl), ozonation, ultraviolet irradiation (UV) as well as combination of ultraviolet irradiation and hydrogen peroxide. Iceberg designs were also employed to assess if the detected chemicals drive the biological effect and which fraction of effect remains unexplained by detected chemicals. This study demonstrated that toxicogenomics-based assay could be applied for detailed and informative toxicity evaluations of water samples and for treatment technology effectiveness assessment.

**Chapter 11 Conclusions and Future Work** summarizes the conclusions, novelty, and major contributions of this study and provides some recommendations for future studies.
Chapter 2

Advances in Quantitative Toxicogenomics for Environmental Applications on Water Quality Monitoring

2.1 Abstract

The emerging field of toxicogenomics promises a revolutionary new ground for evaluating toxic effects, understanding toxicity mechanisms, and obtaining pollutant(s)-specific molecular fingerprints (or biomarkers) for compound classification and identification. This review summarizes the recent advances in various aspects related to the toxicogenomics field with particular focuses on the data quantification for environmental applications and water quality monitoring. The definitions and mainstream toxicogenomics technologies are first described. Then, the efforts and progress made in addressing the most challenging aspects of implementing toxicogenomics-based toxicity assessment in regulation and in risk assessment are encapsulated. They include fundamental knowledge, methodologies and bioinformatics tools required to obtain quantitative toxicogenomics assay endpoints; current status in developing adverse outcome pathway (AOP) concept databases for linking molecular endpoints to phenotypic adverse outcome at various biological organizational levels relevant to disease and regulation; and promises and successful cases for employment of highly-
conservative stress response pathways-related biomarkers for toxicogenomics assays that quantitatively correlate molecular perturbation to biological effects. Finally, explorations, issues and challenges in applying toxicogenomics approach for mixture toxicity evaluation and for environmental applications (i.e. for water toxicity assessment) are discussed and, cases studies are provided.

2.2 Introduction

There is an overwhelmingly large number of unregulated yet widely used chemicals that are found ubiquitously in the environment and, most of them has little or no toxicity data exists (United States Government Accountability Office, 2005; Krewski et al., 2010). Current water monitoring has been focusing only on limit chemicals which have been traditionally considered toxic with regulatory benchmarks exist. Regarding of the toxicity assessment, the current available water toxicity assessment method, such as WET (Whole Effluent Toxicity) (USEPA, 2000) and TIE (Toxicity Identification Evaluation) (USEPA, 1991) adopted those assay only provided limit apical endpoints with limit information on the effects of toxicants. The underlying reasons for these are mainly raised from the limitations of current toxicity testing approach. Current approaches to toxicity testing rely primarily on observing adverse biologic responses in homogeneous groups of animals exposed to high doses of a test agent. However, the relevance of whole-animal studies for the assessment of risks to the populations exposed at much lower concentrations has been questioned. Moreover, the studies are expensive and time-consuming and can use large numbers of animals, so only a small proportion of chemicals have been
evaluated with these methods (NRC, 2006). Adequate coverage of different life stages, of endpoints of public concern, such as developmental neurotoxicity, and of mixtures of environmental agents is also a continuing concern (Krewski et al., 2010). Current tests also provide little information on modes and mechanisms of action, which are critical for understanding interspecies differences in toxicity, also limited the ability to predict toxicity about chemicals that have not been tested (NRC, 2007). In addition, chemicals are ubiquitous in the environment and often occur in mixtures and these mixtures can interact additively, synergistically.

**Box 2.1 Limitation of current regulatory toxicity testing**

Current toxicity assessment method only focus on limit compound with limit threshold regulatory dose/concentrations based on points of departure

Facing with large numbers of existing chemicals, many of which lack basic toxicity data

Rely on in vivo animal test, time-consuming and resource intensive, and sometimes raises ethical issues

Lack of information for much lower doses than those used in whole-animal studies

Limited mechanistic information limit the ability to predict toxicity about similar chemicals that have not been tested

Controversial uncertainty factors must be applied to account for differences between different species

Threshold regulatory doses/concentrations based on points of departure may not be sufficient against combined exposures.

The growing evidence and severely rising public concern of the far-reaching global environmental health problem associated with harmful effects and risks
posed by these chemicals demand a paradigm shift in our toxicological assessment approach (Niemeier et al., 2010; Aardema et al., 2002; Chang et al., 2002; Newton et al., 2004). There are urgent needs for development of toxicity mechanism and pathways-based cost-effective testing scheme as alternatives to the conventional resource-intensive and lengthy whole animal-based testing, then in time we hope to eventually build predictive models and tools to provide toxicological information that transcends the limits of data generation (Chang et al., 2002; Applegate et al., 1998).

The advance in “-omics” technologies, which allows for high-throughput concurrently monitoring of the status of cellular response pathways globally upon exposure to chemical toxicants, can be used to expand the coverage of the universe of new and existing chemicals that need to be evaluated for human health risk assessment (Krewski et al., 2010; Waring et al., 2001). This review summarizes the recent advances in various aspects related to the toxicogenomics field with particular focuses on the data quantification and result interpretation for environmental applications and water quality monitoring.

### 2.2.1. The Emerging Field of Toxicogenomics

Toxicogenomics employs various omics (transcriptomics, proteomics and metabolomics) technologies to assess the variety of effects that specific chemicals (environmental stressors and toxicants) can cause in biological systems. It investigates mRNA expression, cell and tissue-wide protein expression and metabonomics to understand the role of gene-environment interactions in toxicology. The advances in high-throughput technologies and the capacity to
examine large numbers of individual gene/protein/metabolites fragments on small matrices provide an opportunity to identify the early, sensitive genes or proteins responsive to cellular toxicity resulting from toxicant exposure. The high-resolution and information-rich exposure fingerprints can then be used, in combination with bioinformatics tools for toxicants and exposure effects classification, as well as for predicting mode of action. Compared to conventional whole-animal assays, the toxicogenomics approach not only save time, cost and animal use but also enable us to detect early, sub-cytotoxic molecular level effects, gain insights into underlying toxic mechanisms, and allow us to perform studies across multiple species, including human species that could not be carried out at overly toxic exposure (NRC, 2006; Ankley et al., 2006; Ankley et al., 2010).

2.2.1.1 Common Toxicogenomics Technologies

Toxicogenomics is used as a general term referring to analytical technologies that integrate the function of genome and they may include transcriptomics, proteomics and metabolomics or metabolomics that measure expressed genes, proteins and metabolites, respectively (Waters and Fostel, 2004). cDNA (complementary DNA) microarrays is the most commonly used technology for transcriptomics studies and it allows examination of hundreds to thousands genes simultaneously (Nuwaysir et al., 1999). Scientists use DNA microarrays to measure the expression levels of large numbers of genes simultaneously or to genotype multiple regions of a genome (Hamadeh et al., 2002). The advent of inexpensive microarray experiments created numerous chances/applications as well as specific challenges in bioinformatics, such as
standardization (MAQC Consortium, 2006), statistical analysis (Leung, 2003),
data warehousing, etc.

Advances in massively parallel sequencing has led to the development of RNA-Seq technology, that enables a whole transcriptome shotgun approach to characterize and quantify gene expression (Mortazavi et al., 2001; Wang et al., 2009). Unlike microarrays, which need a reference genome and transcriptome to be available before the microarray itself can be designed, RNA-Seq can be also be used for new model organisms whose genome has not been sequenced yet (Wang et al., 2009).

The most common approach for proteomics is gel-based proteomics. In the gel-based approach, proteins are resolved by electrophoresis or another separation method and protein features of interest are selected for analysis. This approach is best represented by the use of two-dimensional sodium dodecylsulfate polyacrylamide gel electrophoresis (2D-SDS-PAGE) to separate protein mixtures, followed by selection of spots, and identification of the proteins by digestion to peptides, MS analysis, and database searching. Comparative 2D-SDS-PAGE with differential fluorescent labeling (for example, differential gel electrophoresis, DIGE) offers powerful quantitative comparisons of proteomes (Tonge et al., 2001; Eggeling et al., 2001). Although 2D gels have been applied most commonly to global analyses of complex proteomes, they have great potential for comparative analyses of smaller subproteomes (for example, multiprotein complexes).

Problems with gel-based analyses including time consuming, labor intensive, and has low reproducibility due to the expertise required and large variations in the
quality of gel run. Another disadvantage is that it can mask the expression of less abundant proteins, resulting in inaccurate quantification and potential misidentification (Gorg et al., 2004; Kussmann et al., 2006). Because of these disadvantages, the simultaneous use of liquid chromatography (LC) as a separation tool and mass spectrometry (MS or MS/MS) as an identification tool results in much higher efficiency (i.e., higher number of proteins identified per unit of time), have become increasingly popular (Monsinjon and Knigge, 2007; Lemos et al., 2010). This approach can generate near-comprehensive sequence analysis of individual protein molecular forms, thus enabling sequence-specific annotation of individual modification variants (NRC, 2007). A limitation of the approach is the requirement for relatively purified proteins and larger amounts of samples than are used in shotgun analyses. However, rapid technology development will make top-down methods increasingly useful for targeted analyses of individual proteins and their modified forms.

Non-MS-based technologies, such as antibody microarrays (in which immobilized antibodies recognize proteins in complex mixtures (NRC, 2007; Wildt et al., 2000), have been applied to proteome analyses, but they have not proven to be as robust and versatile as MS-based methods. The principal technology platforms for metabolomics are NMR spectroscopy and gas chromatography MS (GC-MS) or LC-MS (NRC, 2007). Quantitative analytical methods have been developed to identify metabolites in pathways or classes of compounds. This collective directed approach has been called metabolite profiling or metabolomics. Semi-quantitative,
nuclear-magnetic resonance (NMR) based metabolic fingerprinting has also been applied to high-abundance metabolites and has been termed ‘metabonomics’ (Nicholson et al., 2002). Peaks detected in NMR spectra carry information regarding the structure of the metabolites, whereas peaks detected by mass spectrometry have associated molecular weights. In addition, specific mass spectrometry methods can be established to fragment the parent molecule, allowing metabolites to be identified through investigation of fragmentation patterns (Waters and Fostel, 2004).

As an alternative to the above technologies, whole-cell recombinant bio-reporters consisting of prokaryotic (bacteria) or eukaryotic (yeast, fungi, algae, animal) cells that serve as living sensors have been used to obtain real time measurements of gene/protein expression levels in exposure to environmental stressors (Bulich and Isenberg, 1981; Jennings et al., 2001; Onnis-Hayden et al., 2009; Gou et al., 2010; Gou and Gu, 2010). Bacterial whole-cell biosensors produce measurable gene products encoded by reporter genes, which are either present naturally in the bacterial strain or introduced by genetic manipulation. The most frequently used bacterial reporter genes include the lacZ gene from *Escherichia coli*, the lux genes from, for example, *Vibrio fischeri* or the gfp gene from *Aequorea Victoria* (Bulich and Isenberg, 1981; Jennings et al., 2001; Burmolle et al., 2006; Hansen and Sorensen, 2001; Vollmer and Van Dyk, 2004). The reporter genes are usually fused downstream to specific gene promoters, carried by a plasmid or directly on chromosome, resulting in reporter genes expression when the biosensor strain is exposed to conditions triggering a response that involves the selected genes
Proteomics studies have also been demonstrated using whole-cell reporters with *gfp*-tagged Yeast library that was constructed by oligonucleotide-directed homologous recombination to tag each ORF with *Aequorea victoria* GFP in its chromosomal location at 3’ end, and it allows for continuous and real time measurements of the protein expression (proteomics) with high reliability (Huh et al., 2003; Vogel et al., 2012; O’Connor et al., 2012). Whole cell bioreporters has the advantages of simpler protocol, ease of use and low cost (Onnis-Hayden et al., 2009; Gou et al., 2010; Gou and Gu, 2010. More importantly, the ability of whole cell bioreporters to yield temporal gene/protein expression data can more accurately reflect the chemical-induced cell responses in a temporal pattern (Onnis-Hayden et al., 2009; Gou et al., 2010; Gou and Gu, 2010; Gao et al., 2012).

2.2.1.2 Key Challenges on Applying Toxicogenomics Approach for Regulatory Risk Assessment

Toxicogenomic approaches are distinguished by their ability to reveal patterns of change involving many individual molecules. The resolving power of such patterns, when they can be recognized, will likely be much greater than that provided by individual molecules. However, the benefits of toxicogenomics information and tools in regulatory ecotoxicology are only starting to be elucidated. Existing challenges limited their application for environmental
regulatory and risk assessment (NRC, 2006; NRC, 2007; Waters and Fostel, 2004). First, the biological system is a rather complex system, and most cell-based assays can only target limit cellular targets. Comprehensive risk assessment thus requires a selection of appropriate cellular targets to cover a range of MOAs and/or recipients relevant for the environmental samples/chemicals to be tested. Second, the key scientific challenge is to identify reliable patterns (signatures) that report specific exposures and their intensities. There will be significant statistical challenges in establishing criteria for recognizing transcriptomic, proteomic, and metabolomic signatures of exposure. Third, there will be difficulty of extrapolating data from the simple biologic system of single cells to the complex interactions in whole animals. In addition, to further incorporate these in vitro bioassays into current environmental regulations and decision making frame, comparison and correlation with with conventional whole-animal tests are unavoidable. Last, mixture toxicity evaluation is essential since current threshold regulatory doses/concentrations based on points of departure may not be sufficient against combined exposures.
To address the previous challenges and further incorporate the toxicogenomics into regulatory framework, the keys are the selection of appropriate cellular targets to further link to an adverse outcome as well as the quantitative analysis of high throughput toxicogenomics information (Ankley et al., 2006; Ankley et al., 2010; Waters and Fostel, 2004).

### 2.3 The Concept of Adverse Outcome Pathway (AOP) in Toxicogenomics

#### 2.3.1 AOP Concept

To use mechanistic data derived from toxicogenomics to support chemical assessments, there is a need for effective translation of these molecular response information into endpoints meaningful to ecological risk—effects on survival, development, and reproduction in individual organisms and, by extension,
impacts on populations (Ankley et al., 2010; Boverhof and Zacharewski, 2006; Kramer et al., 2011). A adverse outcome pathway (AOP) is a conceptual construct that provides roadmap for establishing linkage between a direct molecular initiating event (MIE, e.g. a molecular interaction between a xenobiotic and a specific biomolecule) and an adverse outcome at a biological level of organization relevant to risk assessment (NRC, 2006; Ankley et al., 2010; Kramer et al. 2011; “Adverse Outcome Pathway Wiki”, 2013). As stated by Ankley et. al., AOP describe the known linkage along the continuum from the molecular event in which a chemical interacts with a biological target(s); following on through a sequential series of cellular, anatomical and functional changes in biological process, and ultimately culminating in an adverse outcome of relevance to human or ecological risk assessment9. Characterization of AOPs allows for identification of key events based on which high-throughput testing methods can be developed. AOPs also provide important information on the development of structure-activity relationships, i.e. using effects information from one chemical (the source chemical) to predict the effect for another structurally similar chemical (the target chemical) (“Adverse Outcome Pathway Programme”, 2013). Finally, AOPs provide evidence important for qualitative and quantitative predicative models of the adverse outcomes that result from triggering molecular initiating (MIE) or other key events (“Adverse Outcome Pathway Programme”, 2013).

2.3.2 AOP Development Status

To address the needs of identifying new in vitro tests methods, classification of chemicals and developing integrated testing strategies for hazardous endpoints, a
number of ongoing efforts focused on the development, cataloging and application of AOP to support both human health and ecological risk assessments (Krewski et al., 2010; NRC, 2006; Gorg et al., 2004). For example, the Organization for Economic Cooperation and Development (OECD) recently initiated an AOP Development Programme (“Adverse Outcome Pathway Programme”, 2013). A key component of this effort involves contributions from an international group of scientists (several of whom are associated with SETAC) from OECD member countries participating in the OECD Extended Advisory Group on Molecular Screening and Toxicogenomics (EAG MST). Activities of the OECD EAG MST have included issuance of guidance on the development and documentation of AOPs, and the initial assembly and review of a candidate set of AOPs (“Adverse Outcome Pathway Wiki”, 2013; “Adverse Outcome Pathway Programme”, 2013). Currently, there are 18 proposed AOPs (details not available yet) included in the AOP Development Program that includes several AOPs that are conserved across species (e.g. nonpolar Narcosis, cell signaling pathways associated with cell proliferation and differentiation) and others that are more species or pathway-specific (e.g. AOP linking Estrogen Receptor Antagonism, or Steroidogenesis Inhibition, to Impaired Reproduction in Small Repeat-Spawning Fish Species; AOP linking protein alkylation to liver fibrosis) (“Adverse Outcome Pathway Wiki”, 2013).

2.3.3 Limitation of Current AOP for Ecotoxicity Assessment

AOPs provided broad potential to serve as novel tools in toxicology and risk assessment. They are designed to provide a clear-cut mechanistic representation
of critical toxicological effects that span over different layers of biological organization. However, most of the current developed and under developing AOPs are for receptor-mediated toxicity with specific mechanism. For environmental risk assessment, various chemicals with different specific mode of action present in environmental media simultaneous, which leading to biased result if only assessing the specific toxicity mechanism.

To achieve more comprehensive evaluation of environmental risk, batteries of bioassays consisting of a number of separated assays with different AOPs can be applied. However, batteries of bioassays still require a substantial amount of time and efforts. In addition, most of them only provide information for specific mode of action toxicity endpoints without more comprehensive and detailed information of the toxicity profiles and mechanisms. Furthermore, isolated assays that cover only one or a few biomarkers or specific toxicity mechanisms effects cannot be directly anchored translated to an integrated adverse outcome or phenotypic endpoints

2.3.4. Stress Response Pathways Ensemble-Based Toxicogenomics Assays

2.3.4.1 Highly conserved Stress Response Pathways Across Species

Given that it is not practical to design an cost-effective in vitro high throughput assays that can interrogate every potential toxicological target and pathway, an integrative approach that rely on the monitoring of indicative key gene (biomarker) modules in limited and rather conserved pathways (for all cell types of metazoans) linked to adverse outcome seems promising (Onnis-Hayden et al., 2009; Simmons
et al., 2009). Despite the degree of biological organizational level, cells respond to environmental stress (chemical toxicants) through a common architecture with a number of rather conserved specific stress response pathways in an effort to repair the damage to restore homeostasis or make cell fate/death decisions (Kültz, 2005; Simmons et al., 2009). Cellular stress response pathways detect and repair damage caused by toxicants on primary cellular infrastructure, which are the minimal components including the basic macromolecules and the superstructures made of macromolecules (Simmons et al., 2009). One important point is that stress response pathways represent set of homeostatic pathways that are named according to key toxicity events elicited in cells, such as oxidative stress, genotoxic stress, metal stress, ER stress, osmotic stress, heat shock response, hypoxic stress etc. Necrosis, apoptosis, and DNA repair have been studied extensively because there is such a high degree of conservation of the various genes involved in fundamental aspects of these effects (Simmons et al., 2009; Kültz, 2005; Paules, 2003; Taylor and Lehmann, 1998; Hanawalt et al., 1979; Kang et al., 2005; Westerhede and Morimoto, 2005; Lee et al., 2007; Andersen et al., 1987). Therefore, cellular stress response pathways offer a viable solution to the creation of a set of integrative assays as there is a limited and hence manageable set of major cellular stress response pathways to enable the development of high-throughput cell-based assays using the components of the pathways (Onnis-Hayden et al., 2009; Simmons et al., 2009). And, the cellular stress response pathways are thus directly related to the phenotypic changes and
adverse outcomes, which make them potential AOPs (Simmons et al., 2009; Kültz, 2005).

2.3.4.2 Cellular Stress Pathways Ensemble-based Toxicogenomics

Application

Various stress response pathways have been used for ecotoxicity assessment via toxicogenomics approach, as shown in Table 1. Oxidative stress and DNA stress are among the most commonly examined stress responses (Gagne et al., 2013)). Heavy metal stress, protein stress, membrane integrity stress etc. are also been evaluated. Environmental water sample covers almost the whole water cycle, including surface water, drinking water, industry effluent, wastewater treatment plant influent and effluent. Transcriptomics assays appear highly prevalent and interesting tools to underlie the early biological effects of complex mixtures and as biomarkers of exposure to some classes of compounds, although different technologies are applied (Gagne et al., 2013).

Table 2.1 List of Studies Using Stress Response and Pathways-Related Biomarkers for Ecotoxicity Assessment

<table>
<thead>
<tr>
<th>Environmental stressors</th>
<th>Species/Cell line</th>
<th>Technology</th>
<th>Biomarkers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oil sand water and processed water (Gagne et al., 2013b)</td>
<td>Rainbow trout hepatocytes</td>
<td>RT-PCR</td>
<td>16 genes including 13 stress genes covering metal homeostasis, xenobiotic biotransformation, estrogenicity, DNA repair, and oxidative stress</td>
</tr>
<tr>
<td>Municipal wastewater treatment plant influents and effluents (Gagne et al., 2013a)</td>
<td>Rainbow trout hepatocytes</td>
<td>RT-PCR</td>
<td>8 genes covering xenobiotic biotransformation, estrogenicity, heavy metal detoxification, and oxidative stress</td>
</tr>
<tr>
<td>Study</td>
<td>Cell Line/Treatment</td>
<td>Assay/Reporter Gene</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>------------------------------</td>
<td>-----------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Synthetic binary mixtures (Dardenne et al., 2007)</td>
<td><em>E. coli</em> Deletion library</td>
<td>14 stress genes covering oxidative stress, DNA damage/SOS response, Membrane integrity/osmotic stress, Protein perturbation, general stress and heavy metal stress.</td>
<td></td>
</tr>
<tr>
<td>U.S. National Toxicology Program 1,408-compound library (Shukla et al., 2012)</td>
<td>HepG2 cell β-lactamase reporter gene and luciferase reporter gene</td>
<td>Antioxidant response element (ARE)</td>
<td></td>
</tr>
<tr>
<td>Sewage/drinking water (Escher et al., 2012)</td>
<td>Human breast cancer cell line MCF7 Reporter gene assay AREc32</td>
<td>Antioxidant response element (ARE)</td>
<td></td>
</tr>
<tr>
<td>Wastewater, drinking water, surface water (Macova et al., 2011)</td>
<td>Battery test, <em>Vibrio fisheri</em>, algal, <em>E. coli</em>, MCF-7 cell etc. Genotoxicity umuC assay Estrogenicity E-SCREEN assay</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22 antioxidants (Fox et al., 2012)</td>
<td>HEK293T cell Luciferase fusion gene ATAD5 protein</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CdCl₂, NaAsO₂ (Gottschalg et al., 2006)</td>
<td>Rat hepatoma FGC4 cell line, rat hepatocytes, human hepatoma HepG2 cell line ELISA 6 stress proteins including metallothionein (MT) and members of the heat shock protein (HSP) family</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cu (Devi and Prasad, 1998)</td>
<td>Ceratophyllum demersum L. (Coontail) -- antioxidative enzymes and substances</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Pesticides (Pham et al., 2004)</td>
<td><em>E. coli</em> Lux reporter gene</td>
<td>4 Stress genes covering oxidative stress, DNA stress, lipid stress and protein stress</td>
<td></td>
</tr>
<tr>
<td>Benzo(a)pyrene (O’Connor et al., 2012)</td>
<td>Yeast GFP-tagged ORF</td>
<td>123 ORF covering DNA stress, oxidative stress, protein stress, and chemical stress</td>
<td></td>
</tr>
<tr>
<td>Mercury and mitomycin Nanomaterials (Onnis-Hayden et al., 2009; Gou et al., 2010; Gou and Gu, 2010)</td>
<td><em>E. coli</em> GFP infused</td>
<td>91 stress genes covering redox stress, DNA repair/SOS response, detoxification, heavy metal stress, protein stress, membrane stress etc.</td>
<td></td>
</tr>
</tbody>
</table>
2.3.4.3 Anchoring of Stress Response via Toxicogenomics Approach to Adverse Outcome

Phenotypic anchoring is central to the research strategy and concept of AOP since it requires the initial link that relates specific molecular level effects to specific adverse effects of environmental stresses defined by conventional parameters of toxicity (Paules, 2003; Schmidt, 2002; Ung et al., 2010). Paules pointed out that phenotypic anchoring of gene expression data to toxicological and pathological indices removes some of the subjectivity of conventional molecular expression analyses. It also helps to distinguish the toxicological effect signal from other gene expression changes that may be unrelated to toxicity, such as the varied pharmacological or therapeutic effects of a compound (Paules, 2003).

2.3.4.3.1 Specific stress response

Stress response related alternation in gene/protein expression patterns at low, subtoxic level (i.e., no apparent toxicity was detected) may reveal signs of subtle cellular injury that exacerbated at higher doses, or different species. Specifically, in Powell et al, toxicity endpoints and protein adduct formation was used to phenotypically anchor oxidative stress gene expression due to acetaminophen exposure (Powell et al., 2006). This oxidative stress signature was observed that was corroborated in a time-dependent manner with increases in oxidized purines and abasic sites in DNA (Powell et al., 2005). Correlation between cellular response on protein expression with observed DNA lesion by Comet assay were also reported by Lan et al. (Lan et al., 2015). Cellular response were monitored at different species (E.coli, yeast, and human cell line A549)
using different toxicogenomics technologies (transcriptomics, proteomics, rt-PCR) covers various DNA repair pathway, great agreement were found in identification of DNA damaging-agents with a broad range of chemicals. There are many other examples of using the concept of phenotypic anchoring for deriving differentially expressed genes and genomics classifier predictive of the final endpoints (Fielden et al., 2005).

Ung et al. qualitatively linked transcriptomics results to the changes of cell morphology, cell adhesion and macromolecular accumulation (fatty acid and glycogen) in zebrafish liver cell via gene set enrichment analysis (GSEA) (Ung et al., 2010). Bugiak and Weber also qualitative analyzed the relationship between gene expression level of a group of CYP genes and the deformity, fork length of zebrafish egg via RT-PCR for a variety of toxicants (Bugiak and Weber, 2010). Another example of phenotypic anchoring on ecotoxicity is the investigation on Green alga Scenedesmus vacuolatus between biochemical composition and cell volume changes, photosynthesis efficiency inhibition (Sans-Piche et al., 2010). Although there are many researchers who investigated the relationship between variety of gene expression level and organism level phenotypic endpoints, such as organisms weight, cytotoxicity assay and survival rate etc., most of them only performed qualitative analysis rather than quantitative analysis (Gottschalg et al., 2006; Ung et al., 2010; Bugiak and Weber, 2010; Sans-Piche et al., 2010; Ovando et al., 2010; Zheng et al., 2011; Powell et al., 2006).
2.3.4.3.2 Overall stress response

The cellular stress response pathways might be considered the guardians of the primary cellular infrastructure since their ubiquitous presence in all metazoan cells. These guardians are comprised of minimal components including the basic macromolecules such as nucleic acids, proteins, complex carbohydrate polymers and lipids as well as the superstructures made from macromolecules such as the nucleus, mitochondria, endoplasmic reticulum (ER), and lysosomes. Damages to these basic cellular macromolecules and infrastructures initiate activation of one or more stress response pathways in an effort to repair the damage to restore homeostasis, and the activation occurs at doses or exposure times lower than those required to elicit apical toxic events such as apoptosis or necrosis (Simmons et al., 2009; Kültz, 2009). In addition, the role in homeostasis of the stress response pathways also feed into the network whereby a cell decides (cell fate decision) to initiate apoptosis should efforts aimed at restoring homeostasis fail.

2.4. Quantitative Analysis of Toxicogenomics Data

In applying toxicogenomics-based assays to water quality assessment, it is imperative to be able to express the bioassay results in quantitative terms. Being able to express a bioassay result as a value allows, for example, quantitative comparisons and ranking among different water samples, or comparison between “before” and “after” samples to evaluate treatment efficacy. Although the toxicogenomics approach offers great potential for mechanistic toxicity screening and assessment (Krewski et al., 2010; Ankley et al., 2006; Waters and Fostel, 2004), there is none of the toxicogenomics tools currently available for regulatory
decision-making. The benefits of toxicogenomics information and tools in regulatory ecotoxicology are only starting to be elucidated. The challenges and inconsistency in the approaches to convert toxicogenomics information into quantitative toxicity endpoints stem from both scientific and practical factors. It is still not yet clear if molecular level effects can be indeed quantified and, if yes, how to quantify such effects in relevance to toxicity at different biological organizational levels (NRC, 2006, Altenburger et al., 2012). As a result, most risk assessors lack the technical guidance to readily convert this information into a decision-making framework. This situation is also due to numerous situational factors, as summarized by Lobenhofer et al., including a relatively complicated and constantly changing set of terms and definitions; a lack of consented approaches to transform the large amounts of data into a readily usable endpoints analogous to those conventional toxicity endpoints, and difficulty in interpreting the biological and regulatory significance of the information (Lobenhofer et al., 2002). Current approach for quantitate analysis of toxicogemomics result involve several steps as summarized below. First, how to define the quantifier for toxicogenomics result, as the signal of cellular response? Second, how to derive environmental readily usable index with the well-defined quantifier? And further incorporate these indexes into regulatory and risk assessment framework. Since battery test of bioassays is also prevalent in environmental monitoring application, and it can be seen as a “low-density omics”, several approaches for battery test of bioassay
index development are also summarized here. Attempts for quantitate assessment of mixture toxicity via toxicogenomics approach have also been included here.

2.4.1 Cellular Response Signal as Quantifier

Quantitative risk assessment and the elucidation of mechanisms of toxicity requires computational infrastructure and innovative analysis approaches that systematically consider available data at all levels of biological organization. The tradition quantification of cellular response can be classified into two groups: quantification based response intensity, or on range of response. Both approach involved in relative response compared to a matched control group.

The tradition objective of performing omics profiling is to identify biomarkers whose expression is altered by treatment with the agent under study. Intuitively, the most obvious candidates would be genes whose expression exhibited the greatest change between treated and control groups.

2.4.1.1 Signal as Intensity of Response

One approach to making this selection is to assume that those biomarkers showing the greatest fold increase or decrease in expression in the treated group relative to controls most likely associated with the treatment. For example, in microarray study, a gene might be considered differentially expressed if it is more than two fold increased. However, since the selection of a threshold is arbitrary it provides no level of confidence about statistical significance. Further, the reliability of the fold change depends on the spot intensity and may be less reliable for genes with low expression level.
When looking at toxicogenomics data, the most commonly used intensity signal is fold change, which is used to express the relative expression level of specific biomarkers (gene, protein, or metabolites). The scientists determine whether the biomarker is up-regulated (more produced) or down-regulated (fewer produced). For example, microarray data are usually expressed quantitatively as “-fold” changes in gene expression. Usually scientists arbitrarily select a two-fold change in expression as up-regulation or down-regulation. The red-green “heatmap” usually present in literature are used to illustrate patterns of gene expression observed in microarray experiments. These patterns of gene expression are sometimes referred to as a signature or a fingerprint, conveying that patterns seem peculiar to or diagnostic of different chemicals or groups of chemicals.

2.4.1.2 Signal as Range of Response

A statistically based approach that is frequently used is to calculate the mean and standard deviation of the distribution of log(2) intensity ratios and then select as differentially expressed, those genes whose expression level falls outside the 95% confidence interval (Quackenbush, 2002). This is a relatively simple procedure to apply and provides a statistically based method of selection. More recently ANOVA based analyses are being developed that determine the statistical significance of increase or decrease in the expression of level of individual genes (Kerr et al., 2000; Park et al., 2003). This will provide a solid statistical basis, based on p values for identification of differentially expressed genes. Statistically based selection of differentially expressed genes allows selection not only of genes exhibiting a large fold increase or decrease in expression level may be
small in magnitude, but can be measured with high precision. It is anticipated that many important genes will fall into this latter category. An important message for toxicologist is that the selection of differentially expressed genes will depend on a great extent on the method used, and statistically based methods provide the most reliable and unbiased way of making that selection. However, now microarrays probed only semi-quantitative information about changes in gene expression. Therefore, it is important to verify with quantitative PCR the expression level of a representative sample of genes identified on microarrays as being differentially expressed.

The number of differentially expressed genes is always used as another quantification approach for toxicogenomics data. To interpret the number of signals occurring in relation to the degree of contamination, Menzel and co-workers used the percentage of differentially expressed genes for quantification of the overall response in their study of contaminated sediments using microarray gene profiling in Caenorhabditis elegans (Menzel et al., 2001). Underlying notions of this approach is that a higher dose leads to higher number of responding genes or pathways (Judson et al., 2010). The percentage or number of differentially expressed genes has been later employed to quantify the molecular perturbation compared to control by other researchers in various species (Tilton et al., 2011).

Regardless of the method employed, the identification of differentially expressed genes is only a prelude to the real work of masking biological sense of the data. Depending on the particular experiment and issue being studied, the list could
contain hundreds or possibly more than a thousand genes. The task is to organize this mountain of data into a form that lends itself to meaningful interpretation.

2.4.2 Derivation of Molecular Toxicity Endpoints

Exposure to a toxicant is the basis for any risk assessment. In other words, exposure estimation is necessary in order to evaluate the likelihood of toxic event. Toxicological endpoints are values derived from toxicity test that are the results of specific measurements made during or at the conclusion of the test, and can be used to evaluate or predict the effects of toxic agents in natural environments. Since exposure is estimated by using quantification of toxic effect in a target at a specific concentration. It is important to differentiate between dose and concentration. Expression of toxicity as concentration allows the relative toxicity of different substances to be compared and normalizes for the variation in the size of targets exposure. To derive quantitative endpoints, a dose-response curve is required and it displays the central paradigm in toxicology, which is “the dose makes the poison” (Ottoboni, 1991).

2.4.2.1 Point of departure

Classical point of departure derivation, such as No/Low-Observed-Adverse-Effect Level (NOAEL/LOAEL), is determined experimentally by comparing the difference between the exposed population and its appropriate control population to see if there is any statistically or biologically significant increases are seen in the frequency or severity of adverse effect. A NOAEL is the highest exposure level at which no significant different were observed. In cases in which a NOAEL
has not been demonstrated experimentally, the term "lowest-observed-adverse-effect level (LOAEL)" is used, which this is the lowest dose tested.

Mathematical modeling, which can incorporate more than one effect level (i.e., evaluates more data than a single NOAEL or LOAEL), is sometimes used to develop an alternative to a NOAEL known as a Benchmark Dose (BMD) or Benchmark Dose Lower-confidence Limit (BMDL). In developing the BMDL, a predetermined change in the response rate of an adverse effect (called the benchmark response or BMR; generally in the range of 1 to 10% depending on the power of a toxicity study) is selected, and the BMDL is a statistical lower confidence limit on the dose that produces the selected response. When the non-linear approach is applied, the LOAEL, NOAEL, or BMDL is used as the point of departure for extrapolation to lower doses.

The most widely used point of departure derived from dose-response curve in toxicogenomics is No Observed Transcriptional Effect Level (NOTEL). NOTEL can be potentially used as an endpoint and regulatory benchmark for chemical screening, effluent toxicity testing, and environmental monitoring of toxicant.

Using similar concept, NOTEL was first proposed by Lobenhofer (Lobenhofer et al., 2004) for identifying the effects of low dose treatments and determining the threshold dose based on toxicogenomics data from a microarray assay. The dose-response curve, generated by the number of differentially expressed genes vs the toxicants concentration, allowed for the determination of NOTEL. The potential application of NOTEL in toxicology and risk assessment has been demonstrated and discussed by Ankley and Poynton et al. (Ankley et al., 2006; Poynton et al.,
Although NOTEL indicates the relative toxicity level of a compound, it does not fully reflect the rich and specific toxicant-induced genomic information that can be obtained from toxicogenomics assays (e.g., MOA, response at higher concentration above NOTEL). Development of more informative molecular toxicity endpoints is desirable and which requires collaborations among genomic scientists, bioinformatics experts and person involved in risk assessments.

**2.4.2.2 Toxic Potency as Effective Concentration, EC50, IC50**

Establish appropriate, measurable endpoints for all ecological contaminants of concern. Depending upon the dose of a given substance, a spectrum of undesirable effects may be observed that ranges from relatively begin effects at lower doses to serious or life threatening effects at higher doses. The choice of 50% effect/lethality as a benchmark avoids the potential for ambiguity of making measurements in the extremes and reduces the amount of testing required. However, the determination of effect concentration required the existence of a biological meaningful maximal response, which is usually 100% in death, retarded growth, reduce fecundity, etc. Regarding the cellular response in toxicogenomics, an objective effect concentration (ECx) value very often is impossible to obtain from gene-induction measurements, because it is not clear whether there is any fixed maximum induction value for each gene or gene sets, and if the maximum induction level of a gene vary for different toxicants (Dardenne et al., 2007; Gou and Gu, 2011).
2.4.2.3 Toxic unit and Toxic equivalency

Regarding these concentration based toxicity endpoints, it is obvious the lower the values, the greater the toxicity. Because this feature of toxicity data sometimes tends to confuse the non-scientist, an alternative way of expressing toxicity data is as function of the diluted sample. This form of expression is known as Toxic unit (TU) and is defined as follows:

\[ TU = \frac{100}{EC_x} \quad \text{Equation 2.1} \]

This has the advantage that toxicity increases with increasing values of TU. The toxic unit concept is usually used for effluent sample since the concentration of effluent is expressed as % dilution of original sample, not mass weight, or molar weight as single chemicals.

Similar approach, toxic equivalence (TEQ) concept, converts toxicity endpoints of a unknown and unidentified mixture of chemicals, such as an environmental water sample, to be expressed as equivalent concentration of a reference compound that exhibit similar mode of action and mechanism (Macova et al., 2010; Escher et al., 2008; USEPA, 2013).

\[ TEQ = EC_x(\text{reference}) \times TU(\text{sample}) \quad \text{Equation 2.2} \]

TEQ was translated from its original ligand-binding studies to integrative toxicity endpoints such as binding of dibenzodioxins to the arylhydrocarbon receptor and, later, was extended to other compounds that exhibit receptor-mediated toxicity mechanisms such as PCBs (Van den Berg et al., 1998; Tillitt, 1999). Other commonly used TEQs are the estradiol equivalent concentrations (EEQ) used for
estrogenic activity (Behnisch et al., 2001; Rutishauser et al., 2004), dihydrotosterone equivalent (DHTEQ) used in bioassays for androgenic endocrine activity (Wagner et al., 2013), the diuron equivalent (DEQ) used in bioassays for phytoxicity (Bengtson et al., 2005) and the chlorpyrifos equivalent (ChLEQ) used in bioassays for acetylcholinesterase inhibition (Poletika et al., 2003). Although the applicability of the TEQ concept has been mostly limited to receptor-mediated mechanisms when the condition of similar mechanisms was fulfilled, extension of this concept and application of TEQ for baseline toxicity to describe integrative ecotoxicity endpoint- algal growth rate inhibition, was demonstrated by Beate I. Escher et al (Escher et al., 2008) by fixing the dose-response curves with a fixed slope (Motulsky, 2005). TEQ was employed to report the genotoxicity assay of wastewater samples based on umuC activity using 4-nitroquinoline -N-oxide (4NQO) (-S9) and Benzo[a]pyrene (+S9) as reference compounds (Macova et al., 2010; Reungoat et al., 2010; Bi et al., 2011).

Application of TEQ concept for biomarkers or toxicogenomics-based bioassays is still in its exploration stage. In theory, if molecular perturbation (i.e. altered expression) can be quantified by parameters such as TELI described above, it is conceivable that mechanistic pathway-based molecular toxicogenomics assays endpoints can be potentially expressed as TEQ of a selected references chemical with similar MOA. For example, the endpoints of a toxicogenomics assay integrating known oxidative stress response pathways biomarkers could be possibly expressed as TEQ of a known model chemical that induces oxidative stress as main toxicity mechanism. This is analogous to the usage of H$_2$O$_2$ as the
reference chemical for oxidative stress potential and ROS production (Gou and Gu, 2011). DNA damage measured by altered expression changes in genes involved in DNA damage detection and repair via a biomarkers ensemble-based assay can be possibly quantified as TEQ of a reference DNA-damaging chemical or agent (Gou and Gu, 2011; Lan et al., 2015). The information rich and high-resolution toxicity effect fingerprints obtained from a toxicogenomics assay enables classification of samples based on the similarity in their toxicity profiles and thus underlying mechanisms, therefore facilitates the selection of reference chemical for various specific toxicity outcome at finer categorization resolution. More expensive investigations are needed to test against large number and variety of toxicants to confirm these hypotheses.

2.4.2.4 Endpoints for battery of bioassays

For battery test result, endpoints were derived individually for each assay, attempts have been made collapses results of multiply toxicity assay into a single number.

2.4.2.4.1 Potential Ecotoxic Effects Probe (PPEP)

Costan et al. (Costan et al., 1983) developed a potential ecotoxic effects probe (PPEP) index as a simple effects-based hazard assessment scheme to assess industrial and municipal effluent toxicity. PPEP index express the toxic loading of each as a single numerical value that integrates both its toxic potential (determined with a battery of small scale bioassays representing different biological levels and type of toxic effects) and its flow. The PPEP index is calculated by adding the TU of all endpoints, according to the following formula:
\[ P = \log_{10}[1 + n\left(\frac{T_i}{N}\right)Q] \] \textit{Equation 2.53}

where \( P \) is the PEEP numerical value, \( n \) is the number of bioassays exhibiting (geno) toxic responses, \( N \) is the maximum number of obtainable geno toxic responses, \( T \) is the TUs (toxic units) given by a particular bioassay before or after biodegradability testing of the effluent sample, and \( Q \) is the effluent flow (m\(^3\)/h).

The development of PEEP index was intimately linked with the Saint-Lawrence River Action Plan (SLAP), initiated in 1988 by the Government of Canada as part of a national commitment to sustain the biodiversity of its major aquatic environmental (Environment Canada, 1996). Under the first two SLAP (1988-98), the PEEP index was employed to determine and compare the relative toxicity of 106 priority industrial sites all of which discharged their wastewaters to the Saint-Lawrence River (Costan et al., 1983; Environment Canada, 1996; Blaise, 1996; Blaise et al., 2000). The index has also been employed in Australia, France, Lithuania, Janpa and South Africa.

\textbf{2.4.2.4.2 Integrated Biomarker Response (IBR) Index}

Beliaeff \textit{et al.} (Beliaeff and T. Burgeot, 2002) and Guerlet \textit{et al.} (Guerlet et al., 2010) used star plots to display results for the panel of biomarkers used and computed the star plot area as integrated biomarker response (IBR). The IBR was determined based on inner areas delimited by a solid line joining radius coordinates as shown in Figure 2.1. This type of plot not only serves as a visual representation but can also be used to indicate a global or specific toxicity response. Many researchers have shown the application of IBR as a useful tool for
environmental assessment (Damiens et al., 2007; Oliveira et al., 2009; Serafim et al., 2012).

**Figure 2.1.** Examples of a biomarker star plot for four biological markers on mussel (Mytilus edulis) collected in the Baltic sea (AChE=acetylcholinesterase; CAT=catalase; GST=glutathione-S-transferase measured in mussel digestive gland [suffix dg] and gills [suffix g]). Data from each assay are first standardized with the general mean and standard derivation of that given assay among different sampling sites, then star plots are used to represent the scores (standardized data) of the various biomarkers used at a given sampling site as show in the figure. The integrated biomarker response (IBR) is determined based on inner areas delimited by the solid line joining radius coordinates as shown in figure. Reproduced with permission from Beliaeff et al 2002 Environ. Toxicol. Chem. 21 1316. Copyright (2002) by the John Wiley and Sons.
The relation between star plot features and corresponding sample or environmental conditions is highly dependent on a number of a priori choices. The selected biomarkers usually correspond to the objectives of ecotoxicological study and successful application of IBR depends on priori justification for each biomarker used, as well as the prior knowledge known of the physiological significance of the changes to each biomarker. The later is evaluated against the specific objectives of research or monitoring and the characteristics of the environmental samples. Although IBR, as an indicator of environmental stress, appears to be a useful tool for scientists and managers in assessing ecological risk, cautions should be taken when applying IBR. For example, the areas can be sensitive to the type of method used for data standardization and the result may vary with different biomarker position on a star plot.

2.4.3. Mixture Toxicity Assessment

Quantitative analysis of mixture toxicity assessment approach have been a critical problem since many research challenges remain for mixture toxicology, as discussed by a number of researchers (de Zwart and Posthuma, 2005; Altenburger et al., 2003; Syberg et al., 2009; Borgert et al., 2004). There is still a lack of sound theoretical basis for the mechanisms and interactions of mixture toxicity, which inheritably limit both research and policy in an interdependent manner. The unresolved number and complexity of mixture exposures and, the reliance on highly standardized short-term integral adverse effects-based bioassay, have not significantly improve our poor understanding of the relationship between exposure-based effects and internal-received dose metrics (European Commission,
Current predictive modeling of ecological mixture toxicity is adopted from pharmacology (Berenbaum, 1989) and is based on two different theories, which are not mutually exclusive (Greco et al., 1995).

Numbers of experimental studies from the past decade addressed diagnostic and/or mechanistic questions regarding the combined effects of chemical mixtures using toxicogenomic techniques. Visualization tools such as Venn Diagrams, heat maps, or fold expression graphs, are often used for comparative displaying of observations from mixture and individual. Bioinformatics tools, such as hierarchical cluster on gene ontology term, principal components analysis, gene set enrichment analysis etc. were also utilized in recent publications to query observations for toxicogenomics information of mixture toxicity (Vandenbrouck et al., 2010). In these studies, no explicit hypothesis were provided on what an expected combined effect from a mixture exposure would look like, and no quantitative view on the data which allow extrapolation to other mixture compositions.

Mixture toxicity studies with attempt to apply mixture model theories based on quantitative gene expression data are, to our best knowledge, scarce from the literature. For example, quantitative analysis of combined effects, such as linear regressions, were rarely performed (Krasnov et al., 2007; Filby et al., 2007; Zhang et al., 2011). One of the reasons for this gap might be that an objective effect concentration ($EC_x$) value very often is impossible to obtain from gene-induction measurements, because it is not clear whether there is any fixed maximum induction value for each gene or gene sets, and if the maximum
induction level of a gene vary for different toxicants (Gou and Gu, 2011; Dardenne et al., 2007). In recent years, attempts have been made to address this problem. Dardenne et al. (Dardenne et al., 2007) explored the molecular mixture effects for 8 different binary mixtures in dilution series via toxicogenomics assay using 14 stress response related gene reporters using E.coli K12 SF1 deletion library. Differentially expressed genes that were common among individual chemicals and mixture were identified and, dose response curves were developed as fold changes versus concentration. The compliance of molecular mixture toxicity with the current model framework was explored by examination of the prediction of fold change of specific genes in mixture from those in individual chemical. Various addition models were tested and the results demonstrated that at single gene level, the additive mixture model framework might work at molecular level.

2.5. Environmental Application of Toxicogenomics for Water Toxicity Assessment and Quality Monitoring.

The high throughput nature and high-resolution mechanistic information can greatly advance the efficiency of regulatory toxicology by enabling more frequent data collection and toxicity fingerprinting, reducing the uncertainty in risk assessment and optimizing the resource utilization on cases that present the greatest potential risk (Waters and Fostel, 2004; Nuwaysir et al., 1999; Hamadeh et al., 2002; Boverhof et al., 2006; Jayapal et al., 2010; Neumann and Galvez, 2002). One clear advantage of toxicogenomics techniques is in the mode of action (MOA) profiling of toxicants in the assessment of exposure to environmental
stressors. The use of gene (or protein) expression profiling-based mode of action will allow the identification of new toxic substances, by comparing their gene expression profiling with model compounds, as well as molecular biomarkers. The development of toxicogenomics will be particularly useful in toxicity determination of increasing numbers of emerging pollutants and limiting the use of animals for testing purposes. In addition, toxicogenomics techniques detect changes on molecular level and thus are more sensitive towards early warning signs, since cellular stress responses are activated before actual harm occurs, and they indicate the presence of associated stressors (Coen et al., 2004; Hoyt et al., 2003).

Over the last decade, as alternatives to WET type of test that rely on whole animal, toxicogenomics techniques have been successfully applied for toxicity evaluation of various environmental samples including chemical mixtures, industrial waste water, drinking water and wastewater effluents from different treatment processes, as examples shown in Table 6. It could be found that the indicator species employed are very diverse such as yeast, frog, fish and human cells. It is because that environment risk assessment concerns millions of species and it’s impossible to find a single species that is representative for all others. In reality we usually select well-known species, such as budding yeast *Saccaromyces cereviase*, the fish *Oncorhynchus mykiss* (commonly named rainow trout), the frog *Xenopus laevis* and the water flea *Daphnia magna* for standard protocols. For these model organisms, their well-annotated genomic information can help us mapping the effects of a toxicant or a stressor on the affected biological processes (Hoyt et
al. , 2003; Momose and Iwahashi, 2001). However, they may be not the best representatives for their community as a whole. For example, *Daphnia*, which is usually used as a standard test species in both acute and reproduction tests. Under normal laboratory conditions, a nonsexual reproductive strategy (i.e., parthenogenesis) is applied by *Daphnia*. This means that males are absent, and that partner finding, sexual synchronization, mating behavior, etc. are not considered. Hence, hazard identifications partly based on *Daphnia* tests may be of limited value for sexually reproducing species in the environment (Breitholtz et al., 2006).

For this reason, the current toxicogenomics methods, along with other ecotoxicological testing, where a few species are representatives for such diverse groups is of course a gross simplification of an ecosystem (Breitholtz et al., 2006). One approach is to identify stress-induced genes and use them as biomarkers because stress-induced genes are conservative cross species. But further researches are still need on toxicity test species and select more representative species to replace or replenish the species currently applied. Besides, the genetic information, which is lacking for most species of interest, must be developed in the process.

From Table 6, it could be found that microarray and RT-PCR are the most commonly used toxicogenomics techniques. Both technologies offer comprehensive information rather than a single endpoint. However, there is no standardized framework available to provide guidance on how to analyze these hundreds to multiple thousands of data points. A natural first step is to calculate
gene expression change before and after exposure to stress and examine the extremes. However, such analyses do not provide any MOA information and exhibit the full potential of toxicogenomics techniques that may further our understanding of cellular biology. So the next step is to elucidate the biological function of these differentially expressed genes and their related pathways or networks. There are several tools available for gene annotation or pathway/network analysis, Gene Ontology (GO), Kyoto Encyclopedia for Genes and Genomes (KEGG), Munich Information Center for Protein Sequences, GenBank, Ensembl, the Human Gene Organization, TRANSFAC, and TRANSPATH databases (Afshari et al., 2011). In most situations, toxicogenomics data is often a puzzle and require well-trained experts and advanced capabilities to find out the secrets under cover. Bioinformatics is widely recognized as the limiting step in the interpretation of the wealth of data generated using by omics techniques (Miracle and Ankley, 2005).

Another kind of specific hypothesis-driven studies is battery test of bioassay, brought up by Sanchez et al. (Sanchez et al., 1988) and extended by Escher et al. (Escher et al., 2008; Escher et al. 2014; Macova et al., 2011; Macova et al., 2010) as mentioned in former paragraph. Battery of in vitro bioassays (i.e. biomarkers-based assays) have been successfully applied for environmental water samples (Allinson et al., 2011; Fang et al., 2012; Leusch et al., 2014). However, although each individual bioassay in the battery is able to indicate a specific toxicity, how to integrate these separate results to one endpoint evaluating the overall toxicity is an unresolved problem. In addition, these battery tests consisting of a bunch of
separated bioassays still require substantial amount of time and efforts. In addition, limit information were provided on the toxicity mechanisms and profiles.

An additional challenge to apply toxicogenomics on toxicity evaluation and risk assessment is the fact that a change in gene expression does not necessarily lead to a change in protein expression or an adverse event. For this reason, toxicogenomics data must be validated by correlated to other biological events (“phenotypic anchoring”) in order to fully understand the toxicological meaning behind the changes in gene expression (Hahn, 2011). In current stage of toxicogenomics application in water quality assessment, few studies have been done on the correlation between the expression profile of a chemical with a known endpoint or phenotypic effect.

There are also many other challenges remaining in the application of toxicogenomics. For example, in actual aquatic environment, pollutants exist at low concentrations. So concentration is needed in order to derive the dose-response curves. In this case, the original toxicity may be altered by the loss of contained contaminants (e.g. volatile compounds) or introduction of extra ones (e.g. organic solvent used during extraction methods) (WHO, 1997).

Publications on toxicogenomic research are not consistent in level of detail reported on experimental and analytic methods used, such as test organisms, durations of exposure, data analysis methods and so on (Ball et al., 2002; Fielden et al., 2008). Indeed, there will be multiple challenges for regulatory agencies to overcome to integrate toxicogenomics techniques and apply genomic data in
environment risk assessment. However, this just highlights the relatively immature stage of the field and cannot deny the fact that toxicogenomics is becoming a useful research tool particularly for water quality monitoring. The field of toxicogenomics has made headway in understanding the toxicity mechanism of contaminants. To express its full potential, researchers will need to address multiple problems at hand. The imperative aspects that require focused efforts are building fundamental knowledge, theoretical framework and developing necessary bioinformatics tools for quantifying the toxicogenomics data to link measured molecular changes to physiological and toxicological outcomes (e.g. functional analyses). There is also a clear need for more interaction between scientists developing the technologies, bioinformatics, toxicologists and ecotoxicologists, in order to expedite the application of these techniques to urgent environmental problems (Calzolai et al., 2007). One of the most pressing requirements is the standardization of data collection and analysis. Collaboration is also needed to efficiently link expression signatures to the biological end points that are typically measured and to develop bioinformatics tools that consider possible changes in thousands of data points per sample. The new molecular technologies offer sensitivity and specificity, required to determine causality, and in addition, offer an “open” unbiased assessment of the exposure, without the need of a priori knowledge of the conditions. In theory, these tools should improve risk assessment, even though many hurdles still lie ahead.
### Table 2.2 Applications of Toxicogenomics Technologies In Water Toxicity Assessment

<table>
<thead>
<tr>
<th>Sample</th>
<th>Species</th>
<th>Bioassay</th>
<th>Biomarkers</th>
<th>Result exhibition</th>
</tr>
</thead>
<tbody>
<tr>
<td>WWTP influents and effluents</td>
<td>Rainbow trout hepatocytes</td>
<td>RT-PCR</td>
<td>8 stress-response related genes (xenobiotic biotransformation, estrogenicity, heavy metal detoxification and oxidative stress)</td>
<td>Gene expression fold change</td>
</tr>
<tr>
<td>(Gagne et al., 2013a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oil sand water, lixiviate and processed water (Gagne et al., 2013b)</td>
<td>Rainbow trout hepatocytes</td>
<td>RT-PCR</td>
<td>16 genes (xenobiotic biotransformation, metal homeostasis, oxidative stress, estrogenicity, DNA repair, cell growth and glycolysis)</td>
<td>Gene expression fold change</td>
</tr>
<tr>
<td>WWTP effluents, bisphenol A, chlorpyrifos and methylparaben (San Segundo et al., 2013)</td>
<td><em>Xenopus laevis</em> frogs embryo</td>
<td>RT-PCR</td>
<td>4 genes involved in early development and 1 gene involved in the general stress response</td>
<td>Gene expression fold change</td>
</tr>
<tr>
<td>WWTP effluents (Ings et al., 2011)</td>
<td>Rainbow Trout</td>
<td>Microarray</td>
<td>147 genes categorized into functional groups including, metabolism, transport, signaling, chaperones, structural, oncogenes, endocrine, immune and developmental-related genes</td>
<td>Gene/Protein expression differences from control (p-value)</td>
</tr>
<tr>
<td>Textile mill effluents (Kim et al., 2006)</td>
<td>Yeast</td>
<td>Microarray</td>
<td>Global examination of 5956 genes</td>
<td>Gene expression fold change</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RT-PCR</td>
<td>6 stress-response related genes (plasma membrane, oxidative stress, DNA damage)</td>
<td></td>
</tr>
<tr>
<td>Study Description</td>
<td>Species/ Organism</td>
<td>Methodology</td>
<td>Number of Genes/ Measurements</td>
<td>Analysis Method</td>
</tr>
<tr>
<td>----------------------------------------------------------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>-------------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Industrial plant effluent (Moens et al., 2007)</td>
<td><em>Cyprinus carpio</em></td>
<td>Microarray</td>
<td>960 hormone-responsive and gender-related genes</td>
<td>Gene expression fold change</td>
</tr>
<tr>
<td>Chemical mixtures and contaminated groundwater (Garcia-Reyer et al., 2012)</td>
<td><em>Daphnia magna</em></td>
<td>Microarray</td>
<td>Global examination of ~15000 genes</td>
<td>Network analysis, Functional enrichment analysis using the GO annotations</td>
</tr>
<tr>
<td>Effluents from copper mine (Poynton et al., 2008)</td>
<td><em>Daphnia magna</em></td>
<td>Microarray</td>
<td>16 differentially expressed genes</td>
<td>Gene expression fold change</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RT-PCR</td>
<td></td>
<td>NOTEEL</td>
</tr>
<tr>
<td>WWTP effluents (Hara-Yamamura et al., 2013)</td>
<td>Human hepatoma cell line, HepG2</td>
<td>Microarray</td>
<td>Global examination of 5000 genes</td>
<td>Functional enrichment analysis using the GO annotations</td>
</tr>
<tr>
<td>Wastewater treatment plant (WWTP) effluents, surface water (Skelton et al., 2013)</td>
<td>Male fathead minnows (Pimephales promelas)</td>
<td>Metabolomics by NMR spectroscopy and GC-MS</td>
<td>Principal components analysis (PCA)</td>
<td></td>
</tr>
<tr>
<td>WWTP effluents (Fukushima et al., 2014)</td>
<td>Human hepatoma cell line, HepG2</td>
<td>Comet Assay Microarray RT-PCR Metabolomics by FT-MS</td>
<td>Global examination of 8795 genes</td>
<td>Number of altered genes, Biological functional analysis by DAVID</td>
</tr>
<tr>
<td>Drinking water (Zhang et al., 2013)</td>
<td>Mice (<em>Mus musculus</em>)</td>
<td>Microarray</td>
<td>Global examination</td>
<td>Gene expression fold change</td>
</tr>
<tr>
<td>Surface water (Sellin Jeffries et al., 2012)</td>
<td>Fathead minnows (Pimephales promelas)</td>
<td>Microarray</td>
<td>Global examination of 15000 genes</td>
<td>Functional enrichment analysis using the GO annotations</td>
</tr>
</tbody>
</table>

* WWTP: Wastewater treatment plant; FT-MS: Fourier transform mass spectrometer
Chapter 3

A New Transcriptional Effect Level Index (TELI) for Toxicogenomic-based Toxicity Assessment

3.1 Abstract

Examining global effects of toxins on gene expression profiles is proving to be a powerful method for toxicity assessment and for investigating mechanisms of toxicity. To explore the application of toxicogenomics approach for toxicity evaluation and environmental monitoring, this work developed an whole cell array with stress gene ensemble by use of a cell-array library of 91 E. coli K12 strains with transcriptional green fluorescent protein (GFP) fusions covering most known stress response genes. This study proposes and demonstrates the potential application of a new Transcriptional Effect Level Index (TELI) to convert the information-rich toxicogenomic data into integrated and quantitative endpoints. A library of transcriptional fusions of green fluorescent protein (GFP) that includes different promoters for 91 stress-related genes in E. coli K12, MG1655 is employed to evaluate the gene expression alteration induced by exposure to variety of chemicals. TELI is determined for each toxicogenomic assay, and it incorporates the number and identity of genes that had altered expression, the magnitude of alteration, and the temporal pattern of gene expression change in response to toxicant exposure. TELI values exhibit a characteristic “sigmoid” shaped toxicity doseresponse curve, based on which TELIMAX (the maximal value of TELI), TELI50 (concentration that yields half of TELIMAX),
NOTELELI (TELI-based no observed transcriptional effect level), and SlopeTELI (the slope of TELI-dose response curve) are obtained. TELI-based endpoints are compared to currently used endpoints such as EC50 and no observed transcriptional effect level (NOTEL). The agreement of NOTELELI and NOTEL values validates the concept and application of TELI. Multiple endpoints derived from TELI can describe the dose response behavior and characteristics more completely and holistically than single points such as NOTEL alone. TELI values determined for genes in each stress response category (e.g., oxidative stress, DNA repair) indicate mode of action (MOA)-related comparative transcriptional level toxicity among compounds, and it reveals detailed information of toxic response pathways such as different DNA damage and repair mechanisms among the NMs. This study presents a methodology for converting the rich toxicogenomic information into a readily usable and transferable format that can be potentially linked to regulation endpoints and incorporated into a decision-making framework.

3.2 Introduction

The concern of emerging contaminates, such as endocrine disrupting compounds (EDCs), pharmaceuticals and personal care products (PPCP) and nanomaterials (NMs), are anticipated to greatly increase the demands for their ecological effects and risk assessments, as indicated by various new toxicity regulation programs, such as the EPA Tox21 in U.S. and Registration, Evaluation, Authorization and Restriction of Chemicals (REACH) program within the EU (Colborn, Saal, &
Soto, 1993; Daughton & Ternes, 1999; Nel, Xia, Madler, & Li, 2006; Schmidt, 2009; Williams, Panko, & Paustenbach, 2009). Presently, virtually all ecotoxicological tests rely on whole animal exposures with endpoints derived from adverse effects such as survival, growth, and reproduction (e.g. EPA Whole Effluent Toxicity (WET))(USEPA, 2000). These standard tests have been proved to be valuable and relatively efficient in risk assessments and regulatory decision-making. However, it is also recognized that these tests are resource- and time-intensive (e.g. WET test takes up to 96 hrs for acute test and up to 10 days for chronic tests). Unless the approaches can be revised, the time required to handle the anticipated toxicity testing efforts for the large and ever-increasing number of emerging contaminants will be measured in decades(Snape, Maund, Pickford, & Hutchinson, 2004). In addition, there is also an evident growth in the complexity of testing in respect to the environmental sample matrix (e.g. effluents, sediments, synergistic effects of mixtures). Therefore, an urgent demand exists for less costly and more rapid, yet informative and reliable ecotoxicity screening and testing methods.

Recent advancements in the emerging field of toxicogenomics, which examines the global molecular-level activity in response to environmental stressors, provides a significant advance in toxicant evaluation and understanding toxicity mechanisms and pathways(Kawata, Yokoo, Shimazaki, & Okabe, 2007; Newton, Aardema, & Aubrecht, 2004; Yamanaka, Toyoshiba, Sone, Parham, & Portier, 2004). In addition to the better determination of Mode of Action (MOA), another most significant advantage of toxicogenomics is that it provides crucial and
multiple information to help reducing the uncertainty in risk assessment of chemicals (Ankley et al., 2006). The cellular-level and sub-lethal impact observed may also be more indicative of chronic effect than phenotype endpoints. Furthermore, the toxicogenomic data can be potentially used to identify and classify compounds with similar MOAs and gain diagnostic insights into identifying of causal agents (e.g. comparing to “reference” toxicants with established MOAs). At last, incorporation of toxicogenomic-based screening assays into ecological risk assessment framework can help guide and optimize the resources and minimize the animal use (Ankley et al., 2006).

Applications of toxicogenomics on environmental toxicity assessment have been demonstrated. Poynton et al. showed that that gene expression analysis with *Daphnia magna* could predict environmental exposure to metals in effluent from two copper mines in California (e.g Cu, Pb, and Zn) (Poynton, Loguinov, et al., 2008; Poynton, Zuzow, Loguinov, Perkins, & Vulpe, 2008). Recently, our group showed that the toxicogenomics approach using recombinant whole-cell arrays could also be employed for toxicity evaluation and potential identification of NMs (Gou, Onnis-Hayden, & Gu; Onnis-Hayden et al., 2009). Compare with traditionally microarray technology whose procedure includes multiple steps including RNA isolation, PCR amplification, labeling and hybridization, our direct GFP signal detection method is simpler and faster, and therefore it is more feasible for screening a large number of chemicals. In addition, it provides multi-dimensional transcriptional level effect information, by adding a temporal dimension to the gene expression data, and therefore can more accurately reflect
the chemical-induced cell responses that are time-dependent (Gou et al.; Onnis-Hayden et al., 2009).

However, challenges remain on the application of toxicogenomics for ecological assessment and regulatory decision-making. One of the greatest initial challenges is how to convert the rich toxicogenomic information into readily usable and transferable format that can be potentially link to regulation endpoints and incorporated into decision-making framework. NOTEL is the only accepted endpoint that has been used with toxicogenomic data and it indicates the threshold for the altered gene expression induced by a toxicant to be observed (Lobenhofer et al., 2004). Although NOTEL indicates the relative toxicity level of a compound, it does not fully reflect the rich and specific toxicant-induced genomic information that can be obtained from toxicogenomic assays (e.g. MOA, response at higher concentration above NOTEL). Furthermore, there is no accepted quantitative toxicogenomic endpoint that incorporates the important factor---time-dependence of the genomic response.

In this study, we propose and validate a Transcriptional Effect Level index (TELI) that incorporates both the number and magnitude of genes with altered expression induced by chemicals, as well as the temporal pattern of response. Dose-response curves of TELI as function of concentrations are established and the related toxicity endpoints derived from the TELI dose response curves are proposed. To validate the proposed quantification method, the derived TELI from toxicogenomic data are compared to those previously established endpoints including NOTEL, EC50 (half maximal effective concentration), and BOD.
(Biological Oxidative Damage)(Bello, 2009). The potential application and importance of the proposed quantitative ecotoxicogenic endpoints are discussed. A number of nanomaterials are used as representative pollutants due to their recognized as well as unknown environmental impact and health risks(Gou et al.; Nel et al., 2006).

3.3 Material and Methods

3.3.1 Nanomaterials tested for toxicity

Nano silver (nAg, ~60 nm, NanoDynamics Inc., Buffalo, NY, USA, 1424), nano titanium dioxide rutile (nTiO$_2$-r, ~10 nm thick, 40 nm laterally, Sigma-Aldrich, 10024JH), nano titanium dioxide anatase (nTiO$_2$-a, ~10 nm, NanoStructured & Amorphous Materials, Houston, Texas, USA, 5425HT), and fullerene soot (M.E.R.Co, tuscon, AZ, USA) were prepared in M9 medium (minimum medium for bacteria culture) for a stock concentration of 1 mg/mL, which contains 1% of crude Bovine Serum Albumin (BSA) (ACROS, NJ, USA) as a dispersant. The stock solutions are sonicated in a High energy Cup-sonicator, at ~90 Watt power for at least 15 mins to maintain a better dispersion before the toxicity assays. Detailed physical and chemical characterization of these same nanomaterials used are described by Bello et. al.(Bello, 2009) and in Table 3.1.
Table 3.1 Physical and chemical Characterization of the nanomaterials tested in this study (Bello, 2009)

<table>
<thead>
<tr>
<th>Description</th>
<th>Source; (CAT #)</th>
<th>Primary particle size</th>
<th>Zeta potential (mV)</th>
<th>Aggregation size a (nm)</th>
<th>SSA m²g⁻¹</th>
<th>ROSb /mg H₂O₂ Equi</th>
<th>BOD c (TEUs, µmol L⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nano TiO₂ (Anatase)</td>
<td>Nanostructured &amp; Amorphous Materials, Houston, Texas, USA; (5425HT)</td>
<td>10 nm</td>
<td>-33.9</td>
<td>1040 nm</td>
<td>274.2</td>
<td>256.57</td>
<td>64.9</td>
</tr>
<tr>
<td>Nano TiO₂ (Rutile)</td>
<td>Sigma-Aldrich; (10024JH)</td>
<td>10 nm thick, 40 nm laterally</td>
<td>-24.1</td>
<td>5580 nm</td>
<td>189.9</td>
<td>252.13</td>
<td>12.6</td>
</tr>
<tr>
<td>Nano silver</td>
<td>NanoDynamics Inc., Buffalo, NY, USA; (1424)</td>
<td>~60 nm</td>
<td>-26.2</td>
<td>984.1 nm</td>
<td>8.4</td>
<td>381.80</td>
<td>116.4</td>
</tr>
<tr>
<td>F_soot</td>
<td>M.E.R. Co., Tuscon, AZ, USA; (MRST)</td>
<td>&gt;20 nm</td>
<td>-39.1</td>
<td>191.3 nm</td>
<td>194.3</td>
<td>235.07</td>
<td>62.1</td>
</tr>
</tbody>
</table>

a: Aggregation Size are measured by Dynamic Light Scattering
b: ROS production are measured by Dichlorofluorescin diacetate (DCFH) assay, and expressed in H₂O₂ equivalent unit
c: BOD - Biological oxidative damage measured as a decrease in the total antioxidant capacity of human serum by the FRAS assay and expressed in Trolox Equivalent Units (TEUs, µmol L⁻¹, µM).

3.3.2 Measuring the temporal gene transcriptional activity in E.coli.

A library of transcriptional fusions of green fluorescent protein (GFP) that include different promoters for 91 stress-related genes in *E.coli* K12, MG1655 is employed in this study and the detailed information for the library construction and library validation can be found in previous reports (Zaslaver et al., 2006; Zaslaver et al., 2009). The selected stress genes and their main functions are described elsewhere and given in Table 3.2. In this library, each promoter fusion is expressed from a low-copy plasmid, pUA66 or pUA139 that contains a
kanamycin resistance gene and a fast folding gfpmut2 gene, therefore allows for continuous and real time measurements of the promoter activities. The stability of plasmid and the identity of the promoter regions were verified by the inventor of this library (Zaslaver et al., 2006).

**Table 3.2.** Genes included in the stress cell-library for chemical-induced *real time* gene expression profiling analysis (Onnis-Hayden et al., 2009).

<table>
<thead>
<tr>
<th>Category</th>
<th>Gene selected</th>
<th>Known functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Redox stress</td>
<td>soxS, soxR, oxyR, inaA, dps, ahpF, katG, sodA, ahpC, katE, sodB, norR, fpr, tam, yeiG, yniC, uspB, sodC, gst, zntA, yeaE</td>
<td>Increased levels of superoxides, increased levels of peroxides, any other conditions, which alter the redox potential of the cell. Genes response to oxidative stress</td>
</tr>
<tr>
<td>(Membrane Stress)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy stress</td>
<td>sdhC, cyoA</td>
<td>Perturbations of electron transport and exposure to uncoupling agents, which affect ATP levels in the cell</td>
</tr>
<tr>
<td>Cold shock</td>
<td>cspA, cspB</td>
<td>Temperature downshift</td>
</tr>
<tr>
<td>Cell killing</td>
<td>dinJ, slyA, yeeV, yfjG, relB</td>
<td></td>
</tr>
<tr>
<td>General function</td>
<td>phoB, crp, cdaR, ydeO, ybgI, gadX, ompC</td>
<td></td>
</tr>
</tbody>
</table>

To measure the transcriptional level effect induced by the toxicant, bacteria are grown in black 96-well plates with clear bottom (Costar, Bethesda, MD, USA) for 2 hours at 37°C until the cultures reached early exponential growth in M9 media (OD600 about 0.1). Duplicate tests for each treatment were performed. NM stock solution is added into the microplate well for the targeted concentrations. Then the plate is put in a Microplate Reader (SynergyTM Multi-Mode, Biotech, Winooski, VT, USA) for simultaneous absorbance (OD600) measurement (cell growth) and fluorescent readings (GFP level, EX 485nm, EM 528nm) at a time interval of 3 minutes for 2 hrs. We chose 2hr- exposure time for this study because our intention was to develop a relatively fast toxicity assessment and screening methods for evaluating a large number of contaminants.

*Data Processing and Analysis*

All data are corrected for various controls, including blank with medium control (with and without NMs) and promterless bacterial controls (with and without NMs). The alteration in gene expression, also called induction factor I (I=Pe/Pc), for a given gene at each time point due to chemical exposure, is represented by the ratio of the normalized gene expression GFP level (Pe=(GFP/OD)_{experiment}) in the experiments condition with NMs exposure to that (Pc= (GFP/OD)_{control}) in the control condition without any NMs exposure. Then the natural log of I value (ln(I)) at every time point is compiled for further analysis. For down-regulated genes, absolute Ln(I) values [Ln(I)]_{abs} are applied to convert all altered transcriptional effect level to positive values. A conservative cut-off background
noise threshold value of 0.4 ([ln(I)]_{abs}=0.4) is chosen based on previous reproducibility and control tests (Onnis-Hayden et al., 2009).

3.3.3 Concept and the determination of TELI

To quantify the transcriptional effect level induced by a given toxicant using the toxicogenomics data, we developed a Transcriptional Effect Level index (TELI) to convert the information-rich toxicogenomic data into an integrated and quantitative endpoint. The TELI considers and incorporates three factors, (1) the number and identity of genes that had toxicant-induced expression change, (2) the magnitude of altered gene expression for each gene response to the toxicant exposure, and (3) most importantly, the time factor, namely, the temporal pattern of gene expression change. Figure 3.1 illustrates the conceptual construction and determination of TELI for each toxicity assessment. Figure 3.1(A) shows an exemplary temporal pattern of altered gene expression (transcriptional effect) level for a given gene (fnr), and we have shown that the temporal pattern varies for different genes (Gou et al.). For TELI determination, in order to quantify and compare the level of differentially expressed genes, we used absolute change folds values. For example, up-regulation 2 folds (Ln(I)=0.69) will be considered as the same magnitude of change for down-regulation of 50% (Ln(I)=-0.69, [Ln(I)]_{abs}=0.69). Background gene expression level as 1 is subtracted at every time point. Then, the accumulative transcriptional effect of a given gene over a 2 hours exposure period is determined as the area (in blue color) defined by the curve over the X-axis as shown in Figure 3.1(A), which is calculated using the following equation
Equation (1)

\[ TELI_{(genei)} = \left( e^{\frac{\ln(I)}{t=0}} \right) \left( e^{\frac{\ln(I)}{t=2hr}} \right) - 1 \]

To determine the TELI that reflects the overall transcriptional level response of the prokaryotic cell in exposure to the toxicant for a given time period, and over a range of genes (indicative of global response for most genes in the genome or, for any specific functional categories, e.g. stress genes in our library), the overall 3-D accumulative transcriptional effect levels for all the genes of interest is integrated as the volume of the “mountain peaks” as shown in Figure 3.1(B), which is calculated as following equation:

\[ TELI_{(total)} = \sum_{i} TELI_{(genei)} \]

Equation 2

3.3.4 Determination of EC50 and NOTEL

The conventional 24h EC50 values for the three NMs selected are determined using the same E.coli strain based on growth rate inhibition. E.coli K12 were incubated at 37°C for about two hours to obtain OD of about 0.1 before the NMs stock solution was added to obtain various NMs concentrations on microplates. Duplicates were performed for each concentration. After 24 hours incubation, both colony-forming unit (CFU) counting and absorbance (600nm) measurements were conducted to determine the extent of cell growth inhibition at various NM concentrations. EC50 was then calculated based on the dose-response curve of growth inhibition (%) using four-parameter logistic nonlinear regression model equation (GraphPad PRISM 5, La Jolla, CA 92037).

NOTEL has been proposed and applied for toxicogenomics-based toxicity assessment(Gou et al.; Poynton, Loguinov, et al., 2008). We apply the concept of
NOTE as the maximal concentration of a chemical at which less than 5% of the genes in the library are differentially expressed upon chemical exposure compared to control. The NOTE is also determined by a dose-response curve of the percentage of genes that had altered expression in our “stress library” versus the dose concentration of chemicals, using the four-parameter logistic nonlinear regression model equation as described above.

3.4 Results and discussion

3.4.1 TELI as a quantitative toxicogenomic-based toxicity assessment endpoint

Figure 3.1. Determination of Transcriptional Effect Level Index (TELI), as an endpoint for toxicogenomic based toxicity assessment. (A). Exemplary
temporal altered gene expression level \((\text{Ln}(I)_{\text{abs}})\) for a given gene over an exposure time of 2 hrs. Integration of the altered gene expression level over time, indicated as highlighted area, was calculated as accumulative transcriptional effect level for a given gene. Y-axis: altered gene transcriptional effect level, \(\exp([\text{Ln}(I)]_{\text{abs}})\), a transformed value from induction factor I which is calculated as the ratio of gene expression level in the experiment with toxicant to that of the control without any toxicant exposure. (B). 3-D altered gene expression profile compiled from the altered gene expression level for all the genes in the test library. The TELI value was determined as the sum of accumulative transcriptional effect level (determined as shown in Figure 3.1A) for all the genes in the stress genes library used for a toxicogenomic assay. Only selected genes are shown in X-axis.

Three-dimensional real-time gene expression profiles, depicted as 3-D topography, are obtained for four nanomaterials (NMs), including nano-silver (nAg), nano-titanium dioxide anatase (nTiO\(_2\) _a_), nano-titanium dioxide rutile (nTiO\(_2\) _r_), and fullerene and they are shown in Figure 3.2. The 3-D altered gene expression profiles are distinctive for the four NMs evaluated, suggesting compound-specific cellular responses likely resulted from their different MOAs. TELI translates the transcriptional level effect of an organism in response to a toxicant into a quantitative endpoint value. TELI value is determined for each toxicogenomic-based toxicity assay upon NMs exposure and its value allows for quantitative comparison of transcriptional-level effect induced by various NMs at different concentrations.
Figure 3.2-3-D altered gene expression profile compiled from the altered gene expression level for all the genes in the test library. The TELI value was determined as the sum of accumulative transcriptional effect level (determined as shown in Figure 3.1A) for all the genes in the stress genes library used for a toxicogenomic assay.
Currently used ecotoxicogenomics endpoint, namely NOTEL, is quantified based on gene expression data at one chosen time point (Ankley et al., 2006). It has been recognized that the cellular responses, measured as gene expression profiling, are dynamic over time (Ma, Castillo-Davis, Zhong, & Liu, 2006; Ronen, Rosenberg, Shraiman, & Alon, 2002). Although NOTEL has been applied and shown to be a useful indicator for quantifying the transcriptional effect level induced by a toxicant, it is evident that the NOTEL value does not reflect other important dimensions of the toxicogenomic data such as the magnitude of differentially expressed gene levels for specific genes and the variation in response profile depending on the exposure time. The proposed TELI can incorporate all these factors and therefore is expected to quantify the transcriptional effect at higher resolution and with better accuracy.

It is important to note that our toxicogenomic-based toxicity assays record the transcriptional level response of the E.coli cells upon the exposure of various NMIs at the concentrations much lower than that would lead to any noticeable inhibition effect (the maximum concentration examined caused less than 5% growth inhibition compared to control, data not shown). It reflects more subtle cellular response at doses far below those used in conventional toxicity assay that can lead to observable phenotype change such as growth inhibition or death. Therefore, it is conceivable that the TELI values likely entail more potential long-term impact of a toxicant on the cell, although the TELI can be obtained from short-term assay within 2hrs.
Figure 3.3 Dose response curves based on TELI versus dose concentration for nAg, nTiO2_a, nTiO2_r, and fullerene, fitted by Four-parameter Logistic Equation. Data points with error bar represent the TELI value determined at each dose concentration. (A) TELI versus does concentrations for the NMs studied. TELI_{MAX}, the maximum value of TELI for each assay, is determined through model fitting. (B) Transformed dose-response curve based on probability unit. TELI values are normalized to the TELI_{MAX}. TELI50, NOTEL_{TELI}, and Slope_{TELI}
are determined via a linear fitting of *probit unit* versus dose concentration. TEL$_{\text{MAX}}$. TEL$_{50}$ and NOTEL$_{\text{TEL}_1}$ of nTiO$_2$-a are shown in figure.

The TELI values exhibit concentration-dependent pattern and allows for the establishment of dose-response curve, as shown in Figure 3.2 and Figure 3.3. The TELI-dose response curve are analyzed with Four-parameter Logistic Equation (4PL) model and they exhibit a characteristic “sigmoid” shaped toxicity dose-response curve, indicating that TELI is a suitable response quantifying parameter.

In Figure 3.2A, the best fitting curves did not pass theoretical zero points for the NM$_\text{s}$ tested. The reason is likely that NM$_\text{s}$ alone may had background physical interference of the GFP reading, which can not be quantified separately, and it gives a background reading slightly higher than theoretical zero.

![Figure 3.4 Transcriptional effect Level Index (TELI) values determined for toxicity assay for nAg, nTiO$_2$-a, nTiO$_2$-r, and fullerene at various dose concentrations with exposure time of 2 hrs. Y-axis: TELI value. Standard errors are shown as error bar.](chart.png)
In traditional toxicity assays, the phenotype endpoints (e.g. growth inhibition, death) are clearly defined and toxic response is normalized as percentage of the maximal effect (e.g. 100% death). With toxicogenomic tests, we observe that there seemed to be a maximal TELI value that can be determined based on the TELI-dose response curve using 4PL model fitting and, it varies for different toxicant (Figure 3.2 (A)). It is reasonable to think that there may be an threshold condition when the transcriptional level effect, including both direct and compensatory effects here in our study, is at its maximal and, beyond which cell damages start to occur and eventually progresses to observable phenotype damage endpoint such as growth inhibition or even death. This threshold can be indicated by the maximal value of TELI (TELI\text{MAX}). The TELI\text{MAX} values are determined as 83.44(±17.14), 53.95(±2.95), 20.15(±7.12), and 39.76(±116.2) for nAg, nTiO\textsubscript{2-a}, nTiO\textsubscript{2-r}, and fullerene respectively, shown in Figure 3.2(A). The TELI\text{MAX}, referred as the maximal efficacy, reflects the limit of the dose-response relationship on the response axis to a certain chemical. It quantifies the maximal transcriptional level effect for a group of given gene that can be induced in 2 hour exposure.

Based on traditional toxicology, in which quantitative toxicity endpoints such as EC50, LC50 are determined as the concentration that leads to 50% of the maximal adverse effect, we apply the similar approach for determining TELI\text{50}. TELI values are normalized to TELI\text{MAX} determined as described above and, the percent TELI response can be exhibited as *probit units* (*probability unit*) as the dose-response phenomena are usually normally distributed(Curtis D. Klaassen, 2010).
TELI50 is found via a linear fitting of probit unit verse dose, as shown in Figure 3.2(B), when probit unit equals to 5 (corresponds to 50% TELI_{MAX} response). We also calculate TELI-based NOTEL, namely NOTEL_{TELI}, as the concentration at which probit unit equals to 3.355 (corresponds to 5% TELI_{MAX} response). Slope of fitted line (TELIslope) is another endpoint since it reflects the rate at which the transcriptional effect increases with the increasing of concentration. The four endpoints defined, namely, TELI_{max}, TELI50, NOTEL_{TELI} and TELIslope, together describes the characteristics of the transcriptional level response as a function of dose for a given NM.

**Table 3.3 Summary of TELI-based toxicity characterization parameters and endpoints**

<table>
<thead>
<tr>
<th></th>
<th>TELI_{MAX} (mg/L)</th>
<th>Slope_{TELI}</th>
<th>TELI50 (mg/L)</th>
<th>NOTEL_{TELI} (mg/L)</th>
<th>NOTEL (mg/L)</th>
<th>EC50 (mg/L)</th>
<th>BOD (umol/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>nAg</td>
<td>83.44 (±17.14)</td>
<td>1.19 (±0.34)</td>
<td>18.92 (±1.63)</td>
<td>0.79 (±2.51)</td>
<td>1.08 (±1.23)</td>
<td>67.19 (±9.48)</td>
<td>116.4</td>
</tr>
<tr>
<td>nTiO2_a</td>
<td>53.95 (±2.95)</td>
<td>2.19 (±0.87)</td>
<td>7.02 (±1.30)</td>
<td>1.01 (±2.51)</td>
<td>1.01 (±1.32)</td>
<td>74.14 (±8.13)</td>
<td>64.9</td>
</tr>
<tr>
<td>nTiO2_r</td>
<td>20.15 (±7.12)</td>
<td>1.96 (±1.04)</td>
<td>6.23 (±5.19)</td>
<td>1.11 (±6.31)</td>
<td>1.17 (±2.45)</td>
<td>177.91 (±19.87)</td>
<td>12.6</td>
</tr>
<tr>
<td>Fullerene</td>
<td>39.76 (±6.46)</td>
<td>1.33 (±0.33)</td>
<td>15.93 (±2.52)</td>
<td>0.92 (±2.52)</td>
<td>1.06 (±1.37)</td>
<td>87.22 (±12.31)</td>
<td>62.1</td>
</tr>
</tbody>
</table>

* Numbers in parenthesis are the standard error.
Table 3.4 Correlation coefficients between TELI-based toxicity endpoints with other toxicity endpoints

<table>
<thead>
<tr>
<th>Correlation Coefficient</th>
<th>TELIMAX</th>
<th>Slope_{TELI}</th>
<th>TELI50</th>
<th>NOTEL_{TELI}</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOTEL</td>
<td>-0.51</td>
<td>-0.04</td>
<td>0.20</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>(-0.4)</td>
<td>(-0.4)</td>
<td>(-0.2)</td>
<td>(0.2)</td>
</tr>
<tr>
<td>EC50</td>
<td>-0.82</td>
<td>0.39</td>
<td>-0.61</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>(-1.0)</td>
<td>(0.4)</td>
<td>(-0.8)</td>
<td>(0.8)</td>
</tr>
<tr>
<td>BOD</td>
<td>0.98</td>
<td>-0.63</td>
<td>0.80</td>
<td>-0.95</td>
</tr>
<tr>
<td></td>
<td>(1.0)</td>
<td>(-0.4)</td>
<td>(0.8)</td>
<td>(-0.8)</td>
</tr>
</tbody>
</table>

* The values outside the parenthesis are Pearson product-moment correlation coefficients, the values inside the parenthesis are Spearman’s rank-order correlation coefficients.

Table 3.3 summarizes the TELIMAX, TELI50, NOTEL_{TELI} and TELIslope values determined based on the TELI-dose response relationships for nAg, nTiO2-a, nTiO2-r, and fullerene, as shown in Figure 3.2. TELIMax show the highest level for nAg and lowest level for nTiO2-r, indicating relatively higher transcriptional level effect induced by nAg than other three NMs to these stress genes. TELI50 and NOTEL_{TELI} can be referred as “potency”, which refers to the dose required to produce a given response intensity. The NOTEL_{TELI} values for the four NMs do not show significant difference (p>0.28), indicating similar level of dose threshold to produce detectable transcriptional effect among the NMs studied, or NOTEL_{TELI} may not be sensitive enough to differentiate the toxic effects of these NMs. The difference in the dose response curve TELIslope suggests that, for these four NMs, the transcriptional effects are induced at different increasing rates with increase in dose.

For comparison of relative toxicity among toxicants based on transcriptional effect levels, employment of multiple endpoints is necessary to describe the dose response behavior and characteristics more completely and holistically. For
example, except NOTEL-TELI, TELI50 and TELIslope values for TiO$_2$$_{r}$ and TiO$_2$$_{a}$ are comparable, however, the dose-response curves in Figure 3.2 clearly showed distinctive toxic response behaviors, which can be reflected by the TELIMAX. This is consistent with traditional toxicology that more than one endpoint of dose-response curve should be considered to indicate the characteristic of the toxic response over the exposure of a toxicant (Anno, Young, Bloom, & Mercier, 2003). The denser arrangement of atoms and higher stability of TiO$_2$$_{r}$ may explain why apparent lower toxicity of TiO$_2$$_{r}$ compared to TiO$_2$$_{a}$ have been observed (at transcriptional effect level), especially in nanosize (Falck et al., 2009; Warheit, Webb, Reed, Frerichs, & Sayes, 2007). This is further confirmed by comparing their EC50 values with much higher EC50 of TiO$_2$$_{r}$ than that of TiO$_2$$_{a}$ (Table 3.3).

3.4.2 The correlation between TELI with other established toxicity endpoints

Another historical impediment to the use of toxicogenomic information for regulatory decision making has been the lack of consented approach and inability to link responses at lower levels of biological organization to adverse outcomes in the whole animal, so called “phenotypic anchoring” (Ankley et al., 2006). To validate our proposed toxicity assessment index-TELI, we conduct conventional toxicity assay to determine the EC50 and NOTEL for all the NMs and then compared the TELI endpoints with these conventional toxicity endpoints. Table 3.3 compares the TELI50 values with EC50, NOTEL and BOD values for the NMs studied. BOD quantifies the oxidative damage potential, measured by a “Ferric reducing ability of serum (FRAS)” assay in human blood serum as the
difference of total antioxidant capacity between unexposed and NM exposed serum, and was proposed as indicator for quantifying oxidative damage-based nanotoxicity in eukaryotic cells (Bello, 2009). BOD for each NMs is measured with 10 mg NM/ml serum at 37°C after 90 mins mixing and represented as trolox equivalent unit (Bello, 2009). Table 3.4 shows the correlations coefficients between the TELI-based toxicity endpoints with other accepted toxicity endpoints, statistic significance testing for correlation coefficient are not done here as the sample size if too small to be reliable for statistics t-test to judge that correlation is exist or not (Dunlop, 2000).

The values of NOTEL_{TELI} are consistent with the NOTEL values as shown in Table 3.3, and their statistical significance testing p-values for difference range from 0.96~1 for all the NMs tested. Since both NOTEL_{TELI} and NOTEL represent the same toxin level at which no affect can be observed on transcriptional level, the agreement of NOTEL_{TELI} and NOTEL values is expected and it validates the concept and application of proposed NOTEL_{TELI} based on TELI. The results suggested that NOTEL and NOTEL_{TELI} are not sensitive enough for differentiating the toxicity level among NMs in this study, however, the resolution may potentially be improved by extending the exposure time or enlarging the gene library (Lobenhofer et al., 2004).

The TELI_{MAX} values among the NMs seem to be correlated relatively well with the EC50 (r=-0.82). EC50 represents the phenotype changes such as cell growth being inhibited, which can only be observed at higher dose concentrations than those applied in our TELI assays (Poch & Pancheva, 1995). As discussed
previously, $\text{TELI}_{\max}$ is assumed to indicate the threshold beyond which the phenotype damage may occur, which is, therefore, very close to the concentration that may cause phenotype damage. This correlation implies that “phenotype anchoring” is possible with our proposed $\text{TELI}_{\max}$, therefore allowing for the linking of phenotype changes with cellular response.

The relative toxicity order of the four NMs based on values of $\text{TELI}_{\max}$ also agreed well with that suggested by the BOD values, as demonstrated by the high correlation coefficient of $\text{TELI}_{\max}$ with BOD value ($r = 0.98$ with $p=0.023$), respectively. This is consistent with the current understanding that oxidative damage, as resulted by the observed oxidative stress and Reactive Oxidative Species (ROS) production by these NMs, has been found to be the dominant toxic mechanism for these NMs (Hussain et al., 2009; Xia et al., 2006).

The toxicity order among the NMs revealed by the TELI50 seemed to be different from that indicated by the EC50 and BOD. This is because the maximum transcriptional effect ($\text{TELI}_{\text{max}}$ value) varies with different toxicant, whereas, the maximal phenotype effect (such as death) in conventional toxicity tests is the same for all chemicals. Therefore, TELI50 here cannot be employed for direct comparison of relative “potency” or toxicity level among chemicals as EC50 implies. The TELI_{slope}, determined as the shape of dose-response curve, is important in defining the shape of TELI-dose response curve for indicating the rate of relative percentage change in the response intensity (in relative to the maximum TELI) per unit change in dose. For example, comparing to nTiO$_2$, nAg and fullerene exhibited relatively less steep dose-response curves, indicating
slower relative increase in transcriptional effect while dose concentration increases. These result implies that the comparative toxic effect at transcriptional level, exhibited at the concentration much below those that cause phenotype response and damage, may not be evaluated the same way as that at phenotype response level.

Therapeutic index (TI), usually used in pharmacology, exhibited the safety information of substances. Therapeutic index is a comparison of the amount of a therapeutic agent that causes the therapeutic effect to the amount that causes drug toxicity. In traditional medical toxicology, TI is determined as the toxic effect of 50% of the population (TD50) divided by the minimum effective dose for 50% population (ED50)(Curtis D.Klaassen, 2010). We explore the possibility to apply the concept of TI to evaluate the ratio of dose that causes phenotype damage to that cause transcriptional effect. In this case, we define Pheno-to-Transcriptional damage Index (PTI) as the EC50 of growth inhibition divided by TELI50 to indicate the potential effect of NMs. The PTI for nAg, nTiO$_2$-a, nTiO$_2$-r, and fullerene are 3.92, 10.56, 28.41, and 5.46 respectively. In similar analogy, lower PTI for nAg and fullerene suggest that the dose that leads to phenotype damage (growth inhibition) is not much higher than that cause transcriptional effect without phenotype damage. PTI can potentially serve as another index for characterizing toxicity of environmental pollutants.

3.4.3 TELI for effects associated with different stress functional categories

The 91 stress genes in *E.coli* used in our library have been categorized into several functional groups based on their main function and involvement in
different stress mechanism, such as DNA stress (SOS response), protein stress, membrane damage (lipid stress) and oxidative stress (including ROS degradation and redox balancing) (show in table 3.2). Based on the definition and principle of TELI determination, it is reasonable to calculate TELI values for genes in each stress response category to quantify the corresponding transcriptional level response for each stress category. TELI indicates transcriptional effect level based on relative temporal altered gene expression levels for the selected genes in the assay, assuming equal weight for all genes. The TELI of each functional group were calculated and shown in Figure 3.4 (A).

Figure 3.5 Comparison of TELI values for different stress gene functional categories among the four NMs studies. (A) Transcriptional effect level indicator (TELI) determined for different functional stress gene categories, for
toxicity assays of four NMs at two representative concentrations. (B) TELI values determined for further division of DNA-stress related genes indicative of different DNA damage and repair pathways upon the exposure to nAg and nTiO\textsubscript{2-a}, respectively. (The number 10 and 50 indicate the exposure concentration, 10mg/L and 50mg/L)

The lowest TELI of nTiO\textsubscript{2-r} of all the stress categories indicates the lowest toxicity among the NMs examined, as discussed in earlier sections. nAg has the highest TELI for genes involved in oxidative stress, indicating main toxicity mechanism as oxidative damage, which is consistent with previous reports(AshaRani, Mun, Hande, & Valiyaveettil, 2009; Carlson et al., 2008; Hsin et al., 2008). Oxidative damage induced those genes related to ROS degradation and redox balancing. The key regulator oxyR, which serves as the transcriptional dual regulator to those involved in peroxide metabolism, peroxide protection and redox balance(Keseler et al., 2009), was up-regulated at an increasing level over time upon exposure to nAg (data not shown). Both nAg and nTiO\textsubscript{2-a} exposures led to observable DNA stress and membrane Damage. It is understood that oxidative stress and ROS production can lead to damage of DNA, RNA, proteins, and lipids(Apel & Hirt, 2004).

Not only TELI values can be determined for genes belong to different functional categories, indicating MOA-related comparative transcriptional level toxicity among compounds as shown in Figure 3.4(A), TELI values can also be obtained for more in-depth pathways to reveal toxicity details. Figure 3.4(B) shows the TELI values specific for those genes that are indicative of specific DNA damage
and repair pathways. Note that the genes that are common for many DNA damage and repair pathways, such as DNA polymerase (polB, dinB) and helicase (dinG), are not included. As shown in Figure 3.4(B), although both nAg and nTiO$_2$-a seem to cause DNA stress, they lead to different DNA damage type and repair pathways. In prokaryotic cells, DNA damage leads to activation of SOS DNA damage and repair regulation system, which is regulated by two SOS regulator, namely *recA* and *lexA*. The activation of SOS system, is a strong function of ssDNA length, for a minimum site size of 30-50 bases(Brenner et al., 1987).

Some DNA damage type may not produce many ssDNA, such as base damage(Jac A. Nickoloff, 1998). Comparing the two NMs that cause DNA stress, including nAg and nTiO$_2$-a, nAg did not seem to activate the SOS system, however, nTiO$_2$-a led to clear activation of SOS (as show in Figure 3.4(B)).

### 3.5 Conclusion

This study proposed and demonstrated the potential application of a new To Transcriptional Effect Level index (TELI) to convert the information-rich toxicogenomic data into an integrated and quantitative endpoint. The TELI values exhibit concentration-dependent pattern and allows for the establishment of dose-response curve. Multiple transcriptional effect level based endpoints, describing the holistic toxic response characteristics, can therefore be determined. The validity of the toxicity assessment index, TELI, was demonstrated through comparison to other conventional or previously reported endpoints. Continuous advances in genomic technologies and in toxicogenomics field will pose it as a
promising and advantageous approach for mechanistic toxicity evaluation of pollutants. Methodologies for converting the rich toxicogenomic information into readily usable and transferable format that can be potentially link to regulation endpoints and incorporated into decision-making framework is in great need and our proposes TELI may serve as such an effective index. Although we applied gene expression data in this study, the principle and concept of TELI can be potentially employed by proteomics, metabolomics and others that examine global response of cell to toxicants.
Chapter 4

Mechanistic Toxicity Assessment and Classification of Environmental Contaminants of Emerging Concern Via Stress Response Ensemble Based Toxicogenomics Assay

4.1 Abstract

Rapid development of toxicogenomics technologies has created new approaches to screen large number of emergency contaminants for mechanistic toxicity assessment. However, challenges remain in the analysis and utilization of the resulting high-dimensional data. Because of the lack of commonly accepted quantization methods, it is difficult to link the toxicogenomics information to conventional phenotypic endpoints and further incorporate into environmental assessment and regulation framework. We applied a new quantitative toxicogenomics-based toxicity assessment method for mechanistic toxicity evaluation and classification of 30 contaminants of emergency concern (CECs,) covering various group, such as pharmaceuticals, antibiotics, drinking water disinfection byproducts, food, industrial additive, pesticide, herbicide, nanomaterials etc.. Their major toxicity mechanism are identified via different approach and compared with previous studies. Molecular responses expressed by Transcriptional Effect Level Index (TELI) exhibited clear dose-response patterns and allows for multiply endpoints derivation. The derivate endpoints are linked to conventional toxicity endpoint EC50. The chemical classification using consensus
4.2 Introduction

Contaminants of emerging concern (CECs) can be broadly defined as any synthetic or naturally occurring chemical that is not regulated or commonly monitored in the environment but has the potential to enter the environment and cause adverse ecological or human health impacts (Mira et al., 2008). The prominent classes of CECs include pharmaceuticals and personal care products (PPCPs), surfactants, plasticizers, pesticides, flame-retardants, nanomaterials, as well as their transformation products (TPs) (Poynton et al., 2009; Mastroianni et al., 2010; Schwarzenbach et al., 2006). Recent studies have detected CECs in surface waters of rivers, streams, and lakes as well as in ground and drinking water (Trine et al., 2010; Murray et al., 2010; Buxton et al., 2007; Sedlak et al., 2005; Snyder et al., 2008). The growing evidence and rising public concern of harmful environmental and health effects and risks posed by these chemicals demand a paradigm shift in toxicological assessment approach to provide timely and sufficient information needed for risk management (Snyder et al., 2004; Kidd et al., 2004; Aardema et al., 2002; Krewski et al., 2010). Currently, most of the CECs have not been adequately tested for their potential impacts to human and wildlife due to the labor and resources intensiveness of conventional animal-based toxicity assessment as well as the complexity and challenges in toxicity characterization of CECs and their transformation products (Muir and Howard, 2006). There are urgent needs for developing of toxicity mechanism and
pathways-based testing scheme that incorporates cost-effective, reliable high-throughput screening and tired testing to compliment and prioritize the conventional resource-intensive and lengthy whole animal-based testing, as National research council (NRC) and USEPA proposed (Bhattacharya et al., 2011; NRC, 2007). Then in time we hope to eventually build predictive models and tools to provide toxicological information that transcends the limits of data generation (Krewski et al., 2010) Error! Reference source not found.

The advance in “-omics” technologies, which allows for high-throughput and concurrent monitoring of cellular status and response pathways globally upon exposure to chemical toxicants, can be used to expand the coverage of new and existing chemicals that need to be evaluated for human health risk assessment (Krewski et al., 2010; NRS, 2007; Neale et al., 2012; Escher et al., 2013; Macova et al., 2011; Escher et al., 2011; Yeh et al., 2014; Waters and Fostel, 2004; Ankley et al., 2006). When applied to the study of large classes of chemicals or drugs, toxicogenomic information can be used to globally define modes or mechanisms of toxic action (Pennie et al., 2000; Kennedy, 2002; Armitage et al., 2014; Nicholson et al., 2002). Compare to conventional toxicity test, toxicogenomics assay has the advantages of targeting primarily cellular responses, revealing toxicity mechanisms, broader range of test doses, high throughput and quick turn-around, significantly lower cost and no or minimal animals use. Efforts that incorporate in vitro assays into toxicity screening and evaluation, environmental monitoring and risk assessment are in progress among various countries (e.g. USEPA TOXCAST, EU REACH, EU OPENTOX,). These
initiatives and many other studies have employed either targeted or batteries of \textit{in vitro} assays among variety of species to provide timely and comprehensive bioactivity signatures for a larger number of toxic chemicals. In addition, using the adverse outcome pathway (AOP) concept, combined with system biology and bioinformatics, they have successfully demonstrated the potential for establishing \textit{in vitro} and \textit{in silico} data-based predictive toxicity models or relationships among \textit{phenotypic or in vivo} toxicity endpoints.

Given that it is not practical to design an cost-effective in vitro high throughput assays that can interrogate every potential toxicological target and pathway, an integrative approach that rely on the monitoring of indicative key gene (biomarker) modules in limited and rather conserved pathways (for all cell types of metazoans) linked to adverse outcome seems promising (Ankley et al., 2006). The highly conserved cellular stress response systems are involved in detecting stress and repairing damage caused by toxicants on primary cellular infrastructure and macromolecules, therefore offer a variable solution to reduce targets for comprehensive and efficient toxicity evaluation (Simmons, et al., 2009). As the cellular stress responses target on primary cellular infrastructure, they are highly conversed in all metazoan cells, which makes them much predictable cross species. The small set of pathways can help to increase throughput rate for fast screening. Since the stress response pathways are aimed to restore cellular homeostasis and participate in cell fate/death decisions, it is directly related to the phenotype changes and adverse outcome (Kultz, 2005; Simmons, et al., 2009), which makes them potential adverse outcome pathways (AOP) and help to link
the direct molecular initiating events to a sequential series of effects from cells to organ, as well as adverse outcome at organism and population level (Ankley et al., 2010; Kramer et al., 2011).

In this study, we applied a newly developed quantitative toxicogenomics method based on cellular stress response pathways biomarkers ensemble, which enabled fast, cost-effective, yet informative and mechanistic toxicity evaluation and classification of 30 CECs (Gou et al., 2010; Gou and Gu, 2011; Onnis-Hayden et al., 2009; Safe, 1990; ). Comprehensive toxicity evaluation of large number and different classes of CECs using toxicogenomics approach has never been reported. Compared with the conventional whole animal based toxicity assays or other single or battery of in vitro assays, this quantitative toxicogenomics approach in combination with bioinformatics computation detects the overall toxicity level, reveals potential toxicity profiles and mechanisms with one single assay within hours (2 hrs), which greatly reduces the test time, resource and cost (Gao et al., 2012). In addition, the high-resolution molecular toxicity profiles can serve as fingerprints for chemicals clustering analysis and classification. The results provided an initial database of comprehensive molecular toxicity information for a large number and diverse categories of CECs, which can assist future QSAR and predictive toxicity model development. The toxicity results not only confirmed known toxicity mechanism for some CECs, but also revealed new toxicological insights and information for many other CECs. The proposed method may serve as an alternative or complementary approach to the current
toxicity assays for environmental pollutants toxicity screening, evaluations and classification.

4.3 Materials and Method

4.3.1 Chemical selection and evaluation

30 CECs were selected based on their prevalence in wastewater and drinking water, as well as their available toxicological information in the literature and they include pharmaceuticals, industrial and food additive, plasticizers, pesticides, and nanomaterials (Appendix 4.1). Stock solutions were prepared in ethanol or dimethyl sulphoxide (DMSO) and stored at proper temperature as required in the dark. Deionized water (18.0 mΩ·cm, Millipore Milli-Q system) is used in all the experiments. 8 well studied chemicals with known modes of action, namely arsenic, chromium, erythromycin, mitomycin C, hydrogen peroxide, trihaloacetic acid and metabolites of benzo[a]pyrene (with mouse S9) were selected as model compounds for toxicity results comparison as references. Dose concentrations applied in the liquid media and calculated cell/tissue dose exposure concentrations based on dosimetry and mass balance according to Armitage et al. (Armitage et al., 2014) were summarized in Stable 3. Details of the mass balance models are described in Appendix.

4.3.2 Toxicogenomics assay

Assay description

A high-throughput mechanistic toxicity assay method was employed with GFP-fused whole-cell array of E.coli K12, MG1655. Those tested chemicals were diluted with minimum media (M9) to various desired concentrations. The selected
assay library covers a variety of genes involved in different known cellular stress response pathways, such as oxidative stress, DNA stress, protein stress, and membrane stress (including membrane transporter, efflux pump, energy metabolism, flagella metabolism, lipopolysaccharide metabolism), etc. (see biomarkers in Table 3.2). Each fusion is expressed from a low-copy plasmid, pUA66, which contains a kanamycin resistance gene and a fast folding gfpmut2, allowing for real-time measurement of gene expression level changes.

Assay protocol

The protocol to measure the temporal gene expression profile is described in our previous reports. In brief, E.coli was cultivated in 384-well plates (Costar, Bethesda, MD, USA) in dark to avoid GFP photobleaching until the early exponential growth stage is reached (OD600 ~ 0.2). After addition of the aqueous sample, the plate was placed into a micro-plate reader (Synergy Multi-Mode, Biotek, Winooski, VT) for cell growth measurement (absorbance, 600 nm) and fluorescent reading (Excitation: 485 nm, Emission: 528 nm) every 5 minutes over a period of 2 hours. Three biological replicates were performed for each condition. Test and control groups are arranged on the same plate for each replicate.

Data processing

The GFP and OD data were first smoothed using 5-time moving average and then corrected against background (from growth medium only blanks) with and without chemicals exposure, respectively. The population normalized GFP signal was calculated as $P = \frac{GFP}{OD}$ and corrected against background (from E.coli strains without GFP infusion with same OD) with and without chemicals.
exposure, respectively. Induction factor, which measures the gene expression alteration, was calculated as the ratio of normalized expression levels between experimental (with chemical exposure) and control groups (without chemical exposure), \( I = \frac{P_e}{P_c} \), where \( P_e \) and \( P_c \) represents for experimental and control groups, respectively. The induction factor \( I \), determined for a given gene at each time point upon chemical exposure, was then used for the following data analysis, where a gene is up-regulated if \( I > 1 \) and down-regulated if \( I < 1 \). The toxicity mechanisms were elucidated by identifying and analyzing the changes in genes associated with specific stress response pathways.

4.3.3 Dose-response and toxicogenomics endpoints derivation

The molecular response was quantified by the TELI (Transcriptional Effect Level Index) value, which was a recently developed index for toxicogenomics data interpretation that incorporates the number, magnitude and the temporal pattern of genes with an altered expression. TELI can be integrated as equation 1 on various levels, as single gene (TELI\textsubscript{gene}), pathway (TELI\textsubscript{pathway}), stress category (TELI\textsubscript{oxidative}, TELI\textsubscript{geno}, TELI\textsubscript{membrane}, TELI\textsubscript{protein}, and TELI\textsubscript{general}), and overall stress library (TELI\textsubscript{total}). Compared to previous publication, improvement was made to TELI computation. In this paper, TELI is calculated as the average of transformed \( I \) instead of sum of transformed \( I \) to make TELI value more comparable among different pathway, stress category, and overall library.

\[
\text{TELI} = \frac{\sum_{i=1}^{n} \left( \frac{\ln(I)}{e^{ln(I)}} \right)_{t=0}^{t=2hr}}{\text{ExposureTime}^* n}
\]

Equation 1
Where, \( t \) was the exposure time, \( n \) was the number of genes in one particular pathway/stress category/total stress library, \( w_i \) was the weight factor of gene \( i \). For this study, we assigned value of 1 for all the weighing factors.

The TELI-dose response curves were analysis with a 4-paramater logistic equation model using GraphPad 5.0c (Prism) (Prism, 2007). Several toxicoenomics endpoints were derived from TELI-dose response curves (Gou and Gu, 2011).

4.3.4 Stress gene set enrichment analyses (Stress-GSEA)

The stress gene set enrichment analyses (Stress-GSEA) are achieved by ranking a list of genes with the TELI values. For each pathway, GSEA calculates the enrichment score by examining the ranked gene list from high score end to the low score end, giving a rewarding score if a gene belongs to the pathway of interest, and penalizing score otherwise. The significance of each pathway is determined by comparing their ranking scores to the corresponding empirical distributions. The null distributions are generated by randomly permuting the specific pathway and all others 1,000 times. For details, please refer to the work of Subramanian (Subramanian et al., 2005). Only those categories with \( p \)-value less than 0.05 are considered to be altered significantly.

4.3.5 Gene ontology analyses

Gene ontology analysis was performed with Cytoscape software using BiNGO plug-in for enrichment analysis to determine the overrepresented GO biological process annotations with a hypergeometric statistical test and a 5% Benjamini & Hochberg false discovery rate correction (Maere et al., 2005). In this study the
whole stress library was used as the reference set, and activated genes were used as test set, which is defined $\text{T}E\text{L}I_{\text{gene}} > 1.5$. 1.5 was selected since, first, it is widely used for gene overexpression determination, in addition, it is close but larger than the mean+2SD (while mean=1, SD=0.28 for $\text{T}E\text{L}I_{\text{gene}}$) in this study.

4.3.6 Clustering analysis

Consensus clustering is a robust clustering method, which iteratively resamples the sample and performs clustering on the reconstructed dataset. It finds out how the clustered treatments are significantly close to each other using the frequency of pair of treatments clustered together in spite of resampling as the similarity between the pair. This method provides higher confidence level over the resulted cluster. Cluster stability in consensus clustering is measured with consensus index (CI), which is calculated from the frequency of two samples are clustered together.

$$Consensus\ Index, CI(i,j) = \frac{\sum_h \text{M}_{(h)}(i,j)}{\sum_h \text{I}_{(h)}(i,j)} \cdots \cdots \cdots \cdots \cdots (1)$$

CI$(i,j)$ is an indicator of similarity between treatments and ranged between 0 and 1. If the two toxicants always clustered together, CI value is 1 and when they are not clustered at any time CI value will be 0. When sample $i$ and $j$ clustered together $\text{M}_{(h)}(i,j)$ is 1 and $\text{M}_{(h)}(i,j)$ is 0 otherwise. On the other hand, $\text{I}_{(h)}(i,j)$ is defined to be 1 if the $i^{th}$ and $j^{th}$ treatments both appear in $h^{th}$ resampling. Clustering algorithm Self-organizing Map (SOM) is used to perform the iterative clustering. The samples are resampled following bootstrapping method. We use 1000 iteration to find the overall consensus indices.
Finally dendrogram is generated based on average linkage hierarchical clustering with distance function, $D(i,j) = 1 - CI(i,j)$. Consensus clustering is conducted on the full data set which consists of 25 time points of 107 genes. The dimensions of the data set are 2450. Logarithm of induction factor, $\ln(I)$ of different time points of each gene are used as the input variable of SOM. Each time point of each of the gene is treated as independent and has similar weightage on the distance measurement while applying SOM clustering algorithm.

4.4 Results and discussion

4.4.1 Toxicogenomics assay reveals Toxicity mechanisms of CECs

Chemical-specific and dose-dependent temporal altered stress response genes expression profiles were obtained for 30 CECs with 6 different concentrations. The results showed very dynamic and complex toxicant-induced real-time gene expression across the stress response pathways and genes examined, with most of the genes exhibiting different patterns and varying magnitudes of transcriptional activities over time. The information-rich toxicogenomics results provided detailed examination of genes involved in various stress responses pathways and functions, therefore suggesting varying underlying molecular toxicity mechanisms among them. The 107 stress genes in our *E.coli* assay library have been categorized into 5 functional groups based on their main function and involvement in different stress mechanisms, such as DNA stress (SOS response), oxidative stress (including ROS defense and redox balancing), protein stress, membrane (lipid stress and membrane transporter), and general stress (other stress...
as pH, temperature, et al.) (Table 3.2 list of genes). Both GO analysis and GSEA analysis were performed to identify the pathways or functional groups that were significantly activated or altered (as seen in Table 4.1). The specific toxicity mechanism associated stress response pathways that were significantly altered were identified as the potential mode of action for the specific toxicant, and the results are summarized in Table 4.1 and Figure 4.1 for comparison with the known toxicity mechanisms for the CECs evaluated. The stress response categories that have been activated was identified based on TELI values (TELICategory >1.5, p<0.05), the dominant cellular stress response was determined based on GSEA analysis and the key biomarkers indicative of specific stress response pathways was selected based on significant fold changes (TELIgene >2, p<0.05). Note that GSEA analysis identified the most significantly altered categories in relative to others, while, TELI values identified all stress response categories that showed significant altered expression level compared to control. IN addition, GO analysis provided more detailed information on specific biological processes that were over expressed upon chemical exposure (Stables 3). The comprehensive toxicity characterization of these chemical not only confirmed their known toxicity MOA but also discovered more toxicological insights (Table 4.1). Detailed discussed of identified toxicity effects for each class of CECs are presented in the following sections.
Figure 4.1 Toxicity profile based on the quantitative endpoints of different stress response categories of tested chemicals at various concentrations. X-axis top: list of stress response categories. Stress response in relative to the untreated were indicated as the TELI value for each stress category and color-coded with the scale 1-5 (Red spectrum colors indicate the magnitude of the altered gene expression, while black indicates no change). Y-axis left: concentration in unit of uM.
Figure 4.2 Toxicty profiles at selected concentration for tested chemicals to represent the variation among different categories. X-axis bottom: list of stress response categories. Stress responses in relative to the untreated were indicated as the TELI value for each stress category.
Table 4.1 Summary of toxicity mechanism (MOA) from literature and our toxicogenomics assay

<table>
<thead>
<tr>
<th>Class</th>
<th>Chemical</th>
<th>Known toxicological information</th>
<th>Identified stress response</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heavy metal</td>
<td>Arsenic (V)</td>
<td>Disrupts ATP production Form reactive oxygen species Bind to protein, induce unfolded protein response</td>
<td>Protein Membrane General</td>
<td>Consistent finding on protein stress, and oxidative stress. New insight as membrane stress and general stress were also identified.</td>
</tr>
<tr>
<td>Chromium (VI)</td>
<td></td>
<td>Induce DNA damage, gene mutation, sister chromatid exchange, chromosomal aberrations, cell transformation and dominant lethal mutations in a number of targets Form oxygen free radical</td>
<td>DNA Membrane General DNA</td>
<td>Consistent finding on DNA stress. New insight as membrane stress and general stress were also identified.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Oxidative stress Affect intra and intercellular signaling, cell adhesion, protein folding and maturation, apoptosis, ionic transportation, enzyme regulation, release of neurotransmitters</td>
<td>Oxidative Oxidative</td>
<td>Consistent finding on oxidative stress</td>
</tr>
<tr>
<td>Lead</td>
<td></td>
<td>Inhibit of DNA gyrase Form free-radical and cause oxidative stress</td>
<td>General General</td>
<td>Consistent finding on general stress with genes involved in cell growth regulation. New insight on more border range of cellular toxicity effects may indicate global stress response</td>
</tr>
<tr>
<td>Antibiotics</td>
<td>Ciprofloxacin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>Compound</td>
<td>MOA</td>
<td>Pathway(s)</td>
<td>MOA Pathway</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Pharmaceuticals</td>
<td>Erythromycin</td>
<td>Inhibit protein biosynthesis</td>
<td>DNA</td>
<td>cyoA, grpE, lon, yedW, dacA, cspB, dinJ, nfo, ssb, uvrD</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Protein stress not identified as previous report. GSEA identified DNA stress</td>
</tr>
<tr>
<td></td>
<td>Mitomycin C</td>
<td>DNA crosslinker</td>
<td>DNA</td>
<td>soxS, ompC, zntA, gadX, recX, mutH, dnaQ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Membrane DNA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Consistent finding on DNA stress. Membrane stress on transporter of ion may been affected.</td>
</tr>
<tr>
<td></td>
<td>Atovastatin</td>
<td>Inhibit hydroxymethylglutaryl-coenzyme A reductase (HMG) and led to subsequently decrease of (LDL)s in serum.</td>
<td>DNA, General</td>
<td>sodA, rpoD, emrA, pbpG, otsB, cspB, ydgL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No known MOA related pathway in E.coli</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Affect the transporter function of the cell membrane</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Metoprolol</td>
<td>β1 receptor blocker</td>
<td>General</td>
<td>katE, rpoD, zntA, cueR, yeeV, ydgL,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No known MOA related pathway in E.coli</td>
</tr>
<tr>
<td>byproduct</td>
<td></td>
<td>Oxidative stress</td>
<td></td>
<td>Consistent finding in DNA stress and oxidative stress Effects on cell growth may be potentials MOA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Consistent finding in DNA stress and oxidative stress. Effects on cell membrane components/transporter and</td>
</tr>
<tr>
<td>Substance</td>
<td>Effect (EC50)</td>
<td>MOA</td>
<td>Gene Products</td>
<td>Finding</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>---------------</td>
<td>-----</td>
<td>-------------------------------------------------------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Trihalomethane mix</td>
<td>Exhibit carcinogenic</td>
<td>DNA</td>
<td>yebG, polB, umuD</td>
<td>cell growth may be potentials MOAs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DNA</td>
<td>yebG, polB, umuD</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DNA</td>
<td>yebG, polB, umuD</td>
<td></td>
</tr>
<tr>
<td>Nitrosamine mix</td>
<td>Exhibit mutagenic and genotoxic</td>
<td>Oxidative</td>
<td>gst, norR, fpr ycgE ompC, dacA, sbmA dinJ, rpoE, otsB nfo, sulA, polA</td>
<td>Consistent finding in DNA stress and oxidative stress. Effects on cell membrane components/transporter and cell growth may be potentials MOAs.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DNA</td>
<td>yebG, polB, umuD</td>
<td></td>
</tr>
<tr>
<td>Industrial and food additive</td>
<td>Exhibit hormone receptor activity Induce rapid signaling effects Oxidative stress</td>
<td>Oxidative</td>
<td>inaA, ahpC, tam amiC, emrE cspB, slyA,</td>
<td>No specific toxicity receptor in E.coli system Consistent finding in oxidative stress.</td>
</tr>
<tr>
<td>Bisphenol A</td>
<td></td>
<td>DNA</td>
<td>yebG, polB, umuD</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DNA</td>
<td>yebG, polB, umuD</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DNA</td>
<td>yebG, polB, umuD</td>
<td></td>
</tr>
<tr>
<td>Surcolase</td>
<td>No observed ecological effects</td>
<td>Oxidative</td>
<td>katE, sodB, tam fepB bacA, fsr uspA, otsB, bolA</td>
<td>No specific toxicity receptor in E.coli system Consistent finding in oxidative stress.</td>
</tr>
<tr>
<td>4-n-Nonylphenol</td>
<td>Endocrine disruptor</td>
<td>Oxidative</td>
<td>Fpr, tam, grpE, clpB, rpoD zntA, dacB, marC cspA, slyA, otsB yebG, sbmC, mutT</td>
<td>No specific toxicity receptor in E.coli system Consistent finding in oxidative stress. Protein stress may be potential MOA.</td>
</tr>
<tr>
<td>Bis(2-ethylhexyl) phthalate</td>
<td>Bind and activate estrogen receptors</td>
<td>Oxidative</td>
<td>katG, oxyR, sodA lon, grpE, yedW fepB, osmE,</td>
<td>No specific toxicity receptor in E.coli system Consistent finding in oxidative stress.</td>
</tr>
<tr>
<td>Compound</td>
<td>Effect</td>
<td>Location(s)</td>
<td>Genes</td>
<td>Toxicity Receptor</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>-------------</td>
<td>-----------------------------------------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Caffeine</td>
<td>Inhibit cyclic nucleotide phosphodiesterase isozymes, Blocks adenosine receptors, Inhibit phosphodiesterase, Greater expression of NO</td>
<td>DNA</td>
<td>General</td>
<td>No specific toxicity receptor in E.coli system</td>
</tr>
<tr>
<td>Herbicide and insecticide</td>
<td>Inhibit fatty acid synthesis</td>
<td>Oxidative</td>
<td>soxS, soxR, oxyR, rpoD, emrA, marA, emrE, cspB, otsB</td>
<td>Consistent finding in oxidative and membrane stress</td>
</tr>
<tr>
<td></td>
<td>Inhibit (reversible) acetylcholinesterase</td>
<td>General</td>
<td>grpE, zntA, yhjX, cmr, cspB, uspA, otsB</td>
<td>No specific toxicity receptor in E.coli system</td>
</tr>
<tr>
<td></td>
<td>Induce protein adduct</td>
<td>General</td>
<td>katE, sodB, tam, entC, dnaJ, dacB, yhjX, marC, slyA, uspA, bolA</td>
<td>No specific toxicity receptor in E.coli system</td>
</tr>
<tr>
<td>Trichloroacetic acid</td>
<td>Hypomethylation of DNA, Lipid peroxidation (TBARS) and oxidative DNA damage</td>
<td>Oxidative</td>
<td>soxS, dps, cyoA, bacA, yhjX, inaA, uspB, gadX, recA, recX, nfo</td>
<td>Consistent finding in DNA stress</td>
</tr>
<tr>
<td>17-beta-</td>
<td>Endocrine disruptor</td>
<td>DNA</td>
<td>katE, dps,</td>
<td></td>
</tr>
<tr>
<td>Compound</td>
<td>Effect</td>
<td>Location</td>
<td>Genes</td>
<td>Consistency</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>----------------</td>
<td>--------------------------------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>Estradiol</td>
<td>Interaction with specific intracellular receptors Damage to the DNA</td>
<td>Membrane</td>
<td>ycgE, motA, zntA, uspB, otsB, ydgL, mutH</td>
<td>Consistent finding in DNA stress</td>
</tr>
<tr>
<td>Phenol liquor</td>
<td>binding to protein thiols and DNA and also to redox cycling</td>
<td>Membrane DNA</td>
<td>cueR, marR, cls, ompC, Fpr, yeiG, sbmC, polB, mutT</td>
<td>Consistent finding in DNA stress</td>
</tr>
<tr>
<td>Nanomaterials</td>
<td>nAg, nTiO2_a, nTiO2_r</td>
<td>Oxidative DNA</td>
<td>Consistent finding in oxidative stress</td>
<td>Consistent finding in oxidative stress</td>
</tr>
</tbody>
</table>
Heavy metals

Protein stress was identified as the major toxicity response upon arsenic exposure by both GSEA and TELI values analysis, while GO: 42493 Protein unfolding, GO:51789 Response to protein stimulus were identified as significantly overrepresented GO terms. Gene lon (degrade mis-folded proteins), dnaJ (assist in protein folding processes), rpoD (promoter the translation of protein synthesis-related genes) were found to be up-regulated upon arsenic exposure. This is also consistent with previous studies. Ramadan et al. found that As(III) tends to bind protein and further compromise protein folding pathway (Ramadan et al., 2008).

Oxidative stress, which is believed to be the major MOAs responsible for arsenic carcinogenesis (Kitchin et al., 2003; Kitchin et al., 2008), were also identified by TELI values analysis, and key oxidative stress related genes soxR (Superoxide Response protein), sodC (superoxide dismutase), and grxA (reduced glutaredoxin) were found to be up-regulated upon arsenic exposure. However, new insight was also found since membrane stress and general stress were also found to be over-represented with corresponding key biomarkers found to be up-regulated.

DNA damage was revealed as the major cellular effects by chromium exposure by all of the analysis. Cr(VI) is known to cause DNA damages such as strand breaks, DNA–DNA and DNA–protein cross-links nucleotides modification (Liu et al., 2001; Slade et al., 2005). Similar to As(III), membrane stress and general stress were also found to be over-represented with corresponding key biomarkers found to be up-regulated.
Both our assay and previous report suggested that, generation of oxidative free radicals is one of the recognized common toxicity MOA for lead (Ercal et al., 2001; Pande et al., 2002), as evidenced by the alterations in the genes involved in oxidative stress, katG (hydrogen peroxide oxidoreductase), ahpC and ahpF (alkyl hydrogenperoxide resuctase), soxS and soxR (transcriptional activator for superoxide radicals) (Table 3.2).

**Antibiotics**

Ciprofloxacin, which affect DNA gyrases and consequently inhibit DNA translation, and cell division (Goossens et al., 2007), led to significantly overrepresented of general stress by both GSEA and TELI analysis, also led to the up-regulations of genes involved in translation inhibition and cell growth inhibition such as ratA(inhibition of cell growth) and relB (regulate cell growth with limited nutrients). The alteration in many stress pathways caused by ciprofloxacin also evidenced a more boarder range of cellular toxicity effects and global stress response by ciprofloxacin (British Columbia Center for Disease Control, 2010; Drlica et al., 1997; Pommier et al., 2010).

There is no identified category for erythromycin by TELI values since none of the examined categories show great response upon erythromycin exposure. (Figure 4.1, Table 4.1) The identifiable of DNA stress by GSEA may imply that using GSEA only may not be appropriate for MOA identification from our toxicogenomics approach. Previous studies found that erythromycin can led to the inhibition of protein biosynthesis (Chambers, 1998; Brisson-Noël et al., 1988;
Leonard and Ashley, 2005; Gaynor et al., 2003). Protein stress is not identified in our study, however, 2 genes related with protein folding, assembly and maturation, grpE (assist in protein folding processes) and lon (degrade mis-folded proteins), showed up-regulation upon erythromycin exposure.

DNA stress was identified as the major MOA for mitomycin C, which is a potent DNA crosslinker in previous report (Tomasz et al., 1987). DNA crosslinks and other damages introduced by mitomycin C are removed by recombinational repair, which was evidenced by up-regulation of several double strand break repair genes recX, mutH, recA, and recN (Bernstein et al., 2012). TELI analysis and up-regulation of gene ompC (trimeric porin for ions trans-membrane) and zntA (zinc, cadmium and lead efflux system) suggested that cell membrane transporter system may also been affected by mitomycin C exposure.

**Pharmaceuticals**

The 3 pharmaceuticals are all functioned via the specific pathways, which are not present in *E.coli* system. Atorvastatin works as inhibitor of hydroxymethylglutaryl-coenzyme A reductase (HMG) (Malinowski, 1998), ibuprofen inhibits the enzyme cyclooxygenase (Mitchell et al., 1993), and metoprolol is a selective β1 receptor blocker (Bengtsson, 1976). They all have general stress identified as the major stress response from GSEA analysis but not identified via TELI value approach, which may indicates the low toxicity of these 3 pharmaceuticals and the impropriation of using either method along for MOA identification. However, significantly induced membrane stress by GSEA analysis
and up-regulation of gene marR (involved in multiple antibiotic resistance), marC (an inner membrane protein to form porin) upon ibuprofen exposure are in consistence with the effect on the transporter function as influx and efflux of the cell membrane by ibuprofen (Parepally et al., 2006).

**Water disinfection byproducts**

Our results showed consistent finding on DNA stress and oxidative stress for the 4 disinfection byproduct (DBPs). NDMA, nitrosamine mixture, trihalomethane mixture, and bromodichloromethane, all have been reported to exhibit carcinogenic activity as well as oxidative stress (Pereira et al., 1982; Coffin et al. 2000). TELI values analysis as well as up-regulated genes also identified DNA stress and oxidative stress as the major MOA for all the DBPs. However, general stress was also identified upon NDMA trihalomethane mixture, and bromodichloromethane, while membrane stress was identified upon the exposure of bromodichloromethane and trihalomethane mixture. Protein stress was identified for nitrosamine mixture. Similar results were not found in literatures. Our toxicogenomics results provided new insight on the MOAs of DBP, effects on cell membrane components/transporter and cell growth may be potentials MOA for DBPs. Much fewer identified categories by GSEA may be due to the severe stress caused by these DBPs overshadowed their specific responses.

**Food and industrial additive**

Similar with pharmaceuticals, 4 out of 5 of the food and industrial additive are have receptor binding/inhibiting effects, which cannot be captured by our
toxicogenomics approach with *E.coli*. Di(2-ethylhexyl) phthlate (DEHP) can bind and activate oestrogen receptors (ER) (Issemann et al., 1993). Caffeine act as an antagonist at adenosine receptors (Fredholm, 1995). Bisphenol A is well known of its nuclear hormone receptor activity (Takeshita et al., 2001). 4-Nonylphenol (4-NP) has been shown to have an estrogenic activity on synthetic promoters carrying an estrogenic, responsive gene (Angelique and Bailhache, 2006).

Oxidative stress is identified as overrepresented MOA for almost all the chemical within this class, with the only exception, caffeine, consistent with previous studies (Chitra et al., 2003; Kasahara et al., 2002; Di Gregorio et al., 2015). General stress is identified as overrepresented MOA for almost all the chemical within this class, only with the exception of DEHP. 4NNP has protein stress identified as the major stress by both GSEA and TELI values analysis. 3 genes involved in folding and assembly, dnaK, dnaJ, and grpE, are all up-regulated upon 4NNP exposure. No similar conclusion has been found in previous studies.

**Herbicide and insecticide**

2 of 3 chemicals within this class have very consistent results between GSEA analysis and TELI values approach. Oxidative, membrane and general stress are identified as the major MOA upon triclosan exposure. Triclosan is seen to target bacteria mainly by inhibiting fatty acid synthesis and induce the overexpression of multidrug efflux pump (Fan et al., 2002; Heath et al., 1999; Mcmurry et al., 1998). Since fatty acids are necessary for reproducing and building cell membranes, and our membrane stress contains a lot of membrane transporter and efflux pump. It is not surprising that membrane stress is identified as the major response for
triclosan in our toxicogenomics results. Transcriptional activator of superoxide radicals gene soxR is up-regulated in literature as well as our results via toxicogenomics approach (Mcmurry et al., 1998).

Oxidative, protein and general stress are identified as the major MOA upon atrazine exposure. Atrazine (6-chloro-N-ethyl-N’-(1-methylethyl)-1,3,5-triazine-2,4-diamine) is known to induced protein adducts with the formation of atrazine-protein supermolecular by Van der Waals force and hydrogen bonds (Dooley et al., 2007). Up-regulation of 2 antioxidant enzyme coded gene sodB and katE were shown in our results and previous study (Jin et al., 2010).

Carbaryl (1-naphthyl methylcarbamate) is a reversible inhibitor of the enzyme acetylcholinesterase (AChE) (Forsberg and Gertrud, 1984). Similar enzyme does not exist in E.coli system. Only General stress is identified via GSEA analysis.

**Others**

The chemicals in “Others” group includes 3 reference compound, H\textsubscript{2}O\textsubscript{2} for oxidative damage (Ballinger et al., 2000; Halliwell et al., 1994), benzo[a]pyrene metabolites for DNA damage, they both clearly exhibit their known MOA, as indicated by all the analysis approach, GSEA, TELI GO and activated genes analysis.

Trichloroacetic acid, used for the precipitation of macromolecules, showed consistent overrepresented DNA stress with previous study that hypomethylation of DNA is one of the major MOA (Tao et al. 2000).
DNA stress is over expressed for phenol liquor exposure via TELI values approach, and phenol is believed to be toxic to cells by binding to protein and DNA (Hooman and Wright, 2008). Oxidative and general stresses were also found to be overrepresented upon trichloroacetic acid exposure.

17-beta-estradiol is a predominant estrogen during reproductive years, also named as E2. GSEA showed general and membrane stress as the major response for 17β-estradiol, but TELI values approach showed no over expressed category.

**Nanomaterials**

The toxic mechanisms of metal and metal oxidant NMs are still not fully understood and current understanding indicates that the prominent toxic mode of action involves production of Reactive Oxygen Species (ROS) (Oberdörster et al., 2004; Wiesner et al., 2006) which can damage DNA, RNA and proteins, including a multitude of oxidized base lesions, abasic sites, single and double-strand breaks, all of these can be cytotoxic and mutagenic (Diakowska et al., 2007). Our results showed that all 4 nanomaterials induced oxidative stress, detailed MOA can be found in Chapter 3.

For most of the examined toxicants, the toxicity information is limited on human health relevant studies with little or no information on eco-relevant studies. 16 out of the 30 tested chemicals have consistent MOA results between toxicogenomics assay and literature. Our results confirmed the toxicity mechanism and provide more detailed information on cellular response upon the exposure of these chemicals. Then, for the other chemicals, toxicity information generated in this
study have not been mentioned in previous publication, which may guild to further toxicity assessment of these chemicals.

The differences among 4 different MOA identification methods indicate that, using any method along may lead to incompletely conclusion. GSEA identified the most significantly expressed category without consider the absolute expression value, while TELI values approach give out all the over expressed categories with no emphasized on the major one. GO and activated genes provide too detailed information which are challenges for regulation purposes. Combine all these 4 methods for chemical MOA identification, can provide more comprehensive and precise information, especially helpful for environmental assessment and regulation.

4.4.2 Molecular Toxicity Endpoints Derivation and Comparison among CECs

In addition to elucidate the toxicity mechanisms of the CECs by identifying and analyzing the changes in genes associated with specific stress response pathways, we further derived quantitative molecular toxicity endpoints for these CECs from dose response modeling. To derive quantitative endpoints, a dose-response curve is required and it displays the central paradigm in toxicology, which is “the dose makes the poison” (Fawcett et al., 1996). However, quantification of molecular response and development of dose-response curve at molecular level are rare in literature for toxicogenomics data (Burgoon et al., 2008). In our previous work, we have demonstrated that molecular perturbation can be quantified by the TELI
(Transcriptional Effect Level Index) value, which is a recently developed index for toxicogenomics data interpretation that incorporates the number, magnitude and the temporal pattern of genes with an altered expression (Gou and Gu, 2011). TELI values can be determined for a single gene, or for a number of genes representing a specific pathway or the entire stress response assays. Here, our results for the first time, showed that cellular stress response pathway ensemble-based toxicogenomics data derived TEL values exhibited dose-response curves at molecular level for a large and diverse group of CECs.

The TELI values exhibit dose dependent pattern for various toxicants for both overall cellular stress response that cover all known pathways (Figure 4.2), demonstrating that TELI is a suitable response quantification parameter (Gou and Gu, 2011). Note the exposure concentrations used in the dose response modeling are based on dosimetry calculation that considered chemical partition and mass balances, which represent more accurate intracellular concentration.

The TELI dose response curves are analyzed with a 4-parameter Logistic equation (4PL) model, which theoretically exhibit a characteristic “sigmoid” shaped toxicity dose response curve. The 4PL nonlinear regression model is commonly used for curve-fitting analysis in bioassays dose-response curves. For chemicals that are known to have very low toxicity or considered safe to human health and commonly used as over counter-painkiller and food additive, like ibuprofen, metoprolol, bisphenol A, sucralose, and two relatively inert nanomaterials (nTiO2_r and fullerene soot) showed rather flat dose response curves, indicating low toxicity as expected. In contrast, for other chemicals that are recognized to
present high toxicity to *E. coli* such as arsenic, ciprofloxacin, all the drinking water byproduct, and 4NNP induced significantly cellular stress response with the typical “S” or partially “S” shaped dose response curve. For some chemical categories, such as the heavy metals, and the four drinking water disinfection byproducts (DBPs), they seemed to show parallel does response curve, suggesting similar MOA based on the relative potency concept for mixture risk assessment.

With toxicogenomic tests, we observe that there seemed to be a maximal TELI value that can be determined based on the TELI-dose response curve using 4PL model fitting, and it varies for different toxicants (Figure 4.2). It is recognized that cellular stress responses assist in maintaining homeostasis; however, they are also toxicity pathways in that they lead to adverse effects when stress is sufficiently high. It is therefore, reasonable to think that there may be a threshold condition when the transcriptional level effect, including both direct and compensatory effects here in our study, is at its maximal and beyond which cell damages start to occur and eventually progresses to observable phenotype damage endpoint such as growth inhibition or even death. This threshold corresponds to and can be indicated by the maximal value of TELI (TELIMAX).
**Figure 4.3** Dose-response curves based on TELI values versus dose concentration for tested chemicals, fitted by the Four-parameter Logistic Equation. Data points with an error bar represent the TELI value determined at each dose concentration. Chemicals are classified into different group based on their certain function.

The TELIMAX, referred as the maximal efficacy, reflects the limit of the dose_response relationship on the response axis to a certain chemical. It quantifies the maximal transcriptional level effect for a group of given gene that can be induced in a 2 h exposure. For all the chemicals examined, 4 out of 30 have TELIMAX values lower than 2 (ibuprofen, surcolase, nTiO2_rutile, fullerene soot), with nTiO2_r TELIMAX values even lower than 1.5, which is a commonly used cutoff line for gene expression level evaluation. Gene expression level lower than 1.5 is usually considered as no expression. 5 out of 30 have TELIMAX values higher than 4 (mitomycin C, trihalomethane mix, nitrosamine mix, ciprofloxacin, and arsenic) with ciprofloxacin TELIMAX values as high as 11.17, which is a really high value in gene expression tests. Most of the commonly considered low toxic chemicals have their TELIMAX values lower than or close to 2.5 (bisphenol A, carbaryl, atorvaseatin, caffeine, nTiO2_a, and metoprolol, and atrazine). The others with TELIMAX between 2.5 and 4 can be seen as relative moderate toxic. TELI50 was determined as the similar approach of EC50, LC50 (effect concentration, leath concentration) as the concentration that leads to 50% of the maximal adverse effect. We also applied 2 point of departure (POD) approach, namely NOTEL_{TEL1}, as the concentration at which evoked 5% TELIMAX response,
and EC-TELI1.5 which reflect the concentration that causes TELI value to reach 1.5, similar to the approach that has been applied on *umuC* genotoxicity assay by Escher *et al.* TELI50, NOTELELI, and EC-TELI1.5 can be referred to as “potency”, which refers to the dose required to produce a given response intensity. For the EC-TELI1.5 for nTiO2-r is not applicable since its TELI_MAX value is lower than 1.5. For most situation, EC-TELI1.5 is higher than NOTELELI since TELI_MAX value is lower than 10. We consider EC-TELI1.5 as a more reasonable POD approach, since 1.5 is a commonly accepted criteria for gene expression evaluation.

Slope of fitted line (SlopeTELI) is another endpoint since it reflects the rate at which the transcriptional effect increases with the increasing of concentration. The difference in the dose response curve SlopeTELI suggests that, for different chemicals, the transcriptional effects are induced at different increasing rates with an increase in dose. Most of SlopeTELI are lower than 1, which indicates the importance of using the 4PL model instead of 3PL model with a fixed slope (usually equal to 1).

Table 4.2 shows the summary of toxicogenomics endpoints for all the examined chemicals. For comparison of relative toxicity among toxicants based on transcriptional effect levels, employment of multiple endpoints is necessary to describe the dose response behavior and characteristics more completely and holistically.
Table 4.2 Summary of TELI-dose response curve derived toxicogenomics endpoints.

<table>
<thead>
<tr>
<th>Group</th>
<th>Chemical</th>
<th>TELIMAX (mg/L)</th>
<th>TELI_{50} (mg/L)</th>
<th>NOTEL_T_{ELI} (mg/L)</th>
<th>TELI_{1.5} (mg/L)</th>
<th>Slope_{TELI}</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Heavy metal</strong></td>
<td>Arsenic</td>
<td>4.003±0.15</td>
<td>1.19x10^{-5}</td>
<td>1.17x10^{-8}</td>
<td>2.70x10^{-7}</td>
<td>0.4245±0.147</td>
</tr>
<tr>
<td></td>
<td>Chromium (VI)</td>
<td>3.276±1.075</td>
<td>0.0202</td>
<td>3.92x10^{-10}</td>
<td>9.68x10^{-6}</td>
<td>0.1658±0.157</td>
</tr>
<tr>
<td></td>
<td>Lead</td>
<td>3.860±0.8585</td>
<td>0.03686</td>
<td>5.09x10^{-10}</td>
<td>2.65x10^{-6}</td>
<td>0.1627±0.057</td>
</tr>
<tr>
<td><strong>Antibiotics</strong></td>
<td>Ciprofloxacin</td>
<td>11.17</td>
<td>0.018</td>
<td>5.68x10^{-10}</td>
<td>5.12x10^{-10}</td>
<td>0.1705±0.126</td>
</tr>
<tr>
<td></td>
<td>Erythromycin</td>
<td>3.624</td>
<td>891422</td>
<td>2.56x10^{-18}</td>
<td>2.42x10^{-6}</td>
<td>0.05432</td>
</tr>
<tr>
<td></td>
<td>Mitomycin C</td>
<td>5.826</td>
<td>123.3</td>
<td>1.82x10^{-12}</td>
<td>9.02x10^{-9}</td>
<td>0.09246</td>
</tr>
<tr>
<td><strong>Pharmaceuticals</strong></td>
<td>Atovastatin</td>
<td>2.239±0.1896</td>
<td>0.006</td>
<td>1.05x10^{-6}</td>
<td>0.002</td>
<td>0.3380±0.3033</td>
</tr>
<tr>
<td></td>
<td>Ibuprofen</td>
<td>1.937±0.0379</td>
<td>0.00016</td>
<td>3.79x10^{-6}</td>
<td>0.00019</td>
<td>0.7779</td>
</tr>
<tr>
<td></td>
<td>Metoprolol</td>
<td>2.043±0.056</td>
<td>0.00082</td>
<td>0.00057</td>
<td>0.0008</td>
<td>7.988</td>
</tr>
<tr>
<td><strong>Drinking water byproduct</strong></td>
<td>N-Nitrosodimethy lamine</td>
<td>3.542±0.5699</td>
<td>0.016</td>
<td>9.38x10^{-9}</td>
<td>1.68x10^{-5}</td>
<td>0.2051±0.098</td>
</tr>
<tr>
<td></td>
<td>Bromodichlorom ethane</td>
<td>2.916±0.7597</td>
<td>0.023</td>
<td>4.62x10^{-7}</td>
<td>0.0005</td>
<td>0.2727±0.2694</td>
</tr>
<tr>
<td></td>
<td>Trihalomethane mix</td>
<td>5.921</td>
<td>0.3929</td>
<td>8.53x10^{-16}</td>
<td>5.50x10^{-12}</td>
<td>0.08721</td>
</tr>
<tr>
<td></td>
<td>Nitrosamine mix</td>
<td>4.038±1.023</td>
<td>0.0019</td>
<td>1.667x10^{-10}</td>
<td>2.46x10^{-7}</td>
<td>0.1809±0.09</td>
</tr>
<tr>
<td><strong>Industrial and food additive</strong></td>
<td>Bisphenol A</td>
<td>2.310</td>
<td>0.4742</td>
<td>3.40x10^{-13}</td>
<td>0.0048</td>
<td>0.1053</td>
</tr>
<tr>
<td></td>
<td>Surcolase</td>
<td>1.990±0.1367</td>
<td>0.2042</td>
<td>7.967x10^{-4}</td>
<td>0.2121</td>
<td>0.5309</td>
</tr>
<tr>
<td></td>
<td>4-n-Nonylphenol</td>
<td>3.288±0.8679</td>
<td>0.016</td>
<td>6.329x10^{-3}</td>
<td>0.0015</td>
<td>0.5309</td>
</tr>
<tr>
<td></td>
<td>Bis(2-ethylhexyl) phthalate</td>
<td>2.548±0.2227</td>
<td>0.034</td>
<td>2.52x10^{-5}</td>
<td>0.0055</td>
<td>0.4083</td>
</tr>
<tr>
<td></td>
<td>Caffeine</td>
<td>2.332±0.3499</td>
<td>0.0147</td>
<td>4.00x10^{-9}</td>
<td>0.0011</td>
<td>0.1947</td>
</tr>
<tr>
<td><strong>Herbicide and insecticide</strong></td>
<td>5-CHLORO-2-(2,4- DICHLOROPHENOXY)PHENOL</td>
<td>2.651±0.3265</td>
<td>0.011</td>
<td>1.86x10^{-6}</td>
<td>0.0009</td>
<td>0.3388±0.2113</td>
</tr>
<tr>
<td></td>
<td>Carbaryl</td>
<td>2.566</td>
<td>0.0505</td>
<td>1.73x10^{-14}</td>
<td>3.15x10^{-5}</td>
<td>0.1026</td>
</tr>
<tr>
<td></td>
<td>2-chloro-4- (ethylamino)-6- (isopropylamin o)-s-triazine</td>
<td>2.268±0.8689</td>
<td>2.83x10^{-6}</td>
<td>1.73x10^{-14}</td>
<td>1.79x10^{-7}</td>
<td>0.1557</td>
</tr>
<tr>
<td><strong>Others</strong></td>
<td>Hydrogen peroxide</td>
<td>3.235</td>
<td>8.243</td>
<td>0.027</td>
<td>0.739</td>
<td>0.5160</td>
</tr>
</tbody>
</table>
In order to apply toxicogenomics assay for water quality monitoring and toxicity assessment, quantitative endpoints are crucial to integrate the toxicogenomics data into environmental evaluation and regulation framework. However, the developed endpoints from toxicogenomics studies, as TELI, NOTEL, are still in exploration, very few has been accepted and widely used yet because how these molecular endpoints correlate or link to phenotypic endpoints, so called phenotypic anchoring is an active research area under investigation.

Therefore, investigate the anchoring between molecular response and phenotypic response is the next urgent step, but it is beyond the scope of this study. We will further evaluate this anchoring with the increase of toxicological data for a large number of CECs.

For comparison of relative toxicity among toxicants based on transcriptional effect levels, employment of multiple endpoints is necessary to describe the dose response behavior and characteristics more completely and holistically. For example, the benchmark dose, TEL1.5 values for nitrosamine mixture and atrazine are similar with consideration of the variance of the data (Table 4.1), the dose-response curves for these two compounds are quite different as shown in
Figure 4.2, which can only be accurately described by using the multiple endpoints. This is in consistent with traditional toxicology that more than one endpoint of a dose-response curve should be considered to indicate the characteristic of the toxic response over exposure of a toxicant.

Table 4.3 Correlation coefficient between TELI-based toxicogenomics endpoints with conventional toxicity endpoint.

<table>
<thead>
<tr>
<th>Correlation coefficient (r)</th>
<th>TELI_{MAX}</th>
<th>TELI_{50} (mg/L)</th>
<th>NOTEL_{TELI} (mg/L)</th>
<th>TELI_{1.5}</th>
<th>Slope_{TELI}</th>
</tr>
</thead>
<tbody>
<tr>
<td>E.coli 24 hours EC50</td>
<td>0.77</td>
<td>0.01</td>
<td>0.51</td>
<td>0.80</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>(0.00005)</td>
<td>(0.96568)</td>
<td>(0.01818)</td>
<td>(0.00001)</td>
<td>(0.00054)</td>
</tr>
</tbody>
</table>

* The values shown are person product-moment correlation coefficient with log10 of EC50 (M), the values inside the parentheses are significance test p-value.

An additional challenge to apply toxicogenomics on toxicity evaluation and risk assessment has been the lack of consented approach and inability to link responses at molecular level to adverse outcomes at system level. To probe the linkage between our proposed toxicogenomics endpoints with apical endpoints, we quantitatively evaluated the correlation of the gene expression result expressed by TELI from our cellular stress response ensemble toxicogenomics approach with conventional apical endpoints, EC50, based on growth inhibition on *E.coli*, Table 4.3 shows the correlation coefficient between the TELI-based endpoints with the conventional EC50. Except TELI_{50}, all the other 4 endpoints seem to correlate well with EC50 (r>0.50, with p-value <0.05). Among the 4 endpoints, TELI_{1.5} has the strongest correlation with EC50 (r=0.8), while the other benchmark dose NOTEL_{TELI} show much weaker correlation with EC50 (r=0.51).
This may indicate that TELI\textsubscript{1.5} is a better endpoints for linking to phenotypic response.

4.4.3 Consensus Cluster

The information rich and high-resolution toxicity effect fingerprints obtained from a toxicogenomics assay enables classification of samples based on the similarity in their toxicity profiles and thus underlying mechanisms, therefore facilitates the selection of reference chemical for various specific toxicity outcome at finer categorization resolution. More expensive investigations are needed to test against large number and variety of toxicants to confirm these hypotheses.

Clustering methods that consider temporal dimension of the data is rare. The most frequently used methods, as clustering algorithms such as hierarchical clustering (HCl) and self-organizing maps (SOM), is either sensitive to outliers, or sensitive to the random process of map initialization, making comparisons between studies challenging. Current validation schemes include internal approaches, which are based solely on analysis within a data set might not be suitable for biological data which are subject to high noise levels.

Here we explored chemical classification using consensus clustering (CC) approach based on self-organizing maps (SOM) (Afshari et al., 2010; Jiang et al., 2004), in which all the dimensions of the toxicogenomics profiles data were considered including genes with altered, expression, exposure time, and magnitude of gene expression changes. CC serves both as a validation algorithm
for conventional clustering and as an ensemble clustering approach, so it is seen to be applied to cluster the biological data so that it can produce reliable results.

The clustering results using the CC approach are illustrated as consensus matrix as shown in Figure 4.2. Based on the CI value consensus clustering on the original dataset identifies 4 significant clusters, labeled in black and red colors.

The treatments of low toxic chemicals (carbaryl, bisphenol A, ibuprofen, atorvastatin, sucralose, caffeine, etc.) concentrations aggregated closely. For these chemicals, responses do not differentiate largely on their concentration, treatments with different concentrations tend to cluster closely together, indicating that there is conserved similarity in the response patterns for a given chemical at different concentration level. These chemicals include some low toxic chemicals, such as ibuprofen, sucralose, carbaryl, 17-beta estradiol, metoprolol, etc. and some higher toxic chemicals, such as DEHP, erythromycin, H₂O₂, mitomycin, lead etc.. Some chemicals, the treatment of different concentration split into different clusters, may indicating their dose-dependent toxicity mechanism, for example, arsenic, chromium, NDMA, phenol liquor, 4NNP, etc.. This cluster may because that at higher concentrations, the very specific toxicity mechanism transit into a broader global stress response, which lead to cytotoxicity.

The resulting cluster patterns are mostly consistent with prior toxicological knowledge. Chemical specific transcriptional level effects are pronounced. Those with moderate toxic chemicals are cluster together, like triclosan, nitrosamine mix, mitomycin C, trichloracetic acid and lead. And some low to moderate toxic chemicals are not differentiate with low toxic chemicals, such as
triclosan, atrazine, and nitrosamine mix. Noisy and redundant information in the high dimensional data might be a possible reason behind this behavior. In this clustering analysis all the genes and their time points are considered to be of equal weights on separation of the chemicals. No effect of the time series pattern is included since we considered all the dimensions as the features of clusters.

Overall, the results demonstrated that the stress-response pathways ensemble based real time toxicogenomics assay yield chemical-specific and concentration-sensitive transcriptomic profiles. Moreover, statistically reliable clustering such as CC is capable of identifying classes of chemicals according to their underlying toxic mechanisms.
**Figure 4.4** Consensus clustering matrix of tested chemicals with different concentrations. The consensus matrix represented as a heatmap. A dendrogram is based on the CIs, showing the inner structure in the blocks along the main diagonal. The identified clusters are labeled in black or red color in the dendrogram. The numbers indicate the concentration of chemicals, 1 to 6 are from low to high. Those chemicals without number afterwards mean all 6 concentrations are cluster together.

### 4.5 Conclusion

Advanced in high-throughput *in vitro* mechanistic toxicity assay address the challenges in need of timely and informative toxicological information for the large number and ever increasing CECs, here we applied a toxicogeomics assay,
which is fast, feasible, low-cost, make it possible to provide comprehensive
toxicity assessment of a relatively large number of CEC of various categories with
limited resources. The study initiated a database of comparable molecular
toxicological data for CECs, which can be continually enlarged. The information
rich toxicogenomics results provided detailed molecular response information and
relevant screening data that reflect multiple potential toxicity pathway from a
limited number of test for these CECs, thus optimizes resources and limit animal
use.

We for the first time, applied quantitative approach to characterize the molecular
toxicity of these CECs on stress response, and established does-response
relationship. The widely existed and conserve does-dependent response for
molecular toxicity response with stress response ensemble is crucial for further
development and application of toxicogenomics assay, since only if it can drive
quantitative endpoints, it is possible to link to phenotypic and risk assessment
endpoints. The shape of does response at molecular level share high similarity as
phenotypic dose response, and conventional modeling approach may be still
applicable, this allows derivation of molecular endpoints such as TELI_{max},
TLEI1.5, et al.

Finally, the high-resolution toxcigenomics data provide fingerprint of each
chemical (for any given species), and our results demonstrate the possibility to
allow for chemical classification and identification.
Chapter 5

Mixture Toxicity Prediction via Quantitative Toxicogenomics-based Approach

5.1 Abstract

Recent years, toxicogenomic approaches provide tools to improve our understanding and predictability of the mixture toxicity. However, current evidence to support the published mixture assessments from toxicogenomic studies is as yet mainly observational and needs further improvement. In this study, we demonstrated the application of a toxicogenomic-based approach for investigating the combined effects of various binary mixtures with a range from environmental relevant concentration to the benchmark level.

Five binary mixtures were investigated at various concentrations with a fixed ratio (w/w). Cellular level responses (indicated by our quantitative omic-index TELI) evoked by most of the binary mixtures exposure exhibited clear dose-response patterns as single chemicals did. The TELI values were studied using the 4-parameter logistic regression model for further analysis.

To quantitatively analyze the combined effects of mixtures, two additive models were applied, concentration addition (CA) and independent action (IA). Parallel dose-response curves were found among As, Cr and the mixture, which indicate similar mode of action (MOA), and the experimental observed mixture toxicity fit well to both CA and IA model predication (r²>0.8). Significantly higher
Experimental toxicity than both model predications by the mixture of H2O2-sucralose indicated possible synergistic effect. And significantly lower experimental toxicity than both model predications was found by mitomycin C-sucralose indicating possible antagonistic effect.

Our results not only illustrated a rapid, sensitive, and informative approach for mixture toxicity studies, but also for the first time demonstrated the possibility to explicitly predict the combined effects at the molecular level. However, further investigation is required for better understanding of the factors that influence the combined effects.

5.2 Introduction

Chemicals analysis provides ample evidence that large numbers of emerging contaminants are present in the aquatic environment and pose threat to various organisms (Schwarzenbach et al., 2006; Sedlak et al., 2005). Thus, greats challenge exist in addressing the water quality problems associated with such mixture exposures in terms of understanding their harmful impact and risk, and developing effective prediction methods (Krewski et al., 2010; NRC, 2007).

Currently the resulted overall toxicity from mixture of chemicals is often difficult, if possible at all, to predict based on chemical analysis and known toxicity information for the identified individual chemicals, considering the complex mixture interactions such as synergistic or antagonistic effects. Risk assessment focusing on single or detectable chemicals can only be justified if one or a number of chemicals dominate the overall toxicity, and their effects are additive (with
similar MOAs) (Scenihr, 2012; European Commission, 2009; Wilson and Schwarzman, 2011). However, chemicals rarely occur alone in environmental water samples, overall toxicity may raise even though the concentrations of individual chemicals are below detection limit or certain toxicity threshold.

Current predictive modeling of ecological mixture toxicity is adopted from pharmacology and is based on two different theories, which are not mutually exclusive (Berenbaum, 1989). The two models are concentration addition (CA) or independent action (IA), applicable for samples that contain chemicals with similar- or dissimilar- mode of action, respectively (Scenihr, 2012; European Commission, 2009). These models may provide quantitative and reliable predictions of combined effects from existing information of the activity of individual components of complex mixtures, assuming that the mixture components do not interact and the effects are additive. Comparison of model prediction with experimentally observed combined effect can help infer the possible interactions among chemicals, such as antagonism (combined effect from mixture is weaker than that of individual chemical) or synergism effect (combined effect from mixture is stronger than the predication) (Syberg et al., 2009). These models have usually been applied for phenotypic effects, if these models work for molecular effects are unknown.

The challenges in the required resources and time to handle the toxicity testing efforts for the large and ever-increasing number of contaminants in various environmental matrixes have motivated a new vision of toxicity testing strategy as proposed by NRC (National research council) and USEPA (NRC, 2007). It calls
for a systematic transit from current resources-intensive and time-consuming *in vivo* whole animal-based testing to *in vitro* mechanistic toxicity pathway-based assays on cell lines, using cost-effective, reliable and high-throughput screening and tier testing (NRC, 2007). Toxicogenomics approach can provides comprehensive toxicity effects, reveal toxicity mechanisms and discern chemical (or class)-specific toxicity for potential causal agents identification (Ankley et al., 2006; Aardema et al., 2002). In recent years, an increasing amount of efforts have been initiated to perform quantitative assessment of toxicogenomics data, aiming quantify altered molecular profiles changes. For quantitative mixture assessment, it typically relies on monotonous changes in response as represented in sigmoidal concentration response relationships. Defining dose-response relationship using quantitative toxicogenomics data is possible using molecular disturbance quantifiers such as fold changes for specific biomarkers (Hutchins et al., 2010), number or percentage of genes with altered expression (Poynton et al., 2008) or more integrated parameter such as TELI (Transcriptional Effect Level Index) (Gou and Gu, 2011). These methods use signal intensity-based interpretations following the idea that the stressor induced molecular effects should be monotonously related to dose and to the change in differentially expressed signal intensity. The demonstrate dose response relationships using quantitative toxicogenenics data make is possible to further examine the mixture toxicity modeling at molecular levels, which has hardly been explored.
The intrinsic complexity and unresolved issues for conventional phenotypic mixture toxicology present formidable future research challenges for molecular mixture toxicology as well. Altenburger et al. reviewed numbers of experimental studies from the past decade that addressed diagnostic and/or mechanistic questions regarding the combined effects of chemical mixtures using toxicogenomics techniques (Altenburger et al., 2012). For quantitative mixture assessment, it typically relies on monotonous changes in response as represented in sigmoidal concentration response relationships. Majority of mixture toxicity studies via toxicogenomics focused on identifying conservative or novel signals by comparing the pattern to those from the reference cases (i.e. individual component) (Mumtaz et al., 2002; Garcia-Reyero et al., 2012; Mortensen et al, 2006). In these studies, no explicit hypothesis were provided on what an expected combined effect from a mixture exposure would look like, and no quantitative view on the data which allow extrapolation to other mixture compositions.

Few has systematically explored the framework of mixture modeling application at molecular levels, especially rare by using quantitative gene expression data. One of the reasons for this gap might be that an objective effect concentration (ECx) value very often is impossible to obtain from gene-induction measurements, because it is not clear whether there is any fixed maximum induction value for each gene or gene sets, and if the maximum induction level of a gene vary for different toxicants (Dardenne et al., 2008). Dardenne et al. attempted to develop dose-response curves using gene expression fold changes versus concentration (Dardenne et al., 2008). The compliance of molecular mixture toxicity with the
current model framework was explored by examination of the prediction of fold change of a limited number of biomarker genes in the mixture from those in individual chemical. Dardenne’s work showed that the mode of action of toxicants does not determine the optimal choice of model, both addition models (CA and IA) can accurately predict the promoter induction intensity changes of certain individual genes by binary mixture of chemicals regardless their MOAs are similar or dissimilar (Dardenne et al., 2008).

In this study, we demonstrated the application of a quantitative toxicogenomics-based approach for investigating the combined effects of various binary mixtures comprising of a range of pollutants with concentrations ranging from environmental relevant level to the benchmark level. We employed an GFP-fused whole-cell array (E.coli K12, MG1655) that measures temporal altered gene expression and quantifies molecular perturbations in known cellular stress response pathways involved in detecting and responding to stress and repairing damage caused by toxicants on primary cellular infrastructure and macromolecules (i.e. DNA, protein, membrane, oxidative) (Gou et al., 2010; Gou and Gu, 2011; Onnis-Hayden et al., 2009). These selected biomarkers represent set of homeostatic pathways that have been studied extensively and found to have high degree of conservation among various species (Simmons et al., 2009).

Our previous studies showed that the cellular stress responses pathway based quantitative endpoints (TELI) exhibit dose-response patterns and correlated to the phenotypic changes and adverse outcomes (Gou and Gu, 2011). A total of 5 binary mixtures were designed according to the availability of chemical MOAs
information and their relevance in environment. We proposed new molecular mixture toxicity addition models following the analogy and framework of conventional addition mixture models and adapted parameters derived from toxicogenomics assays. Our results demonstrated that the proposed mixture model can explicitly predict the combined effects at the molecular level, and the pathway-based quantitative toxicogenomics approach can be used for mixture toxicity studies.

5.3 Materials and Methods

5.3.1 Chemical selection and mixture combination design relational

Chemicals were selected either because they are well-studied mode compounds with a certain mode of action, as H2O2, mitomycin, lead, arsenic, and erythromycin, or they are prevalent in wastewater and drinking water, like NDMA, atrazine, and sucralose. 5 arbitrary chosen binary mixtures with diagonal/ray design (the components in each mixture are mixed based at a constant mixture ratio and then be diluted) are tested at 6 different concentrations (as shown in Table 5.1). Those five mixture combinations were designed to have chemical with similar structure (Arsenic-chronium, mitomycin C-erythromycin), dissimilar structure (H2O2-sucralose), comparative toxicity (NDMA-atrazine, mitomycin C-erythromycin), and less comparative toxicity (H2O2-sucralose, mitomycin C-sucralose). The fixed mixture ratio for each combination were determined by the ratio of single compound NOEL values on E.coli K12 MG1655 growth inhibition with two hour exposure in minimal medium (M9). To avoid
growth inhibition of mixture, half of the NOEL values were selected as the highest concentration for mixture toxicity assessment.

Chromium reference solution and arsenic reference solution were from Fischer Sci. Solvent used for dissolved those chemicals, such as methanol, ethanol, and dimethyl sulfoxide were from Fischer Sci. Deionized water (18.0 mΩ·cm) obtained from a Millipore Milli-Q system is used in all the experiments. All chemicals used in this study were above analytical grade. Atrazine, NDMA were from CHEM SERVICE (West Chester, PA), sucralose was from AK Scientific ( ), mitomycin C was from Research Products International Corp. (Mt. Prospect, IL), H₂O₂ (35 wt. % solution in water) was obtained from Sigma-aldrich and stored in 4°C.

5.3.2 Quantitative toxicogenomics assay with GFP-infused E.coli whole-cell array

A high-throughput mechanistic toxicity assay method was employed with GFP-fused whole-cell array of E.coli K12, MG1655. The tested chemicals were diluted with minimum media (M9) to desired concentrations. The selected assay library covers a variety of genes involved in different known cellular stress response pathways, such as oxidative stress, DNA stress, protein stress, and membrane stress (including membrane transporter, efflux pump, energy metabolism, flagella metabolism, lipopolysaccharide metabolism), etc. (see biomarkers in Table 2.1). Each fusion is expressed from a low-copy plasmid, pUA66, which contains a kanamycin resistance gene and a fast folding gfpmut2, allowing for real-time measurement of gene expression level changes (Zaslaver et al., 2006).
**Assay protocol**

*E.coli* was cultivated in 384-well plates (Costar, Bethesda, MD, USA) under 37°C in dark to avoid GFP photobleaching until the early exponential growth stage is reached (OD600 ~ 0.2). After addition of the aqueous samples, the plate was placed into a micro-plate reader (Synergy Multi-Mode, Biotek, Winooski, VT) for cell growth measurement (absorbance, 600 nm) and fluorescent reading (Excitation: 485 nm, Emission: 528 nm) every 5 minutes over a period of 2 hours under 37°C. Three biological replicates were performed for each condition. Test and control group are arranged on the same plate for each replicate.

**Data processing**

The *GFP* and *OD* data were first smoothed using 5-time moving average and then corrected against background (growth medium only blanks) with and without chemicals exposure, respectively. The population normalized *GFP* signal was calculated as $P=\frac{GFP}{OD}$ and corrected against background (*E.coli* strains without *GFP* infusion) at same *OD* with and without chemicals exposure, respectively. Induction factor, which measures the gene expression alteration, was calculated as the ratio of normalized expression levels between experimental (with chemical exposure) and control groups (without chemical exposure), $I=\frac{P_e}{P_c}$, where $P_e$ and $P_c$ represents for experimental and control groups, respectively. The induction factor $I$, determined for a given gene at each time point upon chemical exposure, was then used for the following data analysis, where a gene is up-regulated if $I>1$ and down-regulated if $I<1$. The toxicity mechanisms were
elucidated with the identifying and analyzing the changes in genes associated with specific stress response pathways.

### 5.3.3 Dose-response and toxicogenomics endpoints derivation

The molecular response was quantified by the TELI (Transcriptional Effect Level Index) value, which was a recently developed index for toxicogenomics data interpretation that incorporates the number, magnitude and the temporal pattern of genes with an altered expression. TELI can be integrated as equation 1 on various levels, as single gene (TELI\textsubscript{gene}), pathway (TELI\textsubscript{pathway}), stress category (TELI\textsubscript{oxidative}, TELI\textsubscript{geno}, TELI\textsubscript{membrane}, TELI\textsubscript{protein}, and TELI\textsubscript{general}), and overall stress library (TELI\textsubscript{total}). Compared to previous publication, improvement was made to TELI computation. In this paper, TELI is calculated as the average of transformed I instead of sum of transformed I to make TELI value more comparable among different pathway, stress category, and overall library.

\[
TELI = \frac{\sum\limits_{i=1}^{n} w_i e^{\text{ExposureTime}*n}}{\text{ExposureTime}*n}
\]

Where, \( t \) was the exposure time, \( n \) was the number of genes in one particular pathway/stress catgory/total stress library, \( w_i \) was the weight factor of gene\(_i\). For this study, we assigned value of 1 for all the weighing factors.

The TELI-dose response curves were analysis with a 4-paramater logistic equation model using GraphPad 5.0c (Prism). Several toxicogenomics endpoints were derived from TELI-dose response curves.
5.3.4 Gene set enrichment analysis

The gene set enrichment analyses (GSEA) are achieved by ranking a list of genes with the TELI values. For each pathway, GSEA calculates the enrichment score by examining the ranked gene list from high score end to the low score end, giving a rewarding score if a gene belongs to the pathway of interest, and penalizing score otherwise. The significance of each pathway is determined by comparing their ranking scores to the corresponding empirical distributions. The null distributions are generated by randomly permuting the specific pathway and all others 1,000 times. For details, please refer to the work of Subramanian (Subramanian et al., 2005). Categories with p-value less than 0.05 are considered to be significant expressed.

5.3.5 Principal components analysis

Principal component analysis (PCA) was performed for the toxicogenomics data for mixture as well as their single components to simplify the complex data sets of categories by analyzing the components with the greatest amount of variance based on their temporal gene expression profiles. The principal components analysis (PCA) was performed by software MeV 4.0 (MultiExperiment Viewer) based on gene expression levels of each chemical (I, average of triplicates) during 2-hr exposure, with centering mode as mean and number of neighbors for KNN imputation as 10.
5.3.6 Molecular mixture toxicity predication model (CA and IA model)

A molecular mixture addition models were proposed following the analogy and framework of mixture toxicity model with phenotypic observations. Both concentration addition (CA, equation 1) or independent action (IA, equation 2), were applied to the tested five mixture combination using parameters derived from TELI-dose response curves for both individual and mixture chemicals as following.

\[
\frac{c_1 + c_2}{ECx_{(mix)}} = \frac{c_1}{ECx_1} + \frac{c_2}{ECx_2}
\]

\[
E_{mix} = 1 - (1 - E_1)(1 - E_2)
\]

- \(c\) - the molar concentration of each chemical
- \(ECx\) - the concentration that can evoke molecular response \(x\), \(x\) refers to cellular response determined by TELI at certain concentration divided by TELImax (maximal response obtained from TELI-dose response curves) of single chemical/mixture.
- \(E\) - the molecular response at certain concentration, determined by TELI at certain concentration divided by TELImax (maximal response obtained from TELI-dose response curves) of single chemical/mixture.

Note that the determination of comparable ECx levels is not possible for endpoints without a known maximum response level, as is the case for gene expression–based systems here (Dardenne et al., 2008). This is because that the maximum induction level for every gene promoter is not only determined by its intrinsic regulatory mechanism but also dependent on the nature of the compound, therefore a classical toxic unit approach is not applicable. Here, we propose to determined the normalized molecular response at certain concentration as
percentage of TELImax, which is maximal response from TELI-dose response curves, for both single chemical as well as their binary mixture. The TELI dose-response curve made it possible to detect the maximal molecular response using statistics curve fitting tool. The ECx in equation 1 were replaced with EC-TELIx (Gou et al., 2013). Similarly, the effect E in equation 2 was cellular response determined by TELI at certain concentration divided by TELImax (maximal response obtained from TELI-dose response curves) of the chemical or mixture. Comparison of model prediction with experimentally observed combined effect can help infer the possible interactions among chemicals, such as antagonism (combined effect from mixture is weaker than that of individual chemical) or synergism effect (combined effect from mixture is stronger than the predication).

5.4 Results and Discussion

5.4.1 Comparison of toxicity profiles of individual components with their mixture

The temporal altered stress response gene expression profiles were obtained for the 5 mixture combination as well as their single components at 6 different concentrations over 2 h period (Figure 5.1). The results showed rather conserved chemical-specific toxicity for most individual chemicals as indicated by similar chemical-specific toxicity profiles patterns (mechanism) at different does concentrations such as atrazine, sucralose, erythromycin, H2O2, arsenic, and chromium (Figure 5.1). While for some chemicals, the toxicity nature seemed to have more pronounced dose-dependent transit changes such as NDMA and mitomycin C.
Comparison of stress response genes expression profiles of mixtures with that of each individual chemical component revealed similarity or distinction in the detailed toxicity mechanism and characteristics among them (Figure 5.1). To further evaluate the nature of toxicity of mixture in comparison to its components, we performed GSEA analysis that identifies the dominant toxicity effects among the stress response categories (Table 5.1). The metal mixture intended to represent the combination of chemicals with similar MOA and similar toxicity level. Both metals are known to induce ROSs and have cellular effects in membrane stress (i.e. efflux pump, lipid oxidation) and general stress (Kitchin et al., 2003; Kitchin et al., 2008). In addition, chromium is known to cause DNA damages such as strand breaks, DNA–DNA and DNA–protein cross-links nucleotides modification (Liu et al., 2001; Slade et al., 2005) while arsenic bind protein and further compromise protein folding pathway (Ramadan et al., 2008). The mixture of two metals seemed to exhibit the combined molecular effects of both metals across all toxicity categories.

The other three combinations are used to represent the combinations that contain components with both dissimilar MOA and dissimilar toxicity level. Sucralose was mixed with reference compounds for specific MOAs (H2O2 for oxidative stress, mitomycin C for DNA stress) (Tomasz et al., 1987) to represent the most distinctive scenario. The mixture effect differs between these two combinations, H2O2-sucralose led to more severe response (i.e oxidative stress) than each of the single chemical (potential synergistic), and the effect of mitomycin C-sucralose seems to be dominant by the less toxic component (potentially antagonistic).
Figure 5.1 Comparison of temporal altered stress response gene expression profiles of mixture with that of each individual chemical component. X-axis on the top is the different stress pathways/categories, the 107 stress genes in E.coli used in our assay library have been categorized into 5 functional groups based on their main function and involvement in different stress mechanisms, such as DNA
stress (DNA damage repair), oxidative stress (including ROS defense and redox balancing), protein stress, membrane (lipid stress and membrane transporter), and general stress (other stress as pH, temperature, et al. See Table 2.1). Y-axis on the left is the mixture/components, and the exposure concentration is increasing from top to down. Red means up-regulation, black means no response compared to non-exposure control.

For mixture that contain chemical with similar dominant MOAs, such as that contains mitomycin C and erythromycin, both have DNA identified as dominate MOA (Table 5.2). The dominant effect of mixture is the same as each of the two components that led to overrepresented DNA stress as the major MOA. The heavy metal (arsenic-chromium) mixture has both DNA and protein stress overrepresented at various concentrations, and the effect seems to be an addition effect of both single metal. Sucralose, an artificial sweetener with relatively weak toxicity and no known MOAs, is used as an additive composition for two mixture combinations with 2 reference compounds, H2O2 and mitomycin C, respectively. Interestingly, very different results were found for these two binary mixtures. Combination of sucralose with MMC did not exhibit dominant DNA stress although one of the components MMC is a strong DNA-damaging agent, while the mixture of sucralose and H2O2 exhibited oxidative stress as the major MOA for most of the concentration measured. For atrazine and NDMA, there seems to be dissimilar among mixture and single components.
Table 5.1 Overrepresented (p-value < 0.05) stress categories by stress gene set enrichment analysis

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration (uM)</th>
<th>Expressed stress</th>
<th>Concentration (uM)</th>
<th>Expressed stress</th>
<th>Concentration (uM)</th>
<th>Expressed stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arsenic-chromium</td>
<td>0.018</td>
<td>General</td>
<td>0.27</td>
<td>DNA</td>
<td>0.14</td>
<td>DNA</td>
</tr>
<tr>
<td></td>
<td>0.18</td>
<td>Protein</td>
<td>2.7</td>
<td>/</td>
<td>1.4</td>
<td>DNA</td>
</tr>
<tr>
<td></td>
<td>1.8</td>
<td>/</td>
<td>27</td>
<td>DNA</td>
<td>14</td>
<td>DNA</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>Protein</td>
<td>269</td>
<td>DNA</td>
<td>144</td>
<td>DNA, protein</td>
</tr>
<tr>
<td></td>
<td>186</td>
<td>Protein</td>
<td>2692</td>
<td>/</td>
<td>1439</td>
<td>DNA, protein</td>
</tr>
<tr>
<td></td>
<td>1866</td>
<td>Protein</td>
<td>26923</td>
<td>DNA</td>
<td>14394</td>
<td>/</td>
</tr>
<tr>
<td>Mitomycin C-erythromycin</td>
<td>0.000015</td>
<td>/</td>
<td>0.00136</td>
<td>DNA</td>
<td>0.00008</td>
<td>DNA</td>
</tr>
<tr>
<td></td>
<td>0.00015</td>
<td>DNA</td>
<td>0.001364</td>
<td>DNA</td>
<td>0.00075</td>
<td>DNA</td>
</tr>
<tr>
<td></td>
<td>0.001497</td>
<td>DNA</td>
<td>0.01364</td>
<td>DNA</td>
<td>0.00757</td>
<td>DNA</td>
</tr>
<tr>
<td></td>
<td>0.01497</td>
<td>DNA</td>
<td>0.1364</td>
<td>DNA</td>
<td>0.07569</td>
<td>DNA</td>
</tr>
<tr>
<td></td>
<td>0.1497</td>
<td>DNA</td>
<td>1.364</td>
<td>DNA</td>
<td>0.75698</td>
<td>DNA</td>
</tr>
<tr>
<td></td>
<td>1.497</td>
<td>DNA</td>
<td>13.64</td>
<td>DNA</td>
<td>7.5698</td>
<td>DNA</td>
</tr>
<tr>
<td>Atrazine-NDMA</td>
<td>0.000006</td>
<td>/</td>
<td>0.019189</td>
<td>Oxidative</td>
<td>0.0096</td>
<td>General</td>
</tr>
<tr>
<td></td>
<td>0.000058</td>
<td>/</td>
<td>0.19189</td>
<td>DNA</td>
<td>0.0959</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td>0.005814</td>
<td>General</td>
<td>1.9189</td>
<td>/</td>
<td>0.9594</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td>0.05814</td>
<td>/</td>
<td>19.189</td>
<td>DNA</td>
<td>95.945</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td>0.5814</td>
<td>General</td>
<td>1918.9</td>
<td>DNA</td>
<td>959.45</td>
<td>General</td>
</tr>
<tr>
<td>H2O2-sucralose</td>
<td>73.5</td>
<td>Oxidative</td>
<td>0.75</td>
<td>Oxidative</td>
<td>0.39</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td>147</td>
<td>Oxidative</td>
<td>7.5</td>
<td>/</td>
<td>3.9</td>
<td>Oxidative</td>
</tr>
<tr>
<td></td>
<td>294</td>
<td>Oxidative</td>
<td>75</td>
<td>/</td>
<td>39</td>
<td>Oxidative</td>
</tr>
<tr>
<td></td>
<td>588</td>
<td>Oxidative</td>
<td>755</td>
<td>/</td>
<td>389</td>
<td>Oxidative</td>
</tr>
<tr>
<td></td>
<td>1176</td>
<td>Oxidative</td>
<td>7556</td>
<td>/</td>
<td>3896</td>
<td>Oxidative</td>
</tr>
<tr>
<td></td>
<td>2352</td>
<td>/</td>
<td>75566</td>
<td>/</td>
<td>38959</td>
<td>/</td>
</tr>
<tr>
<td>Mitomycin C-sucralose</td>
<td>0.000015</td>
<td>/</td>
<td>0.75</td>
<td>Oxidative</td>
<td>0.38</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td>0.00015</td>
<td>DNA</td>
<td>7.5</td>
<td>/</td>
<td>3.8</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td>0.001497</td>
<td>DNA</td>
<td>75</td>
<td>/</td>
<td>38</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td>0.01497</td>
<td>DNA</td>
<td>755</td>
<td>/</td>
<td>378</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td>0.1497</td>
<td>DNA</td>
<td>7556</td>
<td>/</td>
<td>3778</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td>1.497</td>
<td>DNA</td>
<td>75566</td>
<td>/</td>
<td>37784</td>
<td>Oxidative</td>
</tr>
</tbody>
</table>

* “/” indicates no stress category was significantly (p-value<0.05) overrepresented by GSEA analysis
5.4.2 Comparison of mixture with their components via Principal Components Analysis

The high-resolution toxicogenomics data also allows for principal component analysis (PCA) to reveal the similarity/dissimilarity of the toxicity nature of the mixtures in comparison to their single components. The PCA separation of the binary mixture and its two components, and the distances between the mixture and each of the two components at different exposure concentrations are shown in Figure 5.2. In addition, the weight score of each stress response category within each principal component are also included in Figure 5.2 as well.

The mixture of two heavy metals showed very similar toxicity profile at varying doses as their toxicity profiles could not be separated and overlaid on top of each other and the mixture seems to be in the junction area of arsenic and chromium at lower concentrations. The distances between the mixture and its two individual components are comparable, although the toxicity profile of the mixture is more similar to arsenic than chromium based on the distance calculated by their PCA separation at varying concentrations. The weight score of different stress categories on each PC was determined on eigenvalues from covariance matrix, which reflects the significance of old vectors on each new vector. The protein stress and DNA stress are the dominant toxicity effects of two metals for all three principal components, which were the identified significant stress response effects for Cr, As, respectively (Table 4.1)

Mixture of NDMA and atrazine showed concentration-dependent toxicity profiles as shown by separated data points at different concentrations. The toxicity profiles
of NDMA at different concentrations are mostly separated along the PC2, and PC2 weight score profiles shows dominant DNA damaging effects along with oxidative stress and membrane stress. NDMA has been reported to exhibit carcinogenic activity as well as oxidative stress (Pereira et al., 1982; Coffin et al. 2000). Toxicogenomics data of atrazine at different concentrations are more separated along PC3 and PC1 (Figure 2), with PC3 predominantly exhibiting oxidative stress and PC1 signature by protein stress and oxidative, respective (figure 2d). Atrazine is known to induce protein adducts with the formation of atrazine-protein super-molecular by Van der Waals force and hydrogen bonds (Dooley et al., 2007). Up-regulation of 2-antioxidant enzyme coded genes sodB and katE were observed in this study and previous study (Jin et al., 2010).

Analysis of the distance between mixture and its two components showed that mixture is closer to NDMA than atrazine except the highest concentration.

Mixture of H2O2-sucralose, two chemicals with distinct known MOA and different toxicity level, are dramatically separated (Figure 5.2) and very distant from both individual chemicals in the mixture, although it is more close to sucralose than to H2O2. H2O2 was more separate along PC2 by oxidative stress effects.

The two mixtures containing mitomycin C, MMC-erythromycin and MMC-sucralose, showed similar PCA separation patterns. Both erythromycin and sucralose showed conserved toxicity mechanism and profiles at varying dose concentrations, as indicated by the nearly non-separable data points representing different doses. Dose-dependent separation of MMC was observed along all three
PCs, and all of them were dominated by DNA stress. Both mixtures contained MMC had toxicity profiles more similar to the other component, as the mixture toxicity profiles complete overlaid over the profiles of erythromycin and sucralose, for MMC-erythromycin mixture and MMC-sucralose mixture, respectively.
**Figure 5.2** PCA Graphical representations of mixture and their single components, based on Induction factor (I) during 2 hours exposure with six concentrations. Samples are color coded according to chemicals and each spot represents one treatment (a chemical at a given concentration) with bigger symbol indicating relative higher level of exposed concentration.

### 5.4.3 Dose-response patterns of mixture and its chemical components

The question on whether mixture effects are additive cannot be answered without information on the dose-response relationship for the single components (Lutz et al., 2002). We applied TELI (Transcriptional Effect Level Index) value to quantify the overall molecular stress response activities and they exhibited does response curves for all individual chemicals as well as for mixtures (Figure 3). As described in our previous works, TELI is a recently developed index for toxicogenomics data interpretation that incorporates the number, magnitude and the temporal pattern of genes with an altered expression. 4-parameter logistic regression model was applied to model the dose response curves.

For the two mixtures that contained components with 2 metals (As-Cr), and 2 antibiotics (MMC-erythromycin), parallel molecular dose-response curve were found among mixtures and it components (Figure 5.3). This is consistne with the covnetioanl phenotypic leve understanding that chemical with common toxicity MOA liely exhibit paralel dose repsons ecurves and therefore potential addition effect (European Commission, 2009). The dose response curves for the other three binary mixtures, however, had different curve shape and not parell
with either of the two components. For mixtures of atrazine-DNMA and MMC – surcalose, the mixture toxicity was lower than or very close to the less toxic component, indicating possible antagonistic effect. For the case of H2O2 in mixture with surcalose, the dose responses curves suggested dose-dependent effects with synergistic effect at lower dose range, and then transit to addition effect at higher dose range.
Figure 5.3. Dose-response analysis of mixture with their components
5.4.4 Quantitative prediction of combined effect at molecular level and comparison with observations

To quantitative analysis the combined effects of various binary mixtures, two additive molecular models were applied, concentration addition (CA), and independent action (IA), as described in methods section. Although the CA and IA models are originally derived to predict only chemicals with strictly similar or dissimilar MOAs, it is usually used regardless of knowledge of mode of action (MOA) or for baseline toxicity (Altenburger et al., 2012).

Figure 5.4 shows the comparison of the experimental toxicity data of the 5 binary mixtures with the predicted mixture toxicity using the molecular addition models. R square (r2) was used to represent the agreement of model predictions and experimental data. For comparison, a horizontal straight line was used as null hypothesis for the r2 calculation. r2 equals to 1 means perfect agreement, TELI values for the entire stress response pathways ensemble were used for the results shown in Figures. Table 5.2 lists the summary of model fitting evaluation parameters for all the mixtures. The results indicated that both the CA and IA models are able to predict the mixture response for mixture that contained components with similar MOAs and consequently exhibited parallel dose responses, as in the cases for As-Cr and MMC-erythromycin. For the 2 metals, great agreement was obtained between the experimental observed mixture toxicity and both model predication (r2>0.8 with p-value<0.05), and there is no significant difference between CA and IA models prediction (Figure 4). For the 2 antibiotics, good but not significantly agreement on both model predications, IA model fit
better than CA model. The NDMA – atrazine mixture exposure also exhibited the best-fit results with one model (IA). Clearly higher experimental toxicity than both model predications was found by the mixture of H2O2 at lower concentrations – sucralose indicating possible synergistic effect. However, interactions occurs with concentration increased, addition effect seems occurred at higher concentration. And lower experimental toxicity than both model predications was found by mitomycin C– sucralose at higher concentrations indicating possible antagonistic effect, while at the lower range, the mixture toxicity seems to follow addition model. Similar findings in literature also conclude that different interactions may occur depending on the dose concentration and ratios among chemical components (Altenburger et al., 2012).
**Figure 5.4** Mixture toxicity values indicated by probit unit. The open hole diamond indicates the observed values from toxicogenomics results, the red dashed line means the predicated values from CA model while blue are from IA model.

**Table 5.2** The predication r square for measure of regression models

<table>
<thead>
<tr>
<th>Mixture/Model/r² (p-value)</th>
<th>Concentration Addition (CA)</th>
<th>Independent Action (IA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>As+Cr</td>
<td>0.81 (0.013)</td>
<td>0.87 (0.006)</td>
</tr>
<tr>
<td>mitomycin+erythromycin</td>
<td>0.20 (0.38)</td>
<td>0.51 (0.11)</td>
</tr>
<tr>
<td>NDMA+Atrazine</td>
<td>-0.36 (0.38)</td>
<td>0.57 (0.08)</td>
</tr>
<tr>
<td>H2O2+surcalose</td>
<td>-1.41</td>
<td>-0.26</td>
</tr>
<tr>
<td>Mitomycin C+sucralose</td>
<td>-0.81</td>
<td>-1.43</td>
</tr>
</tbody>
</table>

*Numbers in parenthesis are the p-value of sum of square r²

**5.4.5 Mixture prediction on specific stress pathways**

As discussed earlier, it is generally considered that mixtures containing components with similar model of action comply with the addition mixture prediction model. We therefore further explore mixture model prediction for each individual stress response pathways, with underlying assumption that the effects for each individual stress pathway (indicating similar toxicity mechanism) more likely to comply with addition model. Table 5.3 summarizes the prediction results with both CA and IA models of 5 mixtures for oxidative stress, DNA stress and protein stress response pathways, respectively.

Similar to the total stress pathway prediction, addition models have best predictions on the toxicity of the 2 metal mixture for the all specific stress
categories, however, CA is not applicable for the 2 metal mixture on protein stress. For MMC-erythromycin, addition models work good but not significantly for oxidative and protein stress, however, DNA stress was found to exhibit synergistic effect. For NDMA-atrazine, only antagonistic effect was found on DNA stress. For H2O2-sucralose, both models showed good results on protein stress, however, synergistic effect on oxidative stress while antagonistic effect on DNA stress. Combination of MMC-sucralose exhibited great agreement between both model prediction and experimental toxicity on DNA stress, while antagonistic effect on oxidative stress.

Table 5.3. The prediction r square for measure of regression models for specific stress pathways

<table>
<thead>
<tr>
<th>Mixture model prediction (r²/p-value)</th>
<th>Oxidative Stress</th>
<th>DNA Stress</th>
<th>Protein Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CA</td>
<td>IA</td>
<td>CA</td>
</tr>
<tr>
<td>As+Cr</td>
<td>0.35 (0.21)</td>
<td>0.69 (0.04)</td>
<td>0.91 (&lt;0.01)</td>
</tr>
<tr>
<td>Mitomycin+Erythromycin</td>
<td>0.57 (0.08)</td>
<td>0.56 (0.08)</td>
<td>Synergistic</td>
</tr>
<tr>
<td>NDMA+Atrazine</td>
<td>NA</td>
<td>NA</td>
<td>Antagonistic</td>
</tr>
<tr>
<td>H2O2+Sucralose</td>
<td>Synergistic</td>
<td>Antagonistic</td>
<td></td>
</tr>
<tr>
<td>Mitomycin C+Sucralose</td>
<td>Antagonistic</td>
<td>0.78 (0.02)</td>
<td>0.91 (&lt;0.01)</td>
</tr>
</tbody>
</table>

*NA indicates either the single compound or mixtures have no clear dose-response on the specific stress response pathway with tested concentration range.
5.5 Conclusion

However, mixture toxicity studies based on quantitative gene expression data are, to our best knowledge, scarce from the literature. It is striking that apart from purely descriptive approaches, quantitative analysis of combined effects, such as linear regressions, were rarely performed. One of the reasons for this gap might be that an objective effect concentration (ECx) value very often is impossible to obtain from gene-induction measurements, because it is not clear whether there is any fixed maximum induction value for each gene or gene sets, and if the maximum induction level of a promoter:gene combination vary for different toxicants for a specific gene cassette. Both addition models, however, depend specifically on these accurate ECx values. Looking at both models as being intrinsically different from one another. Dardenne et al. (Dardenne et al., 2008) explored the molecular mixture effects for 8 different binary mixtures in dilution series via toxicogenomics assay using 14 stress response related gene reporters using E.coli K12 SF1 deletion library. Differentially expressed genes shared among individual chemicals and mixture were identified and dose response curves were developed as fold changes versus concentration. The compliance of molecular mixture toxicity with the current model framework were explored by examination of the prediction of fold change of specific genes in mixture from those in individual chemical. Various addition models were employed and demonstrated that at single gene level, the model framework may work.

Our results demonstrated that the concept of combined effects is presented at molecular level and they seem to comply with current mixture toxicity model.
The possibility to explicitly predict the combined effects at the molecular level (when additive effect is present) has been showed. Our toxicogenomics method, which measures the transcriptional level responses of up to 120 stress-related genes, have been illustrated as a rapid, sensitive, and informative approach for mixture toxicity studies. Although most of the mixture exhibited additive effect in our results, further investigation is required for better understanding of the factors that influence the combined effects.
Chapter 6

Link Quantitative Toxicogenomics Endpoints to
Conventional Toxicity Endpoints - Phenotypic Anchoring
and Cross Species Extrapolation

6.1 Abstract

One of the most challenging aspects of implementing toxicogenomics-based toxicity assessment in regulation and in risk assessment involves establishing the theoretical framework and comprehensive knowledge base required to obtain quantitative toxicogenomics assay endpoints and correlate them with relevant phenotypic adverse outcome. Although there are many researchers who investigated the relationship between variety of gene expression level and organism level phenotypic endpoints, such as organisms weight, cytotoxicity assay and survival rate etc., most of them only performed qualitative analysis rather than quantitative analysis. The investigation of the phenotypic anchoring of toxicogenomics derived endpoints to phenotypic adverse outcome for ecotoxicity is only emerging. This study explored and demonstrated the quantitative correlation between endpoint- TELI (Transcriptional Effect Level Index) derived from a stress response pathway ensemble-based toxicogenomics assays in E coli. with conventional apical endpoints, based on growth inhibition in E.coli for 30 environmental chemicals. Furthermore, to explore the derivation of these correlations across different species, we also expanded the evaluation to correlate
the TELI values with growth inhibition of V.fisheri, single cell green algal S.capricornutum and fish cell line RTgillW-1. These species were used since they are all widely used model organisms for water quality monitoring standard methods by ISO and USEAP. Person’s product-moment coefficient and coefficient of determination (R2) are applied to evaluate the correlation between toxicogenomics results and apical endpoints from conventional toxicity test. Our results showed that 2 TELI derived endpoints, TELImax, EC-TELI1.5 correlated well with EC50 based on growth inhibition in E.coli, luminescence inhibition on V.fisheri, and growth inhibition of single cell green algal S.capricornutum. However no significantly correlation with growth inhibition on fish cell line RTgillW-1 was observed. This study explored and suggested possibly link between integrated endpoints derived from temporal altered gene expressions in conserved cell stress response pathways biomarkers with conventional phenotypic endpoints EC50 in single organisms across species. Our results showed that toxicogenomics data can be quantitatively link to conventional toxicity endpoints for water toxicity assessment, thus indicates the possibility of utilizing our quantitatively toxicogenomics approach with cellular stress ensemble for water toxicity assessment.

6.2 Introductions

The emerging filed of toxicogenomics, which allows for high-throughput concurrently monitoring of the status of cellular response pathways globally upon exposure to chemical toxicants, promises a revolutionary new ground the Realization of Tox21Vision (Waters et al., 2004; Pennie et al., 2000). When
applied to the study of large classes of chemicals or drugs, toxicogenomic information can be used to globally define modes or mechanisms of toxic action (Pennie et al., 2000; Sandy, 2002; Armitage et al., 2014; Nicholson et al., 2002). Compare to conventional toxicity test, toxicogenomics assay has the advantages of targeting primarily cellular responses, revealing toxicity mechanisms, broader range of test doses, high throughput and quick turn-around, significantly lower cost and no or minimal animals use.

However, the benefits of toxicogenomics information and tools in regulatory ecotoxicology are only starting to be elucidated. Existing challenges limited their application for environmental regulatory and risk assessment (“Regulation (EC) No 1907/2006”, 2006; NRC, 2007a). One of the most challenging aspects of implementing toxicogenomics-based toxicity assessment in regulation and in risk assessment involves establishing the theoretical framework and comprehensive knowledge base required to obtain quantitative toxicogenomics assay endpoints and correlate them with relevant phenotypic adverse outcome (Krewski et al., 2010; NRC, 2007b) 9,10. This process is necessary to validates the basis of molecular toxicity assays, distinguish more defined and focused molecular toxicity assays, entail possible interspecies extrapolation and to gain greater confidence in animal models for the possible realization of the vision of mechanistic predictive toxicity. In addition, to further incorporate these in vitro bioassays into current environmental regulations and decision-making framework, comparison and correlation with conventional whole-animal tests are unavoidable. Furthermore, cross-species extrapolation of biological response represents a
cornerstone in predictive toxicology, which is intimately dependent on its reliability to accurately predict harmful effects of chemical substances or environmental samples in humans or other recipient species (Lalone et al., 2013; Perkins et al., 2013).

However, great challenges and knowledge gap exists in the emerging field of toxicogenomics with regard to defining the relationships between chemically induced changes in gene expression and alterations in conventional toxicological parameters such as conventional toxicity endpoints, clinical chemistry and histopathology. Adverse outcome pathway (AOP) is an established framework for links among specific adverse outcome to one or multiple molecular effect signatures validates the basis of molecular toxicity assays (NRC, 2006; Ankley et al., 2010; Kramer et al. 2011; “Adverse Outcome Pathway Wiki”, 2013). As stated by Ankley et al., AOP describe the known linkage along the continuum from the molecular event in which a chemical interacts with a biological target(s); following on through a sequential series of cellular, anatomical and functional changes in biological process, and ultimately culminating in an adverse outcome of relevance to human or ecological risk assessment (Ankley et al., 2010). Characterization of AOPs allows for identification of key events based on which high-throughput testing methods can be developed. AOPs also provide important information on the development of structure-activity relationships, i.e. using effects information from one chemical (the source chemical) to predict the effect for another structurally similar chemical (the target chemical) (Ankley et al., 2010; “Adverse Outcome Pathway Programme”, 2013). Finally, AOPs provide evidence
important for qualitative and quantitative predicative models of the adverse outcomes that result from triggering molecular initiating (MIE) or other key events (Ankley et al., 2010; “Adverse Outcome Pathway Programme”, 2013).

AOPs provided broad potential to serve as novel tools in toxicology and risk assessment. They are designed to provide a clear-cut mechanistic representation of critical toxicological effects that span over different layers of biological organization. However, most of the current developed and under developing AOPs are for receptor-mediated toxicity with specific mechanism (“Adverse Outcome Pathway Wiki”, 2013; “Adverse Outcome Pathway Programme”, 2013).

For environmental risk assessment, various chemicals with different specific mode of action present in environmental media simultaneous, which implies that the overall toxicity comprises multiple specific toxicity mechanisms exerted by single chemicals, and complicated with synergistic or antagonistic effects. To achieve more comprehensive evaluation of environmental risk, batteries of bioassays consisting of a number of separated assays targeting different toxicity endpoints and with different AOPs can be applied. However, batteries of bioassays still require a substantial amount of time and efforts. In addition, most of them only provide information for specific mode of action toxicity endpoints without more comprehensive and detailed information of the toxicity profiles and mechanisms. Furthermore, isolated assays that cover only one or a few biomarkers or specific toxicity mechanisms effects cannot be directly anchored translated to an integrated adverse outcome or phenotypic endpoints.
Recently, a number of researchers have investigated the relationship between variety of molecular endpoints, such as altered gene expression level, and organism level phenotypic endpoints, such as organisms weight, cytotoxicity assay and survival rate etc. (Gottschalg et al., 2006; Ovando et al., 2010; Zheng et al., 2011; Powell et al., 2006; Deferme et al., 2013). Powell et al (Powell et al., 2006) reported that toxicity endpoints and protein adduct formation was used to qualitative phenotypically anchor oxidative stress gene expression due to acetaminophen exposure. This oxidative stress signature was observed that was corroborated in a time-dependent manner with increases in oxidized purines and abasic sites in DNA (Powell et al., 2005). Bugiak and Weber also qualitative analyzed the relationship between gene expression level of a group of CYP genes and the deformity, fork length of zebrafish egg via RT-PCR for a variety of toxicants (Bugiak and Weber, 2010). Ung et al. qualitatively linked transcriptomics results to the changes of cell morphology, cell adhesion and macromolecular accumulation (fatty acid and glycogen) in zebrafish liver cell via gene set enrichment analysis (GSEA) (Ung et al., 2011). Another example of phenotypic anchoring on ecotoxicity is the investigation on Green alga Scenedesmus vacuolatus between biochemical composition and cell volume changes, photosynthesis efficiency inhibition (Sans-Piché et al., 2010). Although there are many researchers who investigated the relationship between variety of gene expression level and organism level phenotypic endpoints, most of them only performed qualitative analysis rather than quantitative analysis. Great correlation between cellular responses on protein expression with observed DNA
lesion has been reported by Comet assay, which reported by Lan et al. (Lan et al., 2015).

This study applied a newly developed quantitative stress-response ensemble based toxicogenomics assay to evaluated the toxicity of various classes of CECs, and explore the correlation between the molecular response with stress response pathway perturbation with phenotypic response within and across species on cytotoxicity/adverse outcome. A quantitative molecular endpoint from toxicogenomics assays in E. coli - Transcriptional Effect Level Index (TELI)- was developed from dose response curves. Then the TELI values was correlated with EC50 values determined in E. coli, V. fisheri, single cell green algal S. capricornutum and fish cell line RTgillW-1. These species were used since they are all widely used model organisms for water quality monitoring standard methods by ISO and USEAP.

6.3 Materials and Methods

6.3.1 Chemical selection and evaluation

The 30 CECs were selected either because they are prevalent in wastewater and drinking water, like NDMA, atrazine, and sucralose. Bisphenol A, benzo[a]pyrene, and carbaryl, or they are attracting most research interest for their implication in recent years, the tested 4 nanomaterials, bis(2-ethylhexyl) phthalate. Also 6 well-studied chemicals with known mode of action were selected as mode compounds for reference.
The tested chemicals used in this study are above analytical grade. The detailed information is summarized in Appendix 4.1. Solvent used for dissolved those chemicals, such as methanol, ethanol, and dimethyl sulfoxide are from Fischer Sci. Deionized water (18.0 mΩ·cm) obtained from a Millipore Milli-Q system is used in all the experiments. All chemicals used in this study are above analytical grade.

6.3.2 Toxicogenomics assay

Assay description

A high-throughput mechanistic toxicity assay method was employed with GFP-fused whole-cell array of E.coli K12, MG1655. The tested chemicals were diluted with minimum media (M9) to various desired concentrations. The selected assay library covers a variety of genes involved in different known cellular stress response pathways, such as oxidative stress, DNA stress, protein stress, and membrane stress (including membrane transporter, efflux pump, energy metabolism, flagella metabolism, lipopolysaccharide metabolism), etc. (see biomarkers in STable 1). Each fusion is expressed from a low-copy plasmid, pUA66, which contains a kanamycin resistance gene and a fast folding gfpmut2, allowing for real-time measurement of gene expression level changes.

Assay protocol

The protocol to measure the temporal gene expression profile is described in our previous reports. In brief, E.coli was cultivated in 384-well plates (Costar, Bethesda, MD, USA) in dark to avoid GFP photobleaching until the early exponential growth stage is reached (OD600 ~ 0.2). After addition of the aqueous
sample, the plate was placed into a micro-plate reader (Synergy Multi-Mode, Biotek, Winooski, VT) for cell growth measurement (absorbance, 600 nm) and fluorescent reading (Excitation: 485 nm, Emission: 528 nm) every 5 minutes over a period of 2 hours. Three biological replicates were performed for each condition. Test and control groups are arranged on the same plate for each replicate.

**Data processing**

The $GFP$ and $OD$ data were first smoothed using 5-time moving average and then corrected against background (from growth medium only blanks) with and without chemicals exposure, respectively. The population normalized $GFP$ signal was calculated as $P = GFP/OD$ and corrected against background (from *E.coli* strains without $GFP$ infusion with same $OD$) with and without chemicals exposure, respectively. Induction factor, which measures the gene expression alteration, was calculated as the ratio of normalized expression levels between experimental (with chemical exposure) and control groups (without chemical exposure), $I = P_e / P_c$, where $P_e$ and $P_c$ represents for experimental and control groups, respectively. The induction factor $I$, determined for a given gene at each time point upon chemical exposure, was then used for the following data analysis, where a gene is up-regulated if $I > 1$ and down-regulated if $I < 1$. The toxicity mechanisms were elucidated by identifying and analyzing the changes in genes associated with specific stress response pathways.
6.3.3 Dose-response and toxicogenomics endpoints derivation

The molecular response was quantified by the TELI (Transcriptional Effect Level Index) value, which was a recently developed index for toxicogenomics data interpretation that incorporates the number, magnitude and the temporal pattern of genes with an altered expression. TELI can be integrated as equation 1 on various levels, as single gene (TELI\text{gene}), pathway (TELI\text{pathway}), stress category (TELI\text{oxidative}, TELI\text{geno}, TELI\text{membrane}, TELI\text{protein}, and TELI\text{general}), and overall stress library (TELI\text{total}). Compared to previous publication, improvement was made to TELI computation. In this paper, TELI is calculated as the average of transformed I instead of sum of transformed I to make TELI value more comparable among different pathway, stress category, and overall library.

\[
TELI = \frac{\sum_{i=1}^{n} w_i e^{\frac{I_{\text{gene}}(i=1)}{\text{ExposureTime}^* n}}}{\text{ExposureTime}^* n}
\]

Equation 1

Where, \( t \) was the exposure time, \( n \) was the number of genes in one particular pathway/stress catgory/total stress library, \( w_i \) was the weight factor of gene\( i \). For this study, we assigned value of 1 for all the weighing factors.

The TELI-dose response curves were analysis with a 4-paramater logistic equation model using GraphPad 5.0c (Prism). Several toxicoenomics endpoints were derived from TELI-dose response curves.

**Dose-response and endpoint derivation**
The TELI-dose response curves were analysis with a 4-paramater logistic equation model using GraphPad 5.0c (Prism). Five toxicoenomics endpoints were derived from TELI-dose response curves. Four endpoints were derived from TELI-dose response curves. The TELIMAX, referred as the maximal efficacy, reflects the limit of the dose response relationship on the response axis to a certain chemical. It quantifies the maximal transcriptional level effect for a group of given gene that can be induced in a 2 h exposure. TELI50 was determined as the similar approach of EC50, LC50 (effect concentration, lethal concentration) as the concentration that leads to 50% of the maximal adverse effect. We also applied 2 point of departure (POD) approach, namely NOTELTELI, as the concentration at which evoked 5% TELIMAX response, and EC-TELI1.5 which reflect the concentration that causes TELI value to reach 1.5, similar to the approach that has been applied on umuC genotoxicity assay by Escher et al. TELI50, NOTELTELI, and EC-TELI1.5 can be referred to as “potency”, which refers to the dose required to produce a given response intensity. Slope of fitted line (SlopeTELI) is another endpoint since it reflects the rate at which the transcriptional effect increases with the increasing of concentration.

6.3.5 Phenotypic assays

**E.coli growth inhibition test**

The conventional 24 h EC50 values for the three NMs selected are determined using the same E. coli strain based on growth rate inhibition. E. coli K12 were incubated at 37 C for about 2 h to obtain OD of about 0.1 before the stock solution was added to obtain various concentrations on microplates. Duplicates were performed for each concentration. After 24 h incubation, both colony-
forming unit (CFU) counting and absorbance (600 nm) measurements were conducted to determine the extent of cell growth inhibition at various concentrations. EC50 was then calculated based on the dose response curve of growth inhibition (%) using four-parameter logistic nonlinear regression model equation (Graph-Pad PRISM 5, La Jolla, CA 92037).

**V.fisheri luminescence inhibition test**
Photobacterial Vibrio fisheri (V.fisheri) (ATCC #49387) specific photosynthesis inhibition test (Microtox) was used for nonspecific toxicity evaluation according to the ISO standard method 11348-238. The assay was performed in solid black 96-well microplate (Costar, Bethesda, MD, USA). Briefly, fresh prepared V.fisheri culture was equally mixed with desired water sample, and the luminescence inhibition was calculated by the difference between light intensity at time 0 and 15 minutes measured with a micro-plate reader (Synergy Multi-Mode, Biotek, Winooski, VT). 1 M NaCl solution was used as negative control and ZnSO4 was used as a positive control as guidance. The result from Microtox tests were indicated as the inverse of EC50, which was the concentration causing a 50% photosynthesis inhibition effect with the concentration reported here in units of relative enrichment factor (REF), REF is the product of enrichment factor of lyophilization process and dilution factor of bioassay.

**Green alga S.capricornutum growth inhibition test**
Selenastrum capricornutum (UTEX 1648) was purchased from UTEX Culture Collection of Algae, Austin, TX, USA, and was maintained at 22±1°C with a 16 h light, 8 h dark with mix of cool-white/warm-white fluorescent lamps.
The algal growth inhibition assay was performed according to the U.S. EPA method (1989). Cultures were prepared in 100-mL Erlenmeyer flasks containing 20 mL of the filter-sterilized (0.22μm) test medium in triplicate. Test chemicals were dissolved in DMSO/methanol and added to each culture (final concentration of DMSO: less than 0.1%). Each chemical was tested in a dilution series of at least 5 concentrations. Both controls and test flasks were inoculated with exponentially growing algae at an initial concentration of S. capricornutum of 1x10^4 cells/mL. The cultures were incubated at a temperature of 22±1°C, and shaken at manually twice per day under a periodical (16 h light, 8 h dark) illumination of 4000 ±400 lux. After 96 h cell counts were determined using a microscope and a hemocytometer. The median inhibition concentration (IC50) values were calculated by the linear interpolation method (U.S. EPA, 1989). The concentrations of test chemicals were not analyzed; therefore, their nominal concentration was used as the exposure concentration in the calculation of IC50 values.

**Fish cell line RTgill W-1 growth inhibition test**

An in vitro cytotoxicity test with RTgill-W1 cell line was performed according to the method described by Tanneberger et al (Tanneberger et al., 2012). RTgill-W1 cell line (ATCC® CRL-2523™) acquired from the American Type Culture Collection (ATCC, Manassas, VA, USA) was cultured under standard conditions [17] in white L-15 medium (Sigma Aldrich, Czech Republic) supplemented with 10% fetal bovine serum (Sigma-Aldrich, Czech Republic). The cells were seeded to a 96-well plate (35,000 cells per well) and after 24 hours of incubation at
19.6°C, exposed to chemicals for 24 hours. Cell viability was assessed using cell counting with a microscope and a hemocytometer. The results are reported as a fraction of growth in control, which expresses the viability of cells exposed to chemicals compared to the negative control (i.e. cells in a clean medium with 50% v/v sterile distilled water, growth in control = 1). At least 6 concentrations were tested in triplicate and DMSO (Sigma Aldrich, Czech Republic). Bioassays data were analyzed in GraphPad Prism 5 Software (GraphPad Software, Inc., San Diego, CA, USA), where Hill’s model was applied for sigmoidal dose-response curve fitting to estimate IC50 values.

6.4 Results and discussion

6.4.1 Integrated molecular endpoints based on cellular stress response biomarkers ensemble relate to cellular adverse outcomes EC50

To use mechanistic data derived from toxicogenomics to support chemical assessments, there is a need for effective translation of these molecular response information into endpoints meaningful to ecological risk—effects on survival, development, and reproduction in individual organisms and, by extension, impacts on populations (Ankley et al., 2010; Boverhof et al., 2006; Kramer et al., 2011). There is a lacking of available quantifier for molecular response quantification.

In this study, we examined 30 chemicals as well as binary mixtures via toxicogenomics approach focusing on conserved cellular stress response pathways
that are associated and indicate certain cellular and macromolecular damages such oxidative stress sensor and defense, DNA damage and repair, protein damage signaling and repair, drug resistance, energy stress, cold/heat shock, cell killing and several key general stress regulator (Table 3.2). The assay used yield temporal altered gene expression profiles for each chemical at varying dose concentrations (Chapter 4). This molecular response is quantified by the TELI (Transcriptional Effect Level Index) value, which is a recently developed index for toxicogenomics data interpretation that incorporates the number, magnitude and the temporal pattern of genes with an altered expression. The TELI values exhibit dose dependent pattern for various toxicants from overall cellular stress response to single stress response pathway (Figure 4.3). These TELI-based dose response curves were modeled by a Four-parameter Logistic equation (4PL) model, which theoretically exhibit a characteristic “sigmoid” shaped toxicity dose response curve. A variety of endpoints were derived via TELI dose-response curves. The TELIMAX, referred as the maximal efficacy, reflects the limit of the dose response relationship on the response axis to a certain chemical. TELI50 was determined, in the similar approach of EC50, LC50 (effect concentration, lethal concentration), as the concentration that leads to 50% of the maximal adverse molecular effect. We also derived 2 other molecular endpoints using point of departure (POD) approach, namely NOTELELI, as the concentration at which evoked 5% TELIMAX response, and EC-TELI1.5 which reflect the concentration that causes TELI value to reach 1.5, similar to the approach that has been applied on umuC genotoxicity assay by Escher et al. TELI50, NOTELELI, and EC-
TEL1.5 can be referred to as “potency”, which refers to the dose required to produce a given response intensity. We consider EC-TEL1.5 as a more reasonable POD approach, since 1.5 is a commonly accepted criterion for gene expression evaluation, and a fixed expression level, which is greater than 2 fold standard derivation. Slope of fitted line (Slope\textsubscript{TELI}) is another endpoint since it reflects the rate at which the transcriptional effect increases with the increasing of concentration. Note that the currently developed molecular endpoints for toxicogenomics studies, as TELI, NOTEL, are still in exploration, very few has been accepted and widely used yet because how these molecular endpoints correlate or link to phenotypic endpoints, is an active research area under investigation.

We explored the correlations between molecular endpoints and phenotypic endpoint EC50 of growth inhibition in E coli using different TELI-based endpoints as describe above, including TELImax, TELI50, NOTEL\textsubscript{TELI}, EC-TEL1.5 and Slope\textsubscript{TELI} (Table 4.3). Among the different TELI-derived molecular endpoints as discussed above, the 2 that exhibited the best correlation with EC50 were TELImax and EC-TEL 1.5 for linking of molecular stress response to cytotoxicity. TELImax is recognized as the threshold condition when the transcriptional level effect, including both direct and compensatory effects here in our study, is at its maximal and, beyond which cell damages start to occur and eventually progresses to observable phenotype damage endpoint such as growth inhibition or even death. Therefore, the concentration that may evoke TELImax is very close to the concentration that may cause phenotype damage. EC-TEL1.5 is
cellular benchmark dose (BMD), which represents the threshold that molecular effects are not likely to occur below that.

The significant and good correlation between the above molecular endpoints with cytotoxicity EC50 (24 hour growth inhibition on E.coli) among 25 single chemicals of various structure and categories as well as for 5 binary mixtures, indicate that “phenotype anchoring” is possible for predicting single cellular cytotoxicity based on integrated molecular endpoints that capture the cellular disturbance via stress response pathways.
Figure 6.1 1 Correlation between molecular endpoints derived from 2-hr toxicogenomics assay based on conserved cellular stress response pathway via biomarkers ensemble using GFP-recombinants E coli reporters strains, and phenotypic endpoints EC50 in n E.coli 24 hours growth inhibition. Log scale are used for all concentration-based endpoints.
6.4.2 Cross species extrapolation- correlation of molecular endpoints based on stress response ensemble in E coli to EC50 of prokaryotic bacterial V.fisher luminescence inhibition (Microtox)

Ecological risk assessment usually comes from single-species toxicity tests measuring effects to individuals. However, populations, communities, and ecosystems are generally the entities to be protected (Newman et al., 2000). With rather limiting available toxicity data, underestimation of potential hazard is real possibility. To resolve the incongruity between individual-based data and the complex biological entities addressed in ecological risk assessment, species-sensitivity distribution or cross-species extrapolation methods are being incorporated into assessment of ecological risk and into recommendation for chemical registration, especially pesticide (Solomon et al. 1996; Hall et al., 1998; USWPA, 1998). The types of taxa selected for SSD and extrapolation may span several taxonomic groups, fall within a single group, or share certain habitat preferences (Abell et al., 2009).

In this study, we explored the correlation of cellular stress response quantifier TELI-based molecular endpoints, TELI\textsubscript{max}, EC-TELI1.5, respectively, with phenotypic endpoints obtained from Microtox assay using Vibrio fischeri, a non-pathogenic, marine, bacteria that luminesce as a natural part of their metabolism. The Microtox assay measures the disruption of respiratory process, resulting output related to oxidative stress, and baseline narcosis toxicity (Johnson, 1998, Pery et al., 2013; Poljsak et al., 2013; Saballunas et al. 1998). Significant and good
Great correlations (R>0.7, P<0.01) are found between the TELI derived endpoints with EC50 on Microtox test.

In Microtox assay, response to toxicity in *Vibrio fischeri* is observed as the disruption of respiratory process, resulting output change in luminescence (“Microtox 500”, 2014). The light-emitting reaction in V.fisheri involves the oxidation of reduced flavinmononucleotide (FMNH₂) and a long chain fatty aldehyde with the emission of blue-green light. The FMNH₂ level decline under strong oxidative stress. FMNH₂ also take apart in intracellular ROS production during defense of exogenous stimulation (Poljsak et al., 2013). Microtox is a standard method that has a variety of environmental and industrial applications for its ease of use and high-throughput configuration, economic, rapid response, and ecological relevance. Positive correlations were found by comparing Microtox results with toxicity values for fish, crustaceans and algae (Leeder Consulting, 2014).

Oxidative stress and exogenous stimulations could be responsible for the reduction of light emitting (Pery et al., 2013). Microtox was also used to represent baseline narcosis toxicity since it responds well to a wide array of hydrophobic chemicals [Johnson, 1998, Saballunas et al. 1998] likely due to its sensitivity to membrane-binding or membrane damage chemicals. Chemicals tested in this study include some chemicals known to have baseline toxicity such as organic hydrophobic chemicals including phenol, triclosan, ibuprofen etc.. Some chemicals seemed to induce oxidative stress response including H₂O₂, arsenic, chromium, etc.. Since our assay captures the effect on oxidative stress, membrane
damage, it is conceivable that there was likely correlation between the Microtox results and stress response ensemble-based toxicogenomics endpoints, as shown in Figure 6.2.

This first exploration of the correlation of molecular endpoints derived from E. coli with phenotypic endpoints from Microtox assay confirms our hypothesis that stress responses are highly conserved among species, therefore allowing possible cross species extrapolations among different bacteria. Therefore, the cellular stress response pathway ensemble based toxicogenomics assay in E. coli could provide relevant toxicity endpoints, similar to the standard Microtox assay, and in addition, provides more detailed mechanism toxicity information and “fingerprints”.
Figure 6.2 Correlation between molecular endpoints derived from 2-hr toxicogenomics assay based on conserved cellular stress response pathway via biomarkers ensemble using GFP-recombinants E coli reporters strains, and phenotypic endpoints EC50 in V.fisheri luminescence inhibition, also known as Microtox. Log scale are used for all concentration-based endpoints.
6.4.3 Cross species extrapolation - correlation of molecular endpoints based on stress response ensemble in E coli to EC50 with lower eukaryotic, green alga, S.capricornutum 96 hours growth inhibition

As previously discussed, cross-species extrapolation is essential in ecotoxicity and ecological risks assessment because there is often limited toxicological data with certain organism and translation of the testing results to effect on other organism in the ecological community is often necessary. Alga is the critical primary producer and identity in aquatic systems, and the unicellular green alga Selenastrum capricornutum (S.capricornutum) (Rhapidocelis subcapitata) growth inhibition assay is commonly used by EPA, ISO for water toxicity test.

S.capricornutum is eukaryotic single cell organism and it is the most commonly used organism for bio-monitoring to assess the level of nutrients or toxins in freshwater environments due to its high sensitivity to toxicants in water, which can alert subtle changes in water condition before a problem becomes excessive.

Figure 6.3 shows the significant correlation between S.capricornutum 96 hour growth inhibition assay with both of the TELI derived endpoints, namely TELImax and EC-TELI1.5. These correlations indicate that the across species correlation between stress response pathway perturbation and cytotoxicity was also possible between prokaryotic cell and eukaryotic cell, although the correlation coefficient is not as those between prokaryotic cells. S.capricornutum is single cell organism with similar membrane structure to gram-negative bacterial, such as E coli. However, significant differences exit in the cell structure between eukaryotic S.capricornutum and prokaryotic E coli. For example, the presence of
mitochondria and the essential feature of the photosynthetic apparatus, which may entail distinctive MOAs that are responsible for cytotoxicity in green alga and in E coli. Nevertheless, a significant correlation was observed between the molecular quantifier for disturbance and activation of stress response pathways in E coli and the cytotoxicity impact in unicellular algae, for over 25 single chemicals and 5 mixtures. These results evidenced that the stress response pathways are indeed conserved among species across different taxonomic levels and the integrated endpoint such as TELI was able to capture majority of the molecular responses that predict cytotoxicity in both prokaryotic and eukaryotic single-cell organisms.
Figure 6.3 Correlation between molecular endpoints derived from 2-hr toxicogenomics assay based on conserved cellular stress response pathway via biomarkers ensemble using GFP-recombinants E coli reporters strains, and phenotypic endpoints EC50 in green alga S.capricornutum 96 hours growth inhibition. Log scale are used for all concentration-based endpoints.
6.4.4 Cross species extrapolation - correlation of molecular endpoints based on stress response ensemble in E coli to EC50 with higher eukaryotic fish cell line RTgill W-1 cytotoxicity

Fish are especially important for ecotoxicity assessment since toxicants are often released first into aquatic environments by a variety of routes (Bols et al., 2005). As many biological systems have been preserved throughout evolution, effects on the fish can serve as a warning of possible impacts on human health and fish can serve as laboratory models for studying toxicants of concern to human health (Bols et al., 2005). Recently year, cytotoxicity tests using continuous fish cell lines have been suggested as an alternative tool in ecotoxicology. Gills are unique structure involved in respiration and osmoregulation in piscines as well as in many aquatic invertebrates. The availability of trout-derived gill cell line, RTgill-W1, and its ability to withstand hypo- and hyper- osmotic conditions and their optimal growth capacity at room temperature, make these cells ideal sentinel models for in vitro aquatic toxicology. Excellent correlation was observed between tests with rainbow trout whole fish lethality and the RTgill-W1 cell line cytotoxicity (Dayeh et al. 2002). Here, we explored the potential relationship between conserved molecular cellular stress response and the cytotoxicity in the rainbow trout gill cell line RTgill-W1 for cross-species extrapolation acute toxicity assessment.
Figure 6.4 Correlation between molecular endpoints derived from 2-hr toxicogenomics assay based on conserved cellular stress response pathway via biomarkers ensemble using GFP-recombinants E coli reporters strains, and phenotypic endpoints EC50 in on Fish Cell RTgill-W1 24hours cytotoxicity. Log scale are used for all concentration-based endpoints

No significantly correlations were observed between any of the TELI derived molecular endpoints in E coli and 24-hour cytotoxicity EC50 in RTgill-W1
(Figure 6.4). In addition, as shown in Table 6.1, EC50 measured in trout cell line RTgill-W1 did not significantly correlate with any of the phenotypic endpoints obtained in E.coli, V.fishi, and green alga S.capricornutum. The lack of correlation between RTgill-W1 fish cell cytotoxicity test and TELI derived endpoints could be resulted from the poor sensitivity of RTgill-W1 (Dayeh et al., 2002) and the differences in cell structures and cell functions among fish and other single cell organism. For example, number of chromosome, way for cell division, cell size, and presence of cellular structure, like nucleus, cell wall, mitochondria, chloroplasts, etc. For 19 of the 30 chemicals or mixtures tested in this study (63% of all chemicals), the fish cell exerted low sensitivity and they could not reach EC50 even at their highest soluble concentrations in the media. For those chemicals that can reach measureable EC50, the values do not differs a lot and they were within one order of magnitude range. Previous works also showed that RTgill W-1 appeared up to several orders of magnitude less sensitive than fish (Tanneberger et al., 2012). Similar conclusion has been made for Daphnia lethality bioassay, and there was no correlation between the Daphnia LC50 and the RTgill-W1 EC50 (Maltby, 2007).

We do recognize that the stress response biomarkers we used only reflect the effects at potential associated with cellular toxicity, other toxic impacts, such as apoptosis, receptor binding activity etc., that can be in high level eukaryotic cells such as fish cells that are not captured in our assay. Furthermore, system level impact beyond those can be observe in single cell such as endocrine disrupting, immunological, neurotoxicity could not be capture in our assay either. Therefore
it is not surprising to see the deviation and failure of the E coli based assay to predict fish cell toxicity.

**Table 6.1** Summary of correlation coefficients among variety of endpoints for phenotypic responses

<table>
<thead>
<tr>
<th>r/p-value</th>
<th>V.fisheri</th>
<th>Selenastrum</th>
<th>RTgill-W1</th>
</tr>
</thead>
<tbody>
<tr>
<td>E.coli growth inhibition EC50</td>
<td>0.71 (&lt;0.01)</td>
<td>0.77 (&lt;0.01)</td>
<td>0.33 (0.31)</td>
</tr>
<tr>
<td>V.fisheri luminescence inhibition EC50</td>
<td></td>
<td>0.79 (&lt;0.01)</td>
<td>-0.02</td>
</tr>
<tr>
<td>Selenastrum growth inhibition EC50</td>
<td></td>
<td></td>
<td>-0.06</td>
</tr>
</tbody>
</table>

**6.5 Conclusion**

Advances in high-through put in vitro mechanistic toxicity assay techniques addressed the challenges in need of timely and informative toxicological information for the large number and ever increasing CECs and for effective environmental monitoring, ecological risk assessment. In this study, we applied a toxicogenomics assay based on conserved cellular stress response pathy biomarkers ensemble using GFP-fuses recombinant E coli strains, which is fast, low-cost, therefore feasible for comprehensive toxicity assessment of a relative large number of CEC of various categories with limited resources and time. The assay measures and quantified molecular perturbations in known cellular stress response pathways that are involved in detecting stress and repairing damage caused by toxicants on primary cellular infrastructure. We further applied the newly developed molecular disturbance quantifier – TELI to establish dose response curves, which allowed derivation of molecular endpoints such as TELI\textsubscript{max}, EC-TEL\textsubscript{1.5}. We then, for the first time, explored phenotypic
anchoring of the derived molecular endpoints to phenotypic effects in a number of ecological-relevant single cell biomoniroting species (E.coli, V.fisheri, S.capricornutum) and fish cell line. Our results showed that the stress response pathways-based toxicogenomics assay can potentially capture the overall cellular toxicity potential and they can be quantitatively linked to conventional toxicity endpoints such as EC50 for single-cell micro-organisms across species and representing multiple taxonomic levels. Therefore, cellular stress response pathways offer a viable solution to the creation of a set of integrative assays as there is a limited and hence manageable set of major cellular stress response pathways to enable the development of high-throughput cell-based assays using the components of the pathways. And, the cellular stress response pathways are thus directly related to the phenotypic changes and adverse outcomes, which make them potential AOPs. Previous studies have also demonstrated quantitative correlation between cellular stress responses in DNA damage and repair pathways and phenotypic genotoxicity assay Comet assay (Lan et al., 2015).

These results demonstrated the potential possibility to adopt the propose stress-response pathway ensemble based quantitative toxicogenomics assay as a alternative standard methods for water toxicity assessment and environmental monitoring. Further investigation is warranted for extended study to include larger number and wide range of chemicals and environmental samples and required for explore the correlation of specific toxicity pathway and the underlying of correlation among more different species relevant to human and ecological risk assessment.
Chapter 7

Mechanistic Toxicity Assessment of Nanomaterials by Whole-cell-array Stress Genes Expression Analysis

7.1 Abstract

This study performed mechanistic toxicity assessment of nano-silver (nAg) and nano-titanium dioxide anatase (nTiO2_a) via toxicogenomic approach, employing a whole-cell-array library consisting of 91 recombinated E.coli K12 strains with transcriptional GFP-fusions covering most known stress response genes. The results, for the first time, revealed more detailed transcriptional information on the toxic mechanism of nAg and nTiO2_a, and led to a better understanding of the mode of action (MOA) of metal and metal oxide nanomaterials (NMs). The detailed pathways network established for the oxidative stress system and for the SOS (DNA damage) repair system based on the temporal gene expression profiling data revealed the relationships and sequences of key genes involved in these toxin response systems. Both NMs were found to cause oxidative stress as well as cell membrane and transportation damage. Genotoxicity and DNA damage were also observed, although nTiO2_a induced SOS response via previously identified pathway and nAg seemed to induce DNA repair via a pathway different from SOS. We observed that the NMs at lower concentration tend to induce more chemical-specific toxicity response, while at higher
concentrations, more general global stress response dominates. The information-rich real time gene expression data allowed for identification of potential biomarkers that can be employed for specific toxin detection and biosensor developments. The concentration-dependent gene expression response led to the determination of the no observable transcription effect level (NOTEL) values, which can be potentially applied in the regulatory and risk assessment framework as an alternative toxicity assessment endpoint.

7.2 Introduction

The currently available water toxicity assessment methods, such as WET (Whole Effluent Toxicity) and TIE (Toxicity Identification Evaluation), are labor-intensive and time-consuming, therefore are neither feasible nor sufficient for timely information needed for regulatory decision making to eliminate the potential risks posed by the ever-increasing number of emerging pollutants in our water. The growing production and use of engineered nanomaterials (NMs) makes it inevitable for them to release to the natural environment. However, the toxicity mechanism of these NMs is largely unknown even though there is sound evidence of their toxicity (Kang et al., 2009; Lanone et al., 2009).

Toxicogenomics, in which transcription and expression levels of thousands of genes in an organism in response to environmental toxicants are monitored (Snape et al., 2004; Ankley et al., 2006), promises a revolutionary new ground for monitoring and identifying the chemicals responsible for toxicity (Newton et al., 2004; Poynton et al., 2008), revealing the toxic mechanism and obtaining pollutant-specific molecular fingerprints (or biomarkers) for compound
classification and identification (Watson et al., 2004; Poynton et al., 2008), the No Observable Transcription Effect Level (NOTEL), a new concept generated with the development of toxicogenomics, can potentially serve as a more informative and sensitive endpoint for screening effluents and unknown chemicals for toxicity and be incorporated into ecological risk assessment and regulatory framework (Ankley et al., 2006; Poynton et al., 2008; Lobenhofer et al., 2004). A number of recent studies have demonstrated the application of toxicogenomics technology (e.g. microarray-based methods) for environmental monitoring (Williams et al., 2008; George et al., 2004). However, several limitations of the current microarray-based gene-profiling technology, such as a high-cost, complex procedure and condition-sensitive results, prohibit its further development and wide application in environmental monitoring.

We recently applied prokaryotic real time gene expression profiling by using a comprehensive cell array consisting of transcriptional green fluorescent protein (GFP)-fused recombinant *E. coli* strains for toxicants evaluation (Onnis-Hayden et al., 2009). Our proposed method, comparing to the microarray approach, requires simpler, faster, and more reliable assay procedures, has higher reusability, and provides the desirable flexibility for customization of the cell array. The substantial information available on the functions of the genes of *E. coli* allows for understanding, mapping, and visualizing systematic cellular response pathways and molecular events occurring as a response to chemical exposure.
In this study, we conducted mechanistic toxicity assessment of nano-silver (nAg) and nano-titanium dioxide anatase (nTiO$_{2-a}$) using the prokaryotic real time gene expression profiling method developed. The two NMs were chose because they have wide commercialization application in various fields and therefore most likely to present in the environment. Compound-specific and concentration-sensitive two-dimensional (genes and time) gene expression profiling for nAg and nTiO$_{2-a}$ at various concentrations were obtained. The gene expression alterations as the result of exposure to these two NMs provided insights into the underlying toxic mechanisms of these two NMs. The No Observed Transcriptional Level (NOTEL) for both NMs is determined based on the dose-response curve, and potentially biomarkers have been identified.

7.3 Materials and Method

7.3.1 Nanomaterials

nAg (~60nm, NanoDynamics Inc., Buffalo, NY, USA) and nTiO$_{2-a}$ (10nm, NanoStructured & Amorphous Materials, Houston, Texas, USA) were prepared in M9 medium for a stock concentration of 1000 mg/L, which contains 1 % of crude Bovine Serum Albumin (BSA) as dispersant (Bello et al., 2009). The stock solutions were sonicated in a High energy Cup-sonicator, ~90Watt power for at least 15 mins to maintain a better dispersion before the toxicity assays.

7.3.2 Prokaryotic stress genes cell-library with transcriptional GFP fusions

A library of transcriptional fusions of GFP (Open Biosystem, Huntsville, AL, USA) that includes 91 different promoters controlling the expression of genes
associated with the most known stress responses and other specific function in 
*E. coli K12*, MG1655 was employed in this study. Each promoter fusion is expressed from a low-copy plasmid, pUA66 or pUA139, which contain a kanamycin resistance gene and a fast folding *gfpmut2* that enables measurement of gene expression at a resolution of minutes with high accuracy and reproducibility (Zaslaver et al., 2006). Each category of stress genes and their main functions are briefly described in our previous study (Onnis-Hayden et al., 2009).

7.3.3 *Measuring the temporal gene expression upon chemical exposure*

Cells were grown in black 96-well plates (Costar, Bethesda, MD, USA) for 2 hours at 37 °C until the cultures reached early exponential growth (OD600 about 0.05-0.1), NMs stock solution was added per well to obtain a final concentration of 1, 10, 50mg/L for nAg and nTiO$_2$-a, respectively. Then the plate was put in a Microplate Reader (SynergyTM HT Multi-Mode, Biotech, Winooski, VT, USA) for simultaneous absorbance (OD600) measurement (cell growth) and fluorescence readings (GFP level, filters 485nm, 528nm) at a time interval of 3 minutes. More detailed description of methods is available in our previous study (Onnis-Hayden et al., 2009).

7.3.4 *Data processing and analysis*

The alteration in gene expression, also called induction factor I ($I=\frac{P_e}{P_c}$), for a given gene at each time point due to chemical exposure, was represented by the ratio of the normalized gene expression GFP level ($Pe=(GFP/OD)_{\text{experiment}}$) in the
experiments with NMs exposure to that \( \text{Pc} = (\text{GFP/OD})_{\text{control}} \) in the control condition without any NMs exposure. Then the natural log of I value (lnI) at every time point was compiled for hierarchical clustering (HCL) analysis that is done by MultiExperiment View (MeV) version 4.4. Detailed information of data process and clustering analysis are available in previous chapters.

**Gene Expression Pathway Analysis:**

We combined our real-time gene expression data with the known pathway and function of genes on EcoCyc database (http://ecocyc.org/) and the Gene Ontology Database (http://www.geneontology.org/) to obtain insights and network of the specific genetic pathway upon NMs exposure (Keseler et al., 2009; Ashburner et al., 2000).

**Determination of No Observable Transcription Effect Level (NOTE):**

We applied the concept of NOTE as the maximum concentration of a chemical at which less than 5% of the genes are differentially expressed upon chemical exposure compared to control (Poynton et al., 2008; Lobenhofer et al., 2004). We fit a dose-response curve with the percentage of genes expressed in our “stress library” to the concentrations of chemicals, using a generalized linear model with binomial family. The NOTE is determined, based on the dose-response curve.
7.4 Results and Discussion

7.4.1 Distinctive and Complex Real-time gene-expression profiles for nAg and nTiO$_2$-a

Distinctive temporal gene expression profiles were obtained for nAg and nTiO$_2$-a at three different concentrations and over a 2-hours period (Figure 7.1, Appendix 7.1 and Appendix 7.2). The results showed very dynamic and complex toxin-induced real-time gene expressions across the stress genes examined, with most of the genes exhibiting different patterns and varying magnitudes of transcription activities over time. The temporal change in gene expression level reflected the dynamic of the cellular response system and the time sequence for a particular set of gene to be involved, which may depend on the system-level multiple gene activation, signaling pathways and the roles and time-sequence in which they are involved in the stress-response mechanism. (Exemplary temporal gene expression profiles for several genes are show in Appendix 7.3). These temporal gene expression profiling are compound-specific and concentration-dependent and even the slightest molecular alterations in the cellular system response as result of variation in the concentration of the same toxin, were captured and reflected (Figure 7.1 and Appendix 7.1). This validates and highlights the advantage of our approach using real-time gene expression to gain temporal resolution, in contrast to a possibly biased “snapshot” of the dynamic expression profiles at an arbitrarily selected time point.
Figure 7.1 Real-time (temporal) gene expression profiles of 91 stress genes in *E. coli* in exposure to nAg 10mg/L (left) 50mg/L (right). X-axis top: natural log of induction factor (lnI). (Red spectrum colors indicate up-regulation, green spectrum colors indicate down-regulation) and time in minutes (the first data point shown is at 15 minutes after exposure due to moving average), Y-axis left: clustering of the profiles, Y-axis-right: list of genes color-coded based on functional categorization (nTiO$_2$-a are shown in Appendix 7.1).
7.4.2 Concentration-dependent expression profiles

The specific genes and the number of genes differentially expressed upon exposure to the same NM varied with different concentrations. Figure 7.2A illustrated the number of altered genes that are specific to each concentration as well as the numbers of differentially expressed genes that are common to treatments with different NM concentrations. For both nAg and nTiO$_2$-a, only a few genes showed altered expression compared to control at the lowest concentration (1mg/L) studied, suggesting the threshold of observable molecular alterations at this concentration. For nAg at 1 mg/L, the three toxin-induced genes are completely different from those induced at higher concentration. Many genes related to general stress, cell killing, cold shock and energy stress were only differentially expressed at higher nAg concentration of 50mg/L. For nTiO$_2$-a, 4 genes showed altered expression at 1 mg/L, including gene tam (detoxification) that exhibited over-expression for all three concentrations. Similar conclusions have also been reported following Cu exposure with Daphnia magna and arsenic exposure in a human lung cell line (Poynton et al., 2008; Andrew et al., 2003), in which the profile obtained at low concentration was distinct from that at a higher concentration. Vulpe et. al. mentioned that the specificity of genes in response to a given toxin decreases as the chemical concentrations increases (Poynton et al., 2008). This maybe because that at higher concentrations, more genes related to general stress and cell killing become more prevalent as the cell overall integrity becomes compromised and the specific mode of toxicity would be overshadowed by a common global system-stress response.
The gene expression profiles seem to be not only chemical-specific but also concentration-dependent (Figure 7.2B), indicating that the molecular level genomic activities are very sensitive to not only the type of toxin but also the level of toxin in exposure. For both nAg and nTiO$_2$-a, the number (percentage) of genes with altered expression level increased as the chemical concentrations increased (Figure 7.2B). This relationship allowed us to obtain the dose-response curve using a generalized linear model (Poynton et al., 2008). For nTiO$_2$-a, concentrations at both 10 mg/L and 50mg/L induced more genes than nAg, but the induction level are less than those of nAg at the similar concentrations, suggesting lower toxicity level than nAg. This seems to agree with the NMs toxicity level assessment via Biological Oxidative Damage (BOD) proposed by Bello et. al (Bello et al., 2009).
Figure 7.2 Concentration-dependent responses to NMs. (A) Number of differential expressed genes at each exposure concentration. The genes differentially expressed in more than one concentration are shown in the overlapping part, nAg (left), nTiO$_2$-a (right). (B) Dose-response curve based on the portion of differentially expressed genes upon exposure to NMs in our stress...
library as a function of NMs concentrations. (Data points are represented by the annotation of NMs.)

7.4.3 Determination of No Observed Transcriptional Effect Level (NOTEL)

This concept of NOTEL was proposed by Lobenhofer et al (Lobenhofer et al., 2004) and its potential application in toxicology and risk assessment has been demonstrated and discussed by Ankley and Poynton et. al. (Ankley et al., 2006; Poynton et al., 2008). NOTEL can be potentially used as an endpoint and regulatory benchmark for chemical screening, whole effluent toxicity testing and environmental monitoring of toxicants, similar to the endpoints of traditional toxicity assessment such as EC\textsubscript{50} and LC\textsubscript{50}.

The dose-response curve that related the percentage of differentially expressed genes to the NMs concentration (Figure 7.2B) allowed for the determination of NOTEL (Poynton et al., 2008). In our study, the NOTEL value based on our stress gene library was determined to be 0.658 (± 0.260) mg/L for nAg, and 0.557 (± 0.545) mg/L for nTiO\textsubscript{2}_a. The results indicated that these two NMs seemed to have similar magnitude of toxicity, which is consistent with the Biological Oxidative Damage (BOD) values for nAg and nTiO\textsubscript{2}_a recently proposed as 116.4\textmu mol-TEU (Trolox Equivalent Units)/L and 64.9\textmu mol-TEU/L, respectively (Bello et al., 2009). The NOTEL is expected to be lower and more sensitive than conventional endpoints (e.g. EC50) since it reflects sub-lethal and molecular level response to toxins.
7.4.4 Toxicity mechanism of nAg

The toxic mechanism of metal and metal oxidant NMs are still not fully understood and current understanding indicates that the prominent toxic mode of action involves production of Reactive Oxygen Species (ROS), which can damage DNA, RNA and proteins, including a multitude of oxidized base lesions, abasic sites, single and double-strand breaks, all of these can be cytotoxic and mutagenic (Diakowska et al., 2007). Our results showed that nAg exposure induced many genes that belong to detoxification, SOS response, oxidative/redox stress, drug resistance/sensitivity and protein stress (Figure 7.3). In addition, many of the genes related with inner membrane and transport system, such as cmr, fsr, yajR, and emrE, were up-regulated in exposure to nAg.

Based on our temporal gene expression results and previous pathway framework (Keseler et al., 2009; Ashburner et al., 2000), a more detailed network maps for oxidative damage system was developed as shown in Figure 7.4. They demonstrated the interrelationship among key genes and their temporarily dynamic gene expression in oxide stress regulatory networks upon exposure to nAg. Oxidative damage response system includes genes related to redox response, detoxification and drug resistance. Redox response are regulated by the two redox regulators, namely oxyR and soxR. Gene oxyR was up-regulated at an increasing level over time upon exposure to nAg, indicating the induction of oxidative damage response (Figure 4). OxyR serves as the transcriptional dual regulator to those involved in peroxide metabolism, peroxide protection and redox balance, which involved sodA, sodB, sodC, katE and katG (Keseler et al., 2009). The
protein products of genes sodA, sodB and sodC are all subunit of superoxide dismutase (Alscher et al., 2002), which involve in the degradation of superoxide to hydrogen peroxide, then subsequently hydrogen peroxide transformation into oxygen catalyzed by hydroperoxidase KatG and KatE. The up-regulation of sodA, sodC and katE in our result indicated superoxide radical and hydrogen peroxide generated upon nAg exposure. SoxR, which is known to control the transcription of the regulator involved in the responses against nitric oxide toxicity (Tsaneva et al., 1990), was found to be up-regulated after 30 minutes exposure. Activation of soxR, induces soxS expression, and soxS, in turn, activates transcription and participates in controlling several genes involved in oxidative stress, resistance to antibiotics, organic solvents, and heavy metals, such as nfo, inaA and marR (Keseler et al., 2009). The up-regulation of these genes initiated after 60 mins and the alteration magnitude increased over time upon exposure to nAg (Figure 7.4). The expression level of oxyR remained increasing for 90 minutes then declined after that (Appendix 7.3), suggesting that at the later stage the over production of OxyR protein functions as a repressor for the gene oxyR (Tao et al., 1991).
Figure 7.3 Toxicity mechanisms revealed by the altered gene expression level for genes involved in different stress functional categories in exposure to nAg (10 mg/L). (A) 90% confidence interval of maximum altered gene expression level (Ln(I)) for genes involved in each stress functional category. In our study, we chose Ln (I)=0.4 as the noise cut-off baseline. (B) Distribution of altered gene
expression level (Ln(I)) among the genes involved in each stress functional category.

**Figure 7.4** Oxidative stress response regulator network upon exposure to nAg based on temporally dynamic gene expressions obtained in our study. Black line with arrow indicates the activation/induction due to nAg exposure. The rectangular bar shows the altered gene expression level over time.

Genes related to outer membrane oxidase electron transfer, including *bolA*, *cyoA*, *sdhC*, and *tam*, were also up-regulated upon exposure to nAg (Figure 7.4). Overexpression of *bolA* induces biofilm formation, and alters the properties of outer membrane (Vieira et al., 2004). SdhC is one of two membrane subunits of succinate dehydrogenase, which participates in both the citric acid cycle and the electron transfer chain (Oyedotun et al., 2004).

Most drug resistance genes induced by nAg (at 10mg/L) are Major Facilitator Superfamily (MFS). The MFS transporters are single-polypeptide secondary carriers capable only of transporting small solutes in response to chemiosmotic
ion gradients. *cmr*, whose product also known as MdfA, is a multidrug transporter that is driven by the proton electrochemical gradient (Edgar et al., 1997) and it showed 3.5 folds over-expression in our results. *emrE*, encoding for a multidrug efflux protein with a broad substrate specificity, exhibited up-regulation by 3 folds. In addition, *fsr, yajR*, which were over expressed at 10mg/L nAg exposure, belong to the drug efflux systems of MFS. However, none of these genes showed over-expression at 50mg/L nAg exposure, indicating again that at higher toxin concentration, the entire integrity of bacteria cell is likely compromised such that it lost its specific anti-toxin ability. Or, alternatively, bacteria activate overall global-stress response system at a higher concentration, as discussed previously.

Genes involved in DNA repair, such as *recN, uvrA, ybfE, yebG, ssb, sbmc* and *nfo*, were up-regulated 30 mins after being exposed to nAg. RecN is required for DNA repair when breaks occur at two or more locations. Over-expression of *uvrA*, a subunit of UvrABC nucleotide DNA excision repair complex, indicated its active role in repairing a wide diversity of lesions. The product of the essential *ssb* gene, interacts directly with DNA polymerase II, exonuclease I, and replication protein n, which are all involved in DNA metabolism (Meyer et al., 1990). Endonuclease Nfo catalyze the formation of single-stand breaks in double apurinic (AP) DNA to remove the damaged base in DNA repair. Interestingly, the two key up-stream regulating genes of SOS system, *recA* and *lexA*, showed no observed differential expression within the two hours period. Previous studies of SOS system generally assume that genes recA and lexA should be up-regulated when SOS system is turned on (Yamanaka et al., 2004; McCool et al., 2004). However, not all SOS
genes are regulated by *lexA*, and *recA* are not involved in every physiological DNA repair pathway (WALKER et al., 1984). So the involvement and activity of SOS response genes to different genotoxicity may require further investigation. Nevertheless, altered promoter activity of a number of DNA repair related genes discussed above suggested that nAg causes DNA damage.

### 7.4.5 Toxicity mechanism of nTiO$_2$-a

The genes differentially expressed upon exposure of nTiO$_2$-a are mostly dominated by those belonging to drug resistance/sensitivity, detoxification, DNA damage and protein stress categories (Appendix 7.2). About one third of the differential expressed genes induced by nTiO$_2$-a (10mg/L) are the same as those induced by nAg (10mg/L), including genes in drug resistance/sensitivity, DNA damage categories, suggesting that both NMs cause oxidative stress and DNA damage. However, about two thirds of the altered genes in exposure to nTiO$_2$-a were different for the two NMs, indicating distinctive transcription level stress response these two NMs induce. More than half of the drug resistance/sensitivity genes induced by nTiO$_2$-a (10mg/L) belongs to MFS, some of them are the same genes with nAg (10mg/L). However, at concentration of 50mg/L, less than one fourth drug resistance/sensitivity genes belong to MFS. Cell membrane transportation related genes including *cmr*, *yajr* and *emrE*, and *dnaK*, which are involved in the cytoplasmic cellular processes such as protein folding and protein translocation through membranes, were also up-regulated in exposure to nTiO$_2$-a at 10 mg/L. *sanA*, which involves in cell envelope barrier functions, was up-regualted in exposure to nTiO$_2$-a (10mg/L), indicating outer membrane
permeability defects (Rida et al., 1996). These suggested that nTiO$_2$ a damaged the permeability of cell membrane, which was also evidenced by our recent study that observed compromised cell membrane and surface (data not published). 

AphF, functioned with AphC to catalyze alklydroperoxide into alcohol with the participation of NAD(P)H, different from the hydroperoxidase KatE (up-regulated in nAg), a key step in superoxide redicals degradation, convert hydrogen peroxide directly into water and oxide (Bieger et al., 2000). As gene aphF showed 2 folds over-expression upon nTiO$_2$ a (10mg/L) but no expression with nAg exposure (10 and 50 mg.L), which may indicate that nTiO$_2$ a induced a different peroxide degradation pathway compared to nAg.

nTiO$_2$ a(10mg/L) caused DNA damage and it activated the DNA repair SOS system, as indicated by the up-regulations of the two key up-stream regulons, recA and lexA., which showed 1.5-fold and 3-fold over-expression, respectively. Their activation led to the up-regulations of a number of DNA damage related genes, including recN, mutT, nfo, uvrA, uvrD, umuD, polB and ssb. DNA polymerase II (Pol II), the product of gene polB, is a combined polymerase and exonuclease involved in translesion synthesis and nucleotide excision repair, and it also plays a role in avoiding the mutagenic effects of agents such as peroxide (Sedliaková et al., 2001).

7.4.6 Candidate biomarkers for nAg and nTiO$_2$ a exposure

Compound–specific signature gene expression profile offers the possibility to uncover novel biomarkers of exposure and predict the presence of a class of contaminants (Poynton et al., 2008; Kim et al., 2006), especially the emerging
contaminants such as NMs for which biomarkers are not presently available. Based on the gene expression results, we can propose some genes as potential biomarkers based on the following guidelines: a) genes that show chemical-specific and concentration-dependent expression pattern; b) genes that are more related to MOA of a contaminant rather than those general stress or function genes. We screened and selected genes that seem to show significant alteration in their expression level and exhibited concentration-dependent patterns as candidate biomarkers for nAg and n_TiO₂_a exposure. *oxyR, cls and cspB* are shown to be the potential bio-markers for nAg exposure and *mutT, sodB, pbpG* are the three potential bio-markers for n_TiO₂_a exposure (Appendix 7.3). Of course, the specificity, sensitivity and reliability of these suggested candidate biomarkers require further investigation and evaluation.

**7.4 Conclusion**

In summary, our results, for the first time, revealed more detailed transcriptional information on the toxic mechanism of nAg and n_TiO₂_a and led to a better understanding of the MOA of metal and metal oxide NMs. Both NMs were found to cause oxidative stress as well as cell membrane and transportation damage. But the difference in the specific genes with altered expression upon the exposure to the two NMs, suggested the different toxicity mechanisms. For example, different hydrogen peroxide reduction pathway seemed to be involved in the oxidative stress response, as nAg causes damage to electron transfer pathway, whereas, n_TiO₂_a damages membrane permeability. Both n_TiO₂_a and nAg cause DNA damage, however, differences were observed in the genes and their magnitude of
alterations involved in the SOS system. In addition, we observed that the NMs at lower concentration tend to induce more chemical-specific toxicity response, while at higher concentrations, more general global response dominates. The information-rich real time gene expression results allowed for identification of potential biomarkers that can be employed for specific toxin detection and monitoring applications. The NOTEL values determined for both nAg and nTiO$_2$-a seem to be consistent with other established toxicity assessment endpoints, therefore can be potentially used as regulatory benchmarks and for toxicity screening. Our results demonstrated that real time gene expression profiling yields compound-specific and concentration-sensitive multidimensional “fingerprints” specific to each compound with variables of gene, time and concentration and therefore can be applied as a feasible method for toxicity assessment and screening of a large number of emerging contaminants.
Chapter 8

Application of Quantitative Toxicogenomics Assay for
Wastewater Water Quality Monitoring and Treatment

Efficacy Assessment

8.1 Introduction

The Clean Water Act (CWA) (United States Federal Water Pollution Control Act Amendments, Public Law 92-500 of 1972) prohibits the discharge of “toxic pollutants in toxic amounts” (“The Clean Water Act”, 1972). Until recently, regulation of wastewater treatment plant toxic discharges into receiving waters focused primarily on specific chemical constituents (Schwarzenbach et al., 2006; Sedlak et al., 2005; Snyder et al., 2008). The National Pollution Discharge Elimination System (NPDES) permit program was predicated on achieving technology-based control levels represented best available treatment (BAT), considering technical and economic feasibility (“National Pollutant Discharge Elimination System”). However, it has been known for some time that this approach would not ensure that wastewater discharges would not produce adverse biological impacts in receiving waters (Brooks et al., 2006; Lapworth et al., 2012; “Emerging Contaminants In the Environment”, 2014).

Research by the U.S. Environmental Protection Agency (EPA) has shown that biological problems, associated with point source discharges of substances in toxic concentrations, are related to extensive site-specific factors (“USGS Fact Sheet FS-027-
02”, 2002; Petrović et al., 2003). These include the quantities, types, and sources of toxic materials; the control effectiveness of existing wastewater treatment plants, and the characteristics and dilution capacity of the receiving water.

In 1984 the EPA issued a new national policy that recommended the combined use of biological and chemical data to control toxic substance discharge to receiving waters (“Fed. Reg. 49, 46”, 1984). In 1985 a water-quality-based technical support document was published to support the policy (USEPA, 1985). It included procedures for evaluating toxicity problems in receiving water for locating and identifying the responsible point source discharge. Numerous field studies have been conducted to validate the procedures and use of the new, short-term chronic toxicity test methods has been recommended (Mount et al., 1984; Norberg-King and Mount, 1986; Heber, 1994; Weber, 1991).

However, the challenges in the required resources and time to handle the toxicity testing efforts for the large and ever-increasing number of samples have motivated a new vision of toxicity testing strategy as proposed by NRC (National research council) and USEPA (NRC, 2007). It calls for a systematic transit from current resources-intensive and time-consuming in vivo whole animal-based testing to in vitro mechanistic toxicity pathway-based assays on cell lines, using cost-effective, reliable and high-throughput screening and tier testing (NRC, 2007). Single or battery of in vitro cell-based bioassays using cell lines or biomarkers has been successfully applied to environmental water samples (Neale et al., 2012; Trovo et al., 2010; Macova et al., 2011; Escher et al., 2011). These in vitro bioassays are more sensitive towards early warning signs since cellular responses are activated before the actual harm occurs, and they indicate the presence of
associated stressors. However, these batteries of bioassays consisting of a number of separated assays with different species still require a substantial amount of time and efforts. In addition, most of them only provide information for specific mode of action without more comprehensive and detailed information of the toxicity profiles and mechanisms. Furthermore, isolated assays that cover only one or a few biomarkers or specific toxicity mechanisms cannot be directly anchored to an integrated adverse outcome or phenotypic endpoints.

Recently, we reported a toxicogenomics-based toxicity assessment method that allows a fast, yet informative, mechanistic and quantitative toxicity evaluation of toxicants using transcriptomics technology with GFP-fused whole-cell array of *E.coli* K12, MG1655 (Gou et al., 2010; Gou and Gu, 2011). Compared with the traditional cytotoxicity approach, or the current battery of bioassays approach, this quantitative toxicogenomics method in combination with bioinformatics computation detects the overall toxicity response, and reveals potential toxicity profiles and mechanisms with one single assay within hours (2-4 hours), which greatly reduces the test time, resource and cost. In addition, the measurement is sensitive enough to capture the sub-cytotoxic impact at concentrations much lower than those that can lead to detectable phenotype effects, therefore it requires less or even possibly eliminates the sample extraction and concentration step. The proposed method may serve as an alternative or complementary approach to the current toxicity assays for environmental applications.

In this study, we, for the first time, demonstrated the application of a quantitative toxicogenomics-based approach for evaluation of the toxicity evolution and nature along each treatment processes in two wastewater treatment plants. Toxicity level and profile
evolution along each treatment processes were monitored to reveal the dynamic toxicity changes and mechanisms, as well as their association with treatment effectiveness assessment.

8.2 Materials and Methods

Wastewater Samples.

The two sampled treatment plants were: Deer Island Treatment Plant at Boston and North Attleborough Wastewater Treatment Facility. A list of the plants with their wastewater sources, location, and a brief description of the treatment process is given in Table 8.1.

Table 8.1 Information of Wastewater Treatment Plant

<table>
<thead>
<tr>
<th>WWTP</th>
<th>Influent source</th>
<th>Serve location</th>
<th>Capacity</th>
<th>Treatment</th>
<th>Discharge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deer Island</td>
<td>Business, industrial and resident wastewater</td>
<td>43 greater Boston communities</td>
<td>1270 MGD</td>
<td>Secondary</td>
<td>Ocean</td>
</tr>
<tr>
<td>North Attleborough</td>
<td>Industrial and domestic wastewater</td>
<td>North Attleborough as well as Plainville</td>
<td>3.1 MGD</td>
<td>Tertiary</td>
<td>Surface water</td>
</tr>
</tbody>
</table>

Sample Preparation for Toxicity Assessments.

The influents and effluents, as well as samples taken after every major treatment steps were subjected to the toxicity assessment. All samples were neutralized to pH around 7 prior to enrichment by lyophilization (Freezone 4.5, LABCONCO, Kansas City, MO) under -42°C with 0.1 Torr vacuum. The lyophilization procedure was used since it has been reported to have higher recovery rates of non-volatile compounds than solid phase extraction (SPE).
Microtox Assay.

Photobacterial *Vibrio fisheri* (ATCC #49387) specific photosynthesis inhibition test (Microtox) was used for nonspecific toxicity evaluation according to the ISO standard method 11348-2 (ISO11348-2, 2007). The assay was performed in solid black 96-well microplate (Costar, Bethesda, MD, USA). Briefly, fresh prepared *V. fisheri* culture was equally mixed with desired water sample, and the luminescence inhibition was calculated by the difference between light intensity at time 0 and 15 minutes measured with a microplate reader (Synergy Multi-Mode, Biotek, Winooski, VT). 1 M NaCl solution was used as negative control and ZnSO$_4$ was used as a positive control. The result from Microtox tests were indicated as the inverse of EC50, which was the concentration causing a 50% photosynthesis inhibition effect with the concentration reported here in units of relative enrichment factor (REF), REF is the product of enrichment factor of lyophilization process and dilution factor of bioassay.

Green alga selenastrum growth inhibition test

*Selenastrum capricornutum* (UTEX 1648) was purchased from UTEX Culture Collection of Algae, Austin, TX, USA, and was maintained at 22±1°C with a 12 h light, 12 h dark with mix of cool-white/warm-white fluorescent lamps.

The algal growth inhibition assay was performed according to the ISO and U.S. EPA method (1989) (ISO 8692, 2004). Cultures were prepared in 100-mL Erlenmeyer flasks containing 20 mL of the filter-sterilized (0.22um) test medium in triplicate. Test samples were dissolved in growth medium and added to each culture. Each sample was tested in a dilution series of at least 4 concentrations. Both controls and test flasks were
inoculated with exponentially growing algae at an initial concentration of $S.\ capricornutum$ of $1 \times 10^4$ cells/mL. The cultures were incubated at a temperature of 22±1°C, and shaken at manually twice per day under a periodical (16 h light, 8 h dark) illumination of 4000 ±400 lux. After 96 h cell counts were determined using a microscope and a hemocytometer. The median inhibition concentration (IC50) values were calculated by the linear interpolation method (U.S. EPA, 1989). The concentrations of test chemicals were not analyzed; therefore, their nominal concentration was used as the exposure concentration in the calculation of IC50 values.

**Toxicogenomics Assay and Data Computation.**

A high-throughput mechanistic toxicity assay method was employed that uses with GFP-fused whole-cell array of *E.coli* K12, MG1655 (Zaslaver et al., 2006; Onnis-Hayden et al., 2009). The water samples (after freeze dried) were re-suspended in minimum media (M9) containing *E.coli* culture strains to various desired concentrations. The selected assay library covered a variety of genes involved in different known cellular stress response pathways, such as oxidative stress, DNA stress, protein stress, and membrane stress (including membrane transporter, efflux pump, energy metabolism, flagella metabolism, lipopolysaccharide metabolism), etc.

The protocol to measure the temporal gene expression profile was described in our previous reports. In brief, *E.coli* was cultivated in 384-well plates (Costar, Bethesda, MD, USA) in dark to avoid GFP photobleaching until the early exponential growth stage is reached (OD600 ~ 0.2). After addition of the aqueous sample, the plate was placed into a micro-plate reader (Synergy Multi-Mode, Biotek, Winooski, VT) for cell growth measurement (absorbance, 600 nm) and fluorescent reading (Excitation: 485 nm,
Emission: 528 nm) every 5 minutes over a period of 2 hours. Three biological replicates were performed for each condition. To quantify the magnitude of the altered gene expression level upon the exposure to toxicants, the induction factor I was determined for a given gene at each time point upon chemical exposure by normalized no-toxic exposure control. The toxicity mechanisms were elucidated by identifying and analyzing the changes in genes associated with specific stress response pathways. The molecular endpoint was quantified by the TELI (Transcriptional Effect Level Index) value, which was a recently developed index for toxicogenomics data interpretation that incorporates the number, magnitude and the temporal pattern of genes with an altered expression. TELI values can be determined for a single gene, or for a number of genes representing a specific pathway or the entire stress response assays. The specific or overall toxicity obtained by the toxicogenomics-based approach was represented by the corresponding REF/concentration that causes the TELI value to reach at 1.5 (EC-TEL1.5), similar to the approach that has been applied on \textit{umuC} genotoxicity assay by Escher \textit{et al.}

\textbf{Gene Enrichment Analysis}

The gene set enrichment analyses (GSEA) were achieved by ranking a list of genes with the TELI values according to the work of Subramanian (Subramanian \textit{et al.}, 2005). For each pathway, GSEA calculated the enrichment score by examining the ranked gene list from high score end to the low score end, giving a rewarding score if a gene belongs to the pathway of interest, and penalizing score otherwise. The significance of each pathway was determined by comparing their ranking scores to the corresponding empirical distributions. The null distributions were generated by randomly permuting the
specific pathway and all others 1,000 times. Only those categories with $p$-value less than 0.05 were considered to be altered significantly.

8.3 Results and Discussions

8.3.1 Excellent plant performance for both plants.

The average concentrations for selected conventional pollutants are given in Table 8.2. Total 5-day biochemical oxygen demand (BOD) for both plants were below 18 mg/L in the effluent, and the percent removals were greater than 90%. Results for dissolved organic carbon (DOC) also indicated good performance for both plants. The TOC concentration for both plant effluents were 9 mg/L or less. Removals for all plants were greater than 80%. The removal of total BOD, and DOC were similar to other POTWs surveyed national by the EPA. Overall, the plant operations for conventional performance ranged from satisfactory to excellent for BOD, and DOC.

![Graph](image)

**Figure 8.1** Performance data for two wastewater treatment plants. Dashed lines indicate data missing for corresponding processes.
8.3.2 Temporal Toxicity Evolution and Profile Changes during the CECs Degradation.

Figure 8.2 and Figure 8.3 showed the toxicity shift along the wastewater treatment process for both plants based on the toxicogenomics assay endpoints and the two conventional toxicity test results, respectively. The chronic toxic unit was calculated as 1/EC10 with green alga selenastrum 96 hours growth inhibition test. And the acute toxic unit was calculated 1/EC50 with V.fisheri 15 minutes luminescence inhibition test. Both conventional toxicity tested showed significantly toxicity reduction after primary and secondary treatment for Deer Island WWTP, while toxicity do not decrease, even arise for the disinfection process. Similar results was also exhibited for North Attleborough WWTPs samples, chronic toxicity unit from green alga tests did not differentiate the toxicity of the process after secondary treatment. In comparison, the more sensitive toxicogenomics assay clearly revealed the toxicity reduction trends during the oxidation transformation process, especially for samples from North Attleborough WWTPs.

Figure 8.2 Toxicity changes along the WWTP processes for both plants. Y axis left: the chronic toxic unit as 1/EC10 with green alga selenastrum 96 hours growth inhibition test. Y axis right: the acute toxic unit as 1/EC50 with V.fisheri 15 minutes luminescence inhibition test. DI water is used as blank control for enrichment process.
To quantitatively illustrate the nature of toxicity changes during the treatment processes with toxiciogenomics approach, molecular endpoint 1/EC-TELI1.5 for individual stress response category were calculated and plotted in Figure 3.

For both plants, the general trend of 3 toxicity tests are consistent, all showed significantly reduction of toxicity along the wastewater treatment processes, similar to the reduction of BOD and DOC. Membrane stress seems to be the major stress response category that were evoked by both influent samples, while this major response shifted to general stress response after primary treatment for North Attleborough WWTP.

Disinfection process cause different effects on the toxicity of final effluents. DNA stress response for Deer Island WWTP was raised while all stress responses were reduced for North Attleborough. The raise of DNA stress response of Deer Island WWTP may be contributed by the disinfection byproducts produced with excess organic matters present in secondary effluent. Therefore, the toxicity reduction of North Attleborough WWTP is more complete (1/EC-TELI1.5 values, which is a cellular benchmark dose, less than 1 (the original concentration) is considered non-toxic level according to the USEPA WET and ISO water quality methods) than that of Deer island WWTP since oxidative, membrane and DNA stress are still induced at the final effluent. This seems to be corresponding to the degree of advanced treatments. System design, optimization and operation should incorporate a treatment efficacy assessment to ensure desired toxicity and risk reduction.
Figure 8.3 Toxicity changes along the WWTP processes for both plants. Y axis left:

1/EC-TELI1.5 is the inverse of the relative enrichment factor (REF) that yields TELI value of 1.5, which was determined from dose-response curves.

8.3.3 Toxicogenomics Assay Reveals the Nature and Dynamics of Toxicity Evolution during the Wastewater Treatment Process

The spatial and temporal toxicity profile changes along the wastewater treatment processes were shown in Figure 8.4. The magnitude of gene expression changes was indicated by the absolute TELI value of each category within 2 hours exposure (that considers both up- and down-regulations as altered expression). These profiles revealed the dynamic nature of toxicity resulted from the original influents as well as those from a mixture of transformation products during the course of the treatment processes.

Different with Figure 8.3, Figure 8.4 showed that strong oxidative stress response significantly induced by the influent for Deer Island WWTPs at original concentration (REF=1). The difference between Figure 3 and Figure 4 indicate that more concentrations rather than the original should be considered for water safety management and risk assessment. All of the stress response reduced significantly after primary treatment. These great reductions of toxicity during primary treatment may indicate the removal of
toxic agents, which can penetrate 0.42um filter paper. However, most of the stress responses do not shift a lot after secondary treatment and disinfection process, which may indicate the difficulties of residue toxic agents removal or production of toxic byproducts. Consistently with Figure 8.3, general stress and DNA stress are still significantly expressed for final effluent when compared to blank control, indicate incomplete toxicity reduction with lower discharge limit to ocean.

For North Attleborough WWTP, similar as shown in Figure 8.3, membrane stress and general stress were induced by influent, and the toxicity do not reduced after primary treatment, indicates different influent composition compared to that of Deer Island WWTP. The toxicity of final effluent is almost completely eliminated, especially the general and DNA stress response, indicate the effectiveness of tertiary treatment on residue toxic agents removal and control of DBP productions.

**Figure 8.4** Exemplary temporal toxicity evolution and profile changes along the WWTP processes for both plants as well as the enrichment blank control (at REF=1, the original concentration). Profiles for other concentrations (REF) are not shown. X-axis top: list of
stress response categories. Altered gene expression changes in relative to the untreated control were indicated as the TELI value for each individual gene and color-coded with the scale 1-3 (Red spectrum colors indicate the magnitude of the altered gene expression, while black indicates no change).

8.4 Conclusions

Water-quality-based permit limitations for toxic substances and conventional pollutants require practical assessment and monitoring approaches to support the permitting process. Two monitoring approaches can be used; specific toxic chemical measurement, and overall toxicity measurement as determined by toxicity tests. To effectively control toxic substance discharge, an additional step, called a toxicity reduction evaluation (TRE), is performed at individual municipal or industrial wastewater treatment plants. This work, for the first time, utilized a quantitative toxicogenomics approach as an integrated TRE monitoring approach that identifies the specific toxic substances and toxicity, the pass-through variability of the overall toxicity in the effluent wastewaters, and finally the probable solution for controlling the toxicity discharge.

The toxicity as revealed by variety stress response reduced along the treatment process for both WWTPs, however, the toxicity removal effectiveness differs for 2 facilities, as well as individual process. Previous studies reported similar changes in chronic, phenotype endpoints, but with no or limited insights into the temporal dynamics of the nature of toxicity mechanisms associated with transformation products evolved during the degradation. In comparison, the high throughput mechanistic toxicogenomics-
based toxicity assay used in this study enables a more comprehensive and detailed toxicity evaluation during the treatment processes, thereby improving the understanding of underlying toxicity changes during a treatment process. Therefore, the quantitative toxicogenomics assay may serve as a useful tool for WWTP technology efficacy assessment, and provides guidance on process design and optimization for desired toxicity elimination and risk reduction.
Chapter 9

Quantitative Toxicogenomics Assay Reveals the Evolution and Nature of Toxicity during Environmental Pollutants Transformation

9.1 Abstract

Advanced oxidation processes (AOPs) have been shown promising for effective elimination of contaminants of emerging concerns (CECs). However, the incomplete mineralization of CECs during the AOPs can generate transformation products that may be comparable or even more toxic than the original contaminant. In this study, we demonstrated the application of a novel, fast and cost-effective quantitative toxicogenomics-based approach for evaluation of the toxicity evolution and nature along the electro-Fenton oxidative degradation of three representative CECs whose oxidative degradation pathways have been relatively well studied, and they were bisphenol A, triclosan and ibuprofen. Toxicity evolution as the results of parent chemicals transformation and intermediates production during the course of degradation are monitored, and the quantitative toxicogenomics assay results revealed the dynamic toxicity changes and mechanisms, as well as their association with identifiable intermediates during the electro-Fenton oxidation process of the selected CECs. Although for the three CECs, majority (>75%) of the parent compounds disappeared at 15-
minute reaction time, the nearly complete elimination of toxicity required a minimal 30-minute reaction time, and they seem to correspond to the disappearance of aromatic intermediates. Bisphenol A led to a wide range of stress responses, and most of the identified transformation products from bisphenol A degradation at 15 minutes contained phenolic or quinone group, such as 1,4-benzoquinone and hydroquinone, which contributed to the sustained toxicity exhibited as DNA stress (genotoxicity) and membrane stress. Triclosan is known to cause severe oxidative stress, and although oxidative damage potential decreased concomitantly with the disappearance of triclosan after 15 minutes reaction, the sustained toxicity associated with both membrane and protein stress were likely attributed at least partially to the production of 2,4-dichlorophenol that is known to cause production of abnormal proteins and affect cell membrane. Ibuprofen affects the cell transporter function and exhibited significantly high membrane stress related to both membrane structure and function. Oxidative degradation of ibuprofen led to a shift in toxicity profile from mainly membrane stress to one that exhibited not only membrane stress, but also protein stress and DNA stress. The sustained or even slightly increased toxicity corresponding to membrane and protein stresses were likely contributed by the identified intermediate, 4-isobutylacetophenone, which is known to be toxic to cell walls and membrane. The information-rich high-resolution toxicogenomics results served as ‘fingerprints” that discerned and revealed the toxicity mechanism at the molecular level among the CECs and their oxidation transformation products. This study demonstrated, for the first time, that the quantitative toxicogenomics
assay may serve as a useful tool for remediation technology efficacy assessment, and provide guidance on process design and optimization for desired toxicity elimination and risk reduction.

9.2 Introduction

Contaminants of emerging concern (CECs), such as endocrine disrupting compounds (EDCs), pharmaceuticals and personal care products (PPCPs), and nanomaterials, are frequently detected in aquatic environments. Great scientific and engineering challenges exist in addressing the water quality problems associated with contaminants of emerging concern (CECs) in terms of understanding their harmful impact and associated risk, and developing of cost-effective remediation technologies (Schwarzenbach et al., 2006). Traditional water and wastewater treatment processes are not designed to eliminate most of the CECs, especially at the trace levels as those present in drinking water and aquatic environment (Sedlak et al., 2003; Jones et al., 2005). Various treatment technologies have been explored and advanced oxidation processes (AOPs) such as photocatalysis (Rosenfeldt and Linden, 2004), ozonation (Zhang et al., 2005), dioxide chlorine (van den Heuvel et al., 2006), Fenton-based process (Dirany et al., 2012; Gozmen et al., 2003; Neyens and Baeyens, 2003; Isarain-Chávez et al., 2010; Flox et al., 2006) and other strong oxidants (Jiang et al., 2012) have been reported to be promising for effective elimination of CECs.

In consideration of the energy cost, chemical usage and length of reaction, and within the economically reasonable range of doses and reaction time commonly
applied, the mineralization rates (i.e. TOC removal) by most AOPs are generally low, generating byproducts often with higher polarity and solubility than the parent compounds (Rosenfeldt and Linden, 2004; He et al., 2008). The oxidative transformation pathways of organic contaminants are rather complex and dynamic. The degradation intermediates mixtures may exhibit comparable or even more toxic than the original contaminant(s) (Boxall et al., 2004; Farré et al., 2008). Some recognized toxic byproducts include hydroxylamines (Huber et al., 2003), phenols (Boxall, 2009), quinones (Boxall, 2009), aldehydes (Linden et al., 2007; Miller et al., 2001), and chlorinated or brominated compounds (Li and Crittenden, 2009; Grebel et al., 2010). Genotoxic chemicals produced from photo-oxidation of anthracene and carbamazepine are two examples (Brack et al., 2003; Chiron et al., 2006). Since the ultimate goal of remediation is to eliminate the toxicity threats and risks associated with the contaminants, monitoring and understanding the toxicity evolution during the remediation process is of great importance for remediation technology effectiveness assessment. However, it is not practically feasible, if possible at all, to evaluate the toxicity of a large number of water samples containing a mixture of contaminants using the conventional resources-intensive animal-based toxicity assessment. Due to the lack of feasible, fast and cost-effective methods, studies that investigate the temporal course of transformation intermediates and associated toxicity evolution during remediation processes are very rare (Dirany et al., 2012; Zhao et al., 2008; Chiang et al., 2004). The challenges in the required resources and time participated to handle the toxicity testing efforts for the large and ever-increasing number of contaminants
in various environmental sample matrixes have motivated a new vision of toxicity testing strategy as proposed by NRC (National research council) and USEPA (NRC, 2007). It calls for a systematic transit from current resources-intensive and time-consuming *in vivo* whole animal-based testing to *in vitro* mechanistic toxicity pathway-based assays on cell lines, using cost-effective, reliable and high-throughput screening and tier tier testing (NRC, 2007). Single or battery of *in vitro* cell-based bioassays using cell lines or biomarkers has been successfully applied for to environmental water samples (Neale et al., 2012; Trovo et al., 2009; Macova et al., 2011). These *in vitro* bioassays are more sensitive towards early warning signs since cellular responses are activated before the actual harm occurs, and they indicate the presence of associated stressors (Neale et al., 2012; Escher et al., 2011). However, these batteries of bioassays consisting of a number of separated assays with different species still require a substantial amount of time and efforts. In addition, most of them only provide information for specific mode of action toxicity endpoints without more comprehensive and detailed information of the toxicity profiles and mechanisms. Furthermore, isolated assays that cover only one or a few biomarkers or specific toxicity mechanisms effects cannot be directly anchored to an integrated adverse outcome or phenotypic endpoints (Neale et al., 2012).

Recently, we reported a toxicogenomics-based toxicity assessment method that allows a fast, yet informative, mechanistic and quantitative toxicity evaluation of pollutants toxicants using transcriptomics technology with GFP-fused whole-cell array of *E.coli* K12, MG1655 (Gou et al., 2010; Gou and Gu, 2011). Compared
with the traditional cytotoxicity approach, or the current battery of bioassays approach, this quantitative toxicogenomics method in combination with bioinformatics computation detects the overall toxicity response, level and reveals potential toxicity profiles and mechanisms with one single assay within hours (2-4 hours), which greatly reduces the test time, resource and cost. In addition, the measurement is sensitive enough to capture the sub-cytotoxic impact at concentrations much lower than those that can lead to detectable phenotype effects, therefore it requires less or even possibly eliminates the sample extraction and concentration step. The proposed method may serve as an alternative or complementary approach to the current toxicity assays for environmental applications.

In this study, we, for the first time, demonstrated the application of a quantitative toxicogenomics-based approach for evaluation of the toxicity evolution and nature along the electro-Fenton oxidative degradation of three representative CECs, namely, bisphenol A, triclosan and ibuprofen. These chemicals were selected because their toxicity and oxidative degradation pathways have been relatively well studied (Karim and Husain, 2010; Dann and Hontela, 2011; Trappe et al., 2002). Toxicity evolution as the results of parent chemicals transformation and intermediates production during the course of degradation were monitored to reveal the dynamic toxicity changes and mechanisms, as well as their association with identifiable intermediates during the electro-Fenton oxidation process of the selected CECs.
9.3 Materials and Methods

9.3.1 Chemicals

Bisphenol A, 4-chlorocatechol (97%), 4-chlororesorcinol (98%), chloro-p-benzoquinone (95%) and 4-isobutylacetophenone were purchased from Sigma-Aldrich. Ibuprofen (99%), 2,4-dichlorophenol (99%), 4-ethylbenzadehyde (98%), resorcinol (98%), chlorohydroquinone (90%), catechol (99%) and benzoquinone (99%) were supplied by Acros. Triclosan (99.7%) was from Calbiochem, and hydroquinone was from Fischer Sci. Deionized water (18.0 mΩ·cm) obtained from a Millipore Milli-Q system was used in all the experiments. All chemicals used in this study were above analytical grade.

9.3.2 Electro-Fenton Degradation of CECs (Appendix 9.1).

A 600-mL acrylic cell was used for the electro-Fenton degradation of CECs at ambient temperature (25 ± 1°C). Two pieces of mixed metal oxides (MMO, \( \text{IrO}_2/\text{Ta}_2\text{O}_5 \) coated on titanium mesh type, 3N International, USA) at 85 × 15 × 1.8 mm (length × width × thickness) were used as the anode and cathode with 42 mm spacing in parallel position. More detailed description and characterization of the electro-Fenton system is in our previous report (Yuan et al., 2013). 410 mL of 2 mM Na₂SO₄ solution in DI water was transferred into the cell. Certain volumes of bisphenol A, triclosan and ibuprofen stock solution were then added to the expected initial concentrations (5 mg/L). Reactions were allowed to continue for 2 hours with continuous stirring at 600 rpm using a Teflon-coated magnetic stirring bar. The electrolysis was maintained at constant current. For process optimization,
H₂O₂ production rate was tested with various dosages of Fe(II) added to the electrolytic system under acidic condition, different pH, and varying current (Yuan et al., 2013). The condition used for this study applied dosage of Fe(II) of 6.9 mg/L, with an initial pH 3 and electrical current 40 mA based on previous system optimization (Yuan et al., 2013).

**9.3.3 Transformation Products Identification by Chemical Analysis.**

The detailed methods for transformation products identification and chemical analysis could be found in our previous report (Yuan et al., 2013). Briefly, bisphenol A, triclosan, and ibuprofen were measured by a 1200 Infinity Series HPLC (Agilent) equipped with a 1260 diode array detector (DAD), a 1260 fluorescence detector (FLD) and a Thermo ODS Hypersil C18 column (4.6 × 50 mm) and an Agilent Eclipse AAA C18 column (4.6 × 150 mm). The key transformation intermediates at different time points during the remediation process were identified by HPLC and GC-MS. The Intermediates identified were listed in Table S1 in the supporting information. The proposed transformation pathways for the three CECs based on the identified key intermediates were summarized in Figure S1. TOC concentration was measured by a TOC analyzer (TOC-L CPH, Shimadzu).

**9.3.4 Sample Preparation for Toxicity Assessments.**

The initial samples containing the contaminants, as well as samples taken during the oxidation process at 15, 30 and 60 minutes were subjected to the toxicity assessment. All samples were neutralized to pH around 7 prior to enrichment by lyophilization (Freezone 4.5, LABCONCO, Kansas City, MO) under -42°C with
0.1 Torr vacuum. The lyophilization procedure was used since it has been reported to have higher recovery rates of non-volatile compounds than solid phase extraction (SPE) (Hirsch et al., 1998).

9.3.5 Microtox Assay.

Photobacterial *Vibrio fisheri* (ATCC #49387) specific photosynthesis inhibition test (Microtox) was used for nonspecific toxicity evaluation according to the ISO standard method 11348-2 (“ISO11348-2”, 2007). The assay was performed in solid black 96-well microplate (Costar, Bethesda, MD, USA). Briefly, fresh prepared *V. fisheri* culture was equally mixed with desired water sample, and the luminescence inhibition was calculated by the difference between light intensity at time 0 and 15 minutes measured with a micro-plate reader (Synergy Multi-Mode, Biotek, Winooski, VT). 1 M NaCl solution was used as negative control and ZnSO₄ was used as a positive control. The result from Microtox tests were indicated as the inverse of EC50, which was the concentration causing a 50% photosynthesis inhibition effect with the concentration reported here in units of relative enrichment factor (REF), REF is the product of enrichment factor of lyophilization process and dilution factor of bioassay.

9.3.6 Toxicogenomics Assay and Data Computation.

A high-throughput mechanistic toxicity assay method was employed that uses with GFP-fused whole-cell array of *E.coli* K12, MG1655 (Gou et al., 2010, Gou and Gu, 2011; Zaslaver et al., 2006; Onnis-Hayden et al., 2009). The water samples (after freeze dried) were re-suspended in minimum media (M9) containing *E.coli* culture strains to various desired concentrations. The selected
assay library covered a variety of genes involved in different known cellular stress response pathways, such as oxidative stress, DNA stress, protein stress, and membrane stress (including membrane transporter, efflux pump, energy metabolism, flagella metabolism, lipopolysaccharide metabolism), etc.

The protocol to measure the temporal gene expression profile was described in our previous reports (Gou et al., 2010; Gou and Gu, 2011). In brief, *E.coli* was cultivated in 384-well plates (Costar, Bethesda, MD, USA) in dark to avoid GFP photobleaching until the early exponential growth stage is reached (OD600 ~ 0.2). After addition of the aqueous sample, the plate was placed into a micro-plate reader (Synergy Multi-Mode, Biotek, Winooski, VT) for cell growth measurement (absorbance, 600 nm) and fluorescent reading (Excitation: 485 nm, Emission: 528 nm) every 5 minutes over a period of 2 hours. Three biological replicates were performed for each condition. To quantify the magnitude of the altered gene expression level upon the exposure to toxicants, the induction factor I was determined for a given gene at each time point upon chemical exposure by normalized no-toxic exposure control. The toxicity mechanisms were elucidated by identifying and analyzing the changes in genes associated with specific stress response pathways. The molecular endpoint was quantified by the TELI (Transcriptional Effect Level Index) value, which was a recently developed index for toxicogenomics data interpretation that incorporates the number, magnitude and the temporal pattern of genes with an altered expression (Gou and Gu, 2011; Gao et al., 2012). TELI values can be determined for a single gene, or for a number of genes representing a specific pathway or the entire stress response
assays (Gou and Gu, 2011). The specific or overall toxicity obtained by the toxicogenomics-based approach was represented by the corresponding REF/concentration that causes the TELI value to reach at 1.5 (EC-TELI1.5), similar to the approach that has been applied on umuC genotoxicity assay by Neale et al (Neale et al., 2012). The corresponding oxidative stress and genotoxicity for each sample were also calculated as the equivalent concentration (oxidative_TEQ and geno_TEQ) of reference compounds H$_2$O$_2$ (Verit et al., 2006) and mitomycin (Salamone et al., 1980), respectively.

9.3.7 Gene Enrichment Analysis

The gene set enrichment analyses (GSEA) were achieved by ranking a list of genes with the TELI values according to the work of Subramanian (Subramanian et al., 2005). For each pathway, GSEA calculated the enrichment score by examining the ranked gene list from high score end to the low score end, giving a rewarding score if a gene belongs to the pathway of interest, and penalizing score otherwise. The significance of each pathway was determined by comparing their ranking scores to the corresponding empirical distributions. The null distributions were generated by randomly permuting the specific pathway and all others 1,000 times. Only those categories with $p$-value less than 0.05 were considered to be altered significantly.
9.4 Results and Discussion

9.4.1 Temporal Toxicity Evolution and Profile Changes during the CECs Degradation.

Figure 9.1 and Appendix 9.3 showed the toxicity results during the electro-Fenton transformation of the three chemicals at 0, 15, 30, and 60 minutes based on the toxicogenomics assay endpoints and the Microtox results, respectively. Microtox assay could only detect the toxicity of triclosan at time zero, indicating that triclosan exhibits higher toxicity than bisphenol A and ibuprofen. Bisphenol A and ibuprofen are known to be non-toxic to gram-negative bacteria (Elvers and Wright, 1995). The Microtox results showed significant toxicity reduction of triclosan after 15 minutes of oxidative transformations. However, the changes in toxicity of bisphenol A and ibuprofen could not be detected and discerned by the Microtox assay at the detection limit under our study condition (relative enrichment factor (REF) =10). In comparison, the more sensitive toxicogenomics assay clearly revealed the toxicity reduction trends during the oxidation transformation process.
Figure 9.1. Toxicity changes during the electro-Fenton transformation of (a) bisphenol A, (b) triclosan and (c) ibuprofen. The transformation conditions are based on 5 mg/L of initial concentration for each CEC, 6.9 mg/L Fe(II), pH 3, 40 mA, and 2 mM Na₂SO₄ background electrolyte. Y axis left: 1/EC-TELI1.5 is the inverse of the relative enrichment factor (REF) that yields TELI value of 1.5,
which was determined from dose-response curves; Y axis-right: C/C0 values indicate the percentage reduction of the parent compounds at a given time point.

For the three CECs tested, the parent compound reduced rapidly during the initial 15 minutes with 73%, 96%, and 83% reduction for bisphenol A, triclosan, ibuprofen, respectively. However, toxicity evaluation, based on molecular endpoint 1/EC-TELI1.5 values, indicated that the toxicity did not always decrease proportionally and concomitantly with the disappearance of the parent contaminant (Figure 9.1). For example, although 83% of ibuprofen disappeared after 15 minutes, the overall toxicity did not seem to decrease and it actually showed a slight increase (Figure 9.1c). The sustained or elevated toxicity was likely associated with the production of toxic intermediates. The nearly complete elimination of toxicity (1/EC-TELI1.5 values, which is a cellular benchmark dose, less than 1 (the original concentration) is considered non-toxic level according to the USEPA WET and ISO water quality methods (“ISO 8692”, 2004; USEPA, 1994) for the three CECs required a minimal 30-minute reaction time, and they seemed to correspond to the disappearance of identified aromatic intermediates (Appendix 9.2). The TOC measurements (Appendix 9.4) also suggested limited mineralization even after 60 minutes degradation, with 49.2%, 74.1% and 64.3% TOC remaining for bisphenol A, triclosan and ibuprofen, respectively. This has important implications for AOP applications since the high-energy intensity of these processes makes it economically impractical to target for complete mineralization. Therefore, system design, optimization and operation should
incorporate a remediation efficacy assessment to ensure desired toxicity and risk reduction (Makgato et al., 2010; Esplugas et al., 2007).

9.4.2 Quantitative Toxicogenomics Assay Reveals the Nature and Dynamics of Toxicity Evolution during the CECs Transformation

The temporal toxicity profile changes during the transformation of individual CECs were shown in Figure 9.2 and the significantly altered stress response category ((p-value<0.05) based on gene set enrichment analysis were highlighted. The magnitude of gene expression changes was indicated by the absolute TELI value of each gene within 2 hours exposure (that considers both up- and down-regulations as altered expression) (Gou and Gu, 2011). These profiles revealed the dynamic nature of toxicity resulted from the individual CECs as well as those from a mixture of transformation products during the course of the degradation process. To further quantitatively illustrate the nature of toxicity changes during the transformation process of the CECs, molecular endpoint 1/EC-TELI1.5 for individual stress response category were calculated and plotted in Figure 9.3. Potential oxidative stress and genotoxicity were also calculated as the equivalent concentration of reference compounds H$_2$O$_2$ (Verit et al., 2006) and mitomycin (Salamone et al., 1980), respectively (Figure 9.3). The concept of toxic equivalent concentration was a widely applied method to express the toxicity of complex mixtures of compounds that act via receptor-mediated mechanisms such as induction of the aryl hydrocarbon or estrogen receptors. Escher et al. expanded this concept to molecular assay such as $umuC$ assay for genotoxicity (Neale et al., 2012; Macova et al., 2011). Here we further expanded the TEQ concept for
quantifying toxicogenomics-based molecular assays.

As shown in Figure 9.2 and 9.3, bisphenol A seemed to impact affect genes across all stress response categories. Bisphenol A is known to cause DNA damage as it activates recA gene that serves as a regulatory protein to initiate SOS response to DNA damage (Kim et al., 2007; Keseler et al., 2011). Most of the identified intermediates from bisphenol A degradation at 15 minutes contained phenolic or quinone group. Quinone electrophilic reactive intermediates detected such as 1,4-benzoquinone, hydroquinone could cause DNA damage.
Figure 9.2. Exemplary temporal toxicity evolution and profile changes during the electro-Fenton oxidation transformation of the three CECs (at REF=1, the original concentration). Profiles for other concentrations (REF) are not shown. X-axis top: list of stress response categories and selected genes. Altered gene expression changes in relative to the untreated control were indicated as the TELI value for each individual gene and color-coded with the scale 1-5 (Red spectrum colors indicate the magnitude of the altered gene expression, while black indicates no change). Y-axis left: chemicals, right: reaction time in minutes. Those stress response categories highlighted with green stars (*) were significantly (p<0.05) affected based on the gene enrichment analysis.
Figure 9.3. Toxicity nature profiles based on the quantitative endpoints of different stress response categories for the three CECs, (a) bisphenol A, (b) triclosan and (c) ibuprofen, during the electro-Fenton transformation process. Y axis left: 1/EC-TEL1.5 is the inverse of the relative enrichment factor (REF) that
yields TELI value of 1.5. 1/EC-TELI1.5 value for individual stress response
categories was determined from dose-response curves, and it revealed the specific
toxicity mechanism and nature. Y axis right: Genotoxicity, measured as
Geno_TEQ, was determined as concentration of reference model genotoxic
compound MMC (mitomycin) that yields TELI value of 1.5. Oxidative damage
potential, measured as Oxidative-TEQ, was determined as concentration of
reference oxidant H2O2 that yields TELI value of 1.5.

Therefore leading or at least contributing to the sustained DNA stress (Ooe et al.,
2005; Ludewig et al., 1989; Das et al., 2010). Bisphenol A also induced sustained
membrane stress at 15 minutes, which might be contributed by one of the
intermediates identified - 1,4-benzoquinone. As bio-active quinone, 1,4-
benzoquinone affects the electron transport chain and transporter activity on
membrane (Das et al., 2010).

Triclosan is known to cause severe oxidative stress as well as DNA stress by
permeating the bacterial cell wall and targeting multiple cytoplasmic and
membrane sites (Dann et al., 2011; Tamura et al., 2012). This was consistent with
the significant oxidative stress and DNA stress exhibited by tricolsan at time 0
(Figure 9.2 and 9.3). Both oxidative damage potential and genotoxicity (shown as
geno_TEQ and oxidative_TEQ) decreased concomitantly with the disappearance
of triclosan after 15 minutes reaction. The sustained toxicity associated with both
membrane and protein stress (Figure 9.3b) at 15 minutes were likely attributed at
least partially to the production of 2,4-dichlorophenol. The phenol group in 2,4-
dichlorophenol can cause protein stress by the production of abnormal proteins
Ibuprofen affects the transporter function as influx and efflux of the cell membrane (Parepally et al., 2006), thus it exhibited relatively significantly severe membrane stress at the initial stage (Figure 9.2c and 9.3c). Oxidative degradation of ibuprofen led to shift in toxicity profile from mainly membrane stress to one that exhibited not only membrane stress, but also protein stress and DNA stress at 15 minutes. The sustained or even slightly increased toxicity corresponding to membrane and protein stresses at 15 minutes were likely associated with the identified intermediate, 4-isobutylacetophenone, which was known to be toxic to cell walls and membrane (Castell et al., 1987). Both ibuprofen and 4-isobutylacetophenone can strongly bind to proteins, affecting protein metabolism and causing protein dysfunction, thus inducing protein stress. The slightly elevated DNA stress at 15 minutes indicated that either the identified transformation byproduct 4-isobutylacetophenone also exhibits genotoxicity or there were other unidentified genotoxic intermediate. Oxidative stress showed the relatively weak response for both ibuprofen and its intermediates, in agreement with previous reports (Trappe et al., 2002; Parepally et al., 2006).

For the three CECs evaluated, a minimal 30 minutes seemed to be required for completion of toxicity reduction, although varying levels of residual toxicity existed. The residual of general stress, membrane stress, and oxidative stress after 30 and 60 minutes electro-Fenton transformation of bisphenol A may indicated
potential production of persistent products that still have some toxicity. These results demonstrated that oxidative transformation pathways of organic contaminants were rather complex and they may lead to dynamic toxicity evolutions and changes as the results of the degradation intermediates formation. The toxicogenomics-based assay provided detailed information and insights into the dynamic toxicity changes and mechanisms, as well as their association with identifiable intermediates during the transformation processes.

7.4.3 Insights Into the Distinct Mode of Action of Toxicity During the Transformation of Three CECs

More in-depth examination of gene activities disclosed genetic level information and distinction among the toxic effects and mechanism of the parent testing chemicals and their intermediates evolved during the oxidative transformation process. Figure 9.4 showed the detailed gene activation (TELI value for single gene > 1.5) for oxidative stress, DNA stress and membrane stress response pathways.

**Oxidative Stress**

Figure 9.4a included 10 selected essential oxidative stress biomarker genes, which are involved in defense systems to defend oxidative stress and to scavenge oxidative radicals in the form of enzymes that can detoxify reactive oxygen species (ROS) (Keseler et al., 2011). Three of them are ROS sensor, *oxyR*, *soxR*, and *soxS*. Two alkyl hydroperoxide reductase, *ahpC* and *ahpF*, are the primarily scavenging enzymes for endogenously produced organic hydroperoxide and H$_2$O$_2$. Two catalase/hydroperoxidase, *katG* and *katE*, are the primary scavenger at
high H$_2$O$_2$ concentrations. Three superoxide dismutases (SOD), sodA, sodB, and sodC, represent the first line of defense against ROS, converting superoxide radicals to hydrogen peroxide and water.

For all three CECs, although the parent chemical as well as the resultant mixture of residual parent compound and intermediates exerted oxidative stress, the distinct activation profiles of specific oxidative stress biomarkers revealed the discernable and varying molecular level oxidative damaging pathways evoked by the different chemicals. Both bisphenol A and triclosan altered more than half of the selected oxidative stress genes at time zero and 15 minutes, indicating the oxidative nature of these two chemicals and their transformation intermediates. Both bisphenol A and triclosan led to up-regulation of master hydrogen peroxide stress regulator and sensor oxyR at 0 and 15 minutes, indicating the presence of ROS as hydrogen peroxide radical. In comparison, ibuprofen led to over expression of less number of oxidative stress biomarkers and did not activate oxyR, suggesting weaker extent of oxidative stress. Varying superoxide dismutases (SODs) were up-regulated during the degradation process of the three CECs. All of the catalase/hydroperoxidase and alkyl hydroperoxide reductase showed up-regulation during the degradation of triclosan while only part or none of them showed up-regulation with the other two CECs. This indicated that triclosan and its intermediates likely pose more H$_2$O$_2$–related oxidative stress.
Figure 9.4. Stress response pathways and biomarkers genes that showed altered expression (TELI_gene > 1.5) during the electro-Fenton transformation of the three CECs. (a) oxidative stress pathways, (b) DNA damage and repair pathways, and (c) membrane stress response pathways. The genes were clustered into subcategories based on their functions and involvement in various pathways. Different colors indicated different time during electro-Fenton treatment, blue – 0 minute, red – 15 minutes, green – 30 minutes, and purple – 60 minutes.

**DNA stress**

Impacts on DNA damage pathways (11 biomarker genes) by the tested chemicals and their transformation products during the electro-Fenton reaction were shown in Figure 9.4b. Gene *recA* and *lexA*, which control the transcription of several genes involved in the cellular response to DNA damage, are involved in SOS regulation. Gene *ada*, one of the two separate direct repair mechanisms in *E.coli*,

<table>
<thead>
<tr>
<th>Gene</th>
<th>Bisphenol A</th>
<th>Triclosan</th>
<th>Ibuprofen</th>
</tr>
</thead>
<tbody>
<tr>
<td>emrA</td>
<td>emrA</td>
<td>emrA</td>
<td>emrA</td>
</tr>
<tr>
<td>marR</td>
<td>marR</td>
<td>marR</td>
<td>marR</td>
</tr>
<tr>
<td>cyaA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>enE</td>
<td>emE</td>
<td>enE</td>
<td>enE</td>
</tr>
<tr>
<td>sanA</td>
<td>sanA</td>
<td>sanA</td>
<td>sanA</td>
</tr>
<tr>
<td>dacB</td>
<td>dacB</td>
<td>dacB</td>
<td>dacB</td>
</tr>
<tr>
<td>bacA</td>
<td>bacA</td>
<td>bacA</td>
<td>bacA</td>
</tr>
<tr>
<td>amiC</td>
<td>amiC</td>
<td>amiC</td>
<td>amiC</td>
</tr>
<tr>
<td>clsA</td>
<td>clsA</td>
<td>clsA</td>
<td>clsA</td>
</tr>
<tr>
<td>motA</td>
<td>motA</td>
<td>motA</td>
<td>motA</td>
</tr>
<tr>
<td>figM</td>
<td>figM</td>
<td>figM</td>
<td>figM</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time</th>
<th>0 minute</th>
<th>15 minutes</th>
<th>30 minutes</th>
<th>60 minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bisphenol A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Triclosan</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ibuprofen</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
controls the transcription of the genes involved in the process of reparation of alkylated DNA. Three genes are involved in base excision repair (BER) and they are mutT, nfo, and polB. Nucleotide excision repair (NER) is a generalized DNA repair process that can repair a wide diversity of DNA lesion and, uvrA and uvrD are indicative of this repair pathway. Gene uvrD, which is superfamily I DNA helicase, is also involved in mismatch repair (MMR). Another gene involved in MMR, ssb, is a highly stable single-stranded DNA binding protein that also functions in double strand break repair (DBR) and DNA replication (Figure 9.4b). Bisphenol A activated lexA, recA and other genes involved in multiple DNA damage and repair pathways during the 30 minutes reaction (0, 15, and 30 minutes), especially those genes involved in double strand break repair that repair those breaks occur at two or more locations in the chromosome (Jackson, 2001), indicating relatively severe DNA damage caused by BPA. This indicated that bisphenol A and its intermediates led to the most severe DNA stress among the three CECs and these DNA damaging products were relatively persistent and could only be eliminated after 60 minutes. Triclosan resulted in changes in the gene expression level for most of the DNA repair pathway, including BER, NER, MMR, and DBR. The transformed intermediates of triclosan still exhibited some DNA stress at 15 minutes. The number of up-regulated DNA repair genes increased after 15 minutes degradation of ibuprofen, which indicated the generation of potentially unknown genotoxic intermediates.

Membrane stress

Membrane stress is a less characterized term than oxidative and DNA stress in
this study. Genes were classified into several sub-categories based on the functions related to membrane, including those related to energy (sdhC, cyoA), multidrug efflux (emrA, emrE, marR), cell wall and membrane structure (san, bacA, and dacB), cell envelope (amiC and clsA) and flagella (motA, flgM) (Keseler et al., 2013).

Many genes related to multidrug efflux and cell membrane structure and envelope showed up-regulation along the degradation of three CECs, but with distinct activation profiles as the results of different mode of action (Figure 9.4c). Multidrug efflux is a universal mechanism responsible for extrusion of toxic substances and antibiotics outside the cell. Most of these genes are usually induced by the chemicals with phenolic rings or quinone mediated bactericidal activity (Barabote et al., 2003). As discussed earlier, most of the parent compounds and intermediates of the three CECs contained phenolic or quinone group. Ibuprofen affected many genes related to multidrug efflux, as well as the cell wall and membrane structures, which were believed, could be linked to drug resistance (Keseler et al., 2013). The increasing number of altered genes at 15 minutes during the triclosan oxidation suggested the generation of membrane-damaging intermediates that were further degraded after 30 minutes. For bisphenol A, genes related to multidrug efflux emrA and marR showed over expression at 60 minutes, indicating the potential presence of phenolic rings-containing products even after 60 minutes process.

9.5 Conclusion

There are limited reports that evaluate toxicity changes during the degradation of
CECs by AOPs. Chiang et al. observed an increase in toxicity in the initial stage of photocatalytic oxidation of bisphenol A and a gradual decrease afterwards using the luminescent bacterial *V. fisheri* (Chiang et al., 2004). Dirany et al. noted a sudden increase in the inhibition of luminescence of *V. fisheri* for electro-Fenton degradation of sulfachloropyridazine (Dirany et al., 2012). Zhao et al. observed a slight increase in estrogenic activity in the heterogeneous photo-Fenton degradation of 17β-estradiol in simulated drinking water using yeast-based enzyme expression (Zhao et al., 2008). These previous studies reported the changes in phenotype endpoints, but with no or limited insights into the temporal dynamics of the nature of toxicity mechanisms associated with transformation products evolved during the degradation. In comparison, the high throughput mechanistic toxicogenomics-based toxicity assay used in this study enables a more comprehensive and detailed toxicity evaluation during the remediation process, which can be linked with key intermediates, thereby improving the understanding of underlying toxicity changes during a degradation process. Therefore, the quantitative toxicogenomics assay may serve as a useful tool for remediation technology efficacy assessment, and provides guidance on process design and optimization for desired toxicity elimination and risk reduction.
Chapter 10

Quantitative Toxicogenomics Assay Revealed the Impact of Disinfection Technologies On Effluent Toxicity

10.1 Abstract

Disinfection technologies have been extensively practiced in U.S. for control of infectious disease. Recent years, the potential adverse effects resulted from disinfection process has been a great concern. Toxicogenomics approach are increasingly used for rapid, relatively inexpensive toxicity screening that can be used with analytical chemistry data for water quality evaluation and the effectiveness assessment of water treatment technologies. In this study, we applied a novel stress gene-based toxicogenomic approach, using GFP-fused E coli reporter cell library, to evaluate the mechanistic toxicity of municipal wastewater effluents disinfected by different technologies, including chlorination (Cl2), chloramination (NH2Cl), ozonation (1.5 mg/L; 6.0mg/L), ultraviolet irradiation (UV; 250 mJ/cm2; 1000 mJ/cm2) as well as combination of ultraviolet irradiation and hydrogen peroxide (UV/H2O2; 10 mg/L&250 mJ/cm2; 10 mg/L&1000 mJ/cm2). The overall toxicity of disinfected secondary effluents was first evaluated by Microtox assay with V. fischeri bioluminescence inhibition tests, and the results indicated that no cytotoxicity could be detected with the effluents with corresponding criterion. The toxicogenomics-based molecular toxicity assay
showed more sensitive results than the Microtox assay. Chloramination lead to a remarkable increase in the total toxicity compared to the raw effluent. Interestingly, chlorination ($p < 0.01$) and ozonation ($p < 0.01$) seemed to remove the original toxicity from raw water effectively. The altered gene expression profiles for various effluent samples exhibited distinctive patterns unique for different disinfection processes. Effluents treated by chloramination showed strong responses over all stress categories, especially oxidative stress and DNA damage. Comparing with UV alone, addition of H2O2 led to different effect on toxicity reduction. The contribution of detected chemicals to the observed overall toxicity was revealed by a designed iceberg experiment. There results showed that for some samples, the detected chemicals can only explain part of the whole water sample toxicity, which may indicate the production of transformation byproducts during the treatment. In summary, this study demonstrated that toxicogenomics-based assay could be applied for detailed and informative toxicity evaluations of water samples and for treatment technology effectiveness assessment.

10.2 Introduction

With the growing of human population and diminishing access to freshwater resources, the use of alternative water resources for drinking continues to expand. However, the quality of these water sources has been a great concern. Effluents from wastewater treatment plant are discharged to water body, which is usually used as drinking water resources. Wastewater disinfection plays a critical role in control of infectious disease and protection of water source safety (Brook et al., 2006; Blatchley et al., 2007). Chlorination, one of the most widely used
wastewater disinfectant in the U.S., has been criticized for the generation of
disinfection byproducts (DBPs), which include hundreds or thousands of DBPs
that have been identified such as trihalomethanes (THMs), haloacetic acids
(HAAs), haloacetonitriles (HANs), haloketones, haloaldehydes, and
halonitromethanes, haloamides, carbonyls and halophenol, as well as other likely
un-identified new ones (Gopal et al., 2007). These DBPs have been reported to
associate with adverse health effects including genotoxicity, estrogenic effects,
carcinogenicity (bladder, colon and rectal cancers) (Richardson et al., 2007;
Graves et al., 2001), developmental toxicity, reproductive toxicity and others
(Nieuwenhuijsen et al., 2000; Hrudey et al., 2009; Morris et al., 1992; Hidesheim
et al., 1998).

These health concerns raised from DBPs have motivated exploration of
alternative disinfectants such as chloramine, ozone, UV radiation, peracetic acid,
bromine and advanced oxidation (e.g., O3/H2O2 and UV/H2O2) (Tardiff et al.,
2006; Hwang et al., 2003; 12. Escobar-Hoyos et al., 2013). Although these
alternative disinfectants are considered to reduce the generation of certain types of
regulated DBPs levels, they may produce other emerging DBPs (Seidel et al.,
1987). The formation of iodinated THMs is reported to be the highest in water
treated by chloramination (Glaze, 1987), while iodine-containing DBPs are
generally known as the most toxic DBP (Siddiqui et al., 1997).

Nitrosodimethylamine (NDMA), a carcinogen (Gehr et al., 2003), can also form
at higher levels during chloramination (Graves et al., 2001). Ozonation, which can
reduce or eliminate formation of THMs and HAAs effectively, resulted in,
however, increase of bromate, especially for water with high bromide (Graves et al., 2001). Bromate is known to have high potent carcinogenicity in animal tests (Graves et al., 2001). Ozonation also leads to the formation of carboxylic acids, aldketoacids, aldehydes, ketones, hydroxy acids, alcohols, esters, and alkanes (Glaze, 1987; Hijnen et al., 2006). Many of those would act as precursors of halonitromethanes (HNMs) (Richardson, 2003), iodo-THMs (Weinberg et al., 2002), haloaldehydes (Siddiqui et al., 1997; Krasner, 2009) as well as other genotoxic and carcinogenic emerging DBPs5. Although UV is commonly recognized as DBP-free disinfectant, a few recent studies indicated that UV leads to the generation of HNMs and trichloropropanone (USEPA, 2015).

The health concerns for the DBPs and other transformation products generated during the employment of various disinfection technologies demands for better understanding of DBPs formation mechanism as well as their consequent toxicity and implicated risks. Toxicity testing, served as a supplement to chemical analysis, is able to provide more comprehensive and accurate information on ecological or human health impacts resulted from disinfection practice. A limited numbers of studies have investigated toxicological impacts brought by disinfection or advanced oxidation processes during drinking water or wastewater treatment (Richardson et al., 1999; Hoigne and Bader, 1988; Bichsel and von Gunten, 2000; Krasner et al., 2006; Reckhow et al., 2010; Tothill and Turner, 1996; Reemtsma, 2001). Most of these studies focused on one or two disinfection processes. These studies provided evidence of toxicity evolution during disinfection and advanced oxidation treatment processes, however, detailed toxicity mechanism and
knowledge on the relationship between toxicity and chemical evolution haven’t been explored. One major barrier for advancing this understanding is the lack of standardized, cost-effective and feasible yet comprehensive toxicity assay that can be widely applied for water effluent toxicity and technology efficacy assessment.

The recognized inherent limitations in current toxicological approach (i.e. cost-prohibitive, long test duration, low detection limit etc.) demands a systematic transit from solely relying on expensive and lengthy in vivo whole animal-based testing to a tiered testing strategy that employs in vitro mechanistic toxicity pathway-based high-throughput bioassays (Krewski et al., 2011; NRC, 2007). Recently, single or a battery of in vitro bioassays (i.e. Ames test (Farreand Barcelo, 2003), umuC assay (Waston et al., 2012), reactive oxygen species (ROS) assay (Neale et al., 2012), AChE inhibition assay (Misik et al., 2011), yeast estrogen screen (YES) (Macova et al., 2010)) has been successfully applied for environmental water samples by Beate’s group (Macova et al., 2010; Wang et al., 2011; Metz et al., 2011; Escher et al., 2011) as well as others (Stalter et al., 2010; Shi et al., 2009). However, several limitations still exist for these reported methods. The battery tests consisting of a set of separated assays with different species still require substantial amounts of time and efforts. In addition, most of them only provide certain toxicity endpoints without yielding comprehensive toxicity information and profiles. Furthermore, single assay that covers only one or a few biomarkers or specific effects cannot be directly translated to an integrated adverse outcome or phenotypic endpoints therefore cannot be easily linked to regulatory relevant endpoints.
Recently, we reported a toxicogenomics-based toxicity assessment method that allows a fast but informative, mechanistic and quantitative toxicity evaluation of pollutants. This method uses (GFP)-fused whole cell array of in E.coli K12, MG1655 43 which measures temporal gene expression changes of over a hundred genes involved in various cellular stress response pathways that are known to be highly conserved among species (Gou et al., 2010; Gou and Gu, 2011). In combination with bioinformatics computation, this assay detects overall toxicity level and reveals potential toxicity profiles and provides quantitative endpoints for multiple potential modes of actions (MOAs) with one single assay within hours (2-4 hours), which greatly reduces the test time, resources and costs. The objective of this study is to apply the toxicogenomics assay to compare the toxicity level and nature of secondary effluents treated with different disinfection processes. Toxicity level and mechanistic profiles for the original secondary effluent and for those that were treated with different disinfection technologies, including chlorination, chloramination, ozonation, UV radiation and UV combined with H2O2, were assessed to reveal potential toxicity reduction and/or evolution by the various disinfection processes. Iceberg designs were also employed to assess if the detected chemicals drive the biological effect and which fraction of effect remains unexplained by detected chemicals.
10.3 Materials and Methods

10.3.1 Wastewater Samples Collection and Preparation.

Wastewater samples are collected from the advanced oxidation process (AOP) piloting plants incorporated in Green Valley wastewater treatment plant (WWTP) (Pima County, Arizona, United States). The treatment processes include primary treatment, clarification, sand filtration and disinfection. WWTP secondary effluents after sand filtration is served as influents of eight parallel AOP pilot plants (Wedeco, Xylem Water Solutions, Germany) that employed different disinfection processes, including ozone (1.5mg/L; 6.0 mg/L), UV (250mJ/cm2; 1000mJ/cm2), UV/H2O2 (UV 250 mJ/cm2, H2O2 10 mg/L; UV 1000mJ/cm2, H2O2 10 mg/L), chlorination (5 mg/L free Cl2 residual after 5 min, 24 hour reaction time) and chloramination (5 mg/L free Cl2 residual after 5 min, 24 hour reaction time). Samples before and after each treatment unit are collected in amber glass bottles. Effluents treated by chlorine and chloramines are quenched with sodium thiosulfate (50 mg/L) for 24 h after collection. 1g/L of sodium azide is added and samples are stored at 4 °C until extraction (within one week). The extraction was performed using an AutoTrace™ 280 Solid-Phase Extraction (SPE) device (Dionex, Sunnyvale, CA). In brief, after 0.7 μm glass fiber filtration (GF/F, Whatman, Maidstone, U.K.), 1 L of each sample was passed through the Oasis Hydrophilic-Lipophilic Balance (HLB, 500mg/6cc) cartridge (Waters Corporation, Millford, MA), which was preconditioned by 5 mL of MTBE, 5 mL of methanol, and 5 mL of reagent water. The samples were loaded onto the cartridges at 10mL/min, and after being rinsed with 10 mL of reagent water, the
cartridges were then dried with a stream of nitrogen for one hour. Next, the cartridges were eluted with 5 mL of methanol followed by 5 mL of 10/90 (v/v) methanol/MTBE into 15 mL calibrated conical centrifuge tubes. The extracts were then concentrated with a gentle stream of nitrogen using a TurboVap (Zymark Corporation, Hopkinton, MA) to an approximate volume of 0.2 mL and brought to a final volume of 1.0 mL using methanol. The MilliQ water generated lab blank (LB) and field blank (FB), which were held in the same storage condition, was extracted the same way as the samples.

10.3.2 Microtox Assay.

The cytotoxicity of various disinfected effluent samples is evaluated by using bioluminescent bacterial *Vibrio fisheri* (ATCC #49387) inhibition test (Microtox) according to the ISO standard method 11348-2. The assay is performed in solid black 96-well microplate (Costar, Bethesda, MD, USA). Briefly, the luminescence intensity of fresh prepared *V. fisheri* culture is measured before the addition of testing water samples, as 0 minute, with a micro-plate reader (Synergy Multi-Mode, Biotek, Winooski, VT). Then water samples are volume equally mixed with *V. fisheri* culture and measure the luminescence intensity after 15 minutes reaction. 2% NaCl solution is used as negative control. Bioluminescence inhibition is determined with the intensity reading at 0 and 15 minutes, by comparing the response to a NaCl with the response to the samples. Triplicate dose-response assays are conducted for each sample. The endpoint of Microtox is indicated as EC50 (the sample’s relative enrichment factor (REF) at which it causes 50% bioluminescence inhibition effect).
10.3.3 Toxicogenomics assay with stress response pathway ensemble

Assay description

A high-throughput mechanistic toxicity assay method was employed with GFP-fused whole-cell array of *E.coli* K12, MG1655. The tested chemicals were diluted by minimum media (M9) to desired concentrations. The selected assay library covers a variety of genes involved in different known cellular stress response pathways, such as oxidative stress, DNA stress, protein stress, and membrane stress (including membrane transporter, efflux pump, energy metabolism, flagella metabolism, lipopolysaccharide metabolism), etc. (see biomarkers in STable 1). Each fusion is expressed from a low-copy plasmid, pUA66, which contains a kanamycin resistance gene and a fast folding *gfpmut2*, allowing for real-time measurement of gene expression level changes.

Assay protocol

*E.coli* was cultivated in 384-well plates (Costar, Bethesda, MD, USA) under 37°C in dark to avoid GFP photobleaching until the early exponential growth stage is reached (OD600 ~ 0.2). After addition of the aqueous samples, the plate was placed into a micro-plate reader (Synergy Multi-Mode, Biotek, Winooski, VT) for cell growth measurement (absorbance, 600 nm) and fluorescent reading (Excitation: 485 nm, Emission: 528 nm) every 5 minutes over a period of 2 hours under 37°C. Three biological replicates were performed for each condition. Test and control group are arranged on the same plate for each replicate.
**Data processing**

The *GFP* and *OD* data were first smoothed using 5-time moving average and then corrected against background (growth medium only blanks) with and without chemicals exposure, respectively. The population normalized *GFP* signal was calculated as $P = \frac{GFP}{OD}$ and corrected against background (*E.coli* strains without *GFP* infusion) at same *OD* with and without chemicals exposure, respectively. Induction factor, which measures the gene expression alteration, was calculated as the ratio of normalized expression levels between experimental (with chemical exposure) and control groups (without chemical exposure), $I = \frac{P_e}{P_c}$, where $P_e$ and $P_c$ represents for experimental and control groups, respectively.

The induction factor $I$, determined for a given gene at each time point upon chemical exposure, was then used for the following data analysis, where a gene is up-regulated if $I > 1$ and down-regulated if $I < 1$. The toxicity mechanisms were elucidated by identifying and analyzing the changes in genes associated with specific stress response pathways.

**Dose-response and toxicogenomics endpoints derivation**

The molecular response was quantified by the TELI (Transcriptional Effect Level Index) value, which was a recently developed index for toxicogenomics data interpretation that incorporates the number, magnitude and the temporal pattern of genes with an altered expression. TELI can be integrated as equation 1 on various levels, as single gene (TELI\textsubscript{gene}), pathway (TELI\textsubscript{pathway}), stress category (TELI\textsubscript{oxidative}, TELI\textsubscript{geno}, TELI\textsubscript{membrane}, TELI\textsubscript{protein}, and TELI\textsubscript{general}), and overall stress library (TELI\textsubscript{total}). Compared to previous publication, improvement was
made to TELI computation. In this paper, TELI is calculated as the average of transformed I instead of sum of transformed I to make TELI value more comparable among different pathway, stress category, and overall library.

$$
TELI = \frac{\sum_{i=1}^{n} w_i e^{[l/o(t)]}}{\text{ExposureTime} \times n}
$$

Equation 1

Where, \( t \) was the exposure time, \( n \) was the number of genes in one particular pathway/stress category/total stress library, \( w_i \) was the weight factor of gene. For this study, we assigned value of 1 for all the weighing factors.

The TELI-dose response curves were analysis with a 4-paramater logistic equation model using GraphPad 5.0c (Prism). Several toxicoenomics endpoints were derived from TELI-dose response curves.

10.4 Results and Discussion

10.4.1 Removal of detected CECs through different disinfection processes

23 CECs of various categories were analyzed in both the influent and effluent through the disinfection piloting plants (Table 10.1, Figure 10.1). Ozone processes yielded the highest removal of the parent CECs monitored, with removal efficiencies ranging mostly from 20% to 100%, at an average of 60%. UV treatment at higher doses (1000 mJ/cm2), with or without the addition of hydrogen peroxide, also led to relatively higher removal (average 40–50%). UV treatment at lower doses (250 mJ/cm2), with or without H2O2 enhancement, did not result in significant (>50%) removal for most (18 out of 23) of CECs except
for sulfamethoxazole, ditiazem, naproxene, triclocarban, and triclosan. Chlorination also lead to relatively high removal (average 50%). In comparison, the chloramination (NH2Cl) process lead to nearly no removal of most of the CECs detected, except for ditiazem and triclosan.

The evaluated CECs exhibited varying degree of susceptibility to different disinfectants (Figure 10.1, Table 10.1). Tricolsan, diclofenac and ditiazem seemed to have the highest removal efficiency by all the disinfection processes, with nearly completely removal by ozonation. This is consistent with previous report (Buth et al., 2011; Kim et al., 2009). TCEP, PEBS, TCPP, PFOA and PFOS showed high recalcitrance and resistant to nearly all treatments. Caffeine, acetaminophen, carbamazepine and primidone, were also rather persistent and they were only susceptible to removal by ozonation or UV-high dose /H2O2, with 20-70% removal. Investigations found that carbamazepine is recalcitrant and its removal efficiencies by the WWTPs using activated sludge processes are mostly below 10% (ZHANG 2008). For triclocarban, naproxene, trimethoprim, and sulfamethoxazole, the most effective removal method seemed to be chlorination in this study.
Figure 10.1. Removal efficiency of 23 detected chemicals by 8 disinfection technologies: UV, UV/H₂O₂, O₃ and Cl₂.

Table 10.1 Removal efficiency of 23 detected chemicals by 8 disinfection technologies: UV, UV/H₂O₂, O₃ and Cl₂.

<table>
<thead>
<tr>
<th>Sample</th>
<th>NH₂Cl</th>
<th>Cl₂</th>
<th>UV₂₅₀</th>
<th>UV₁₀₀₀</th>
<th>UV₂₅₀/H₂O₂</th>
<th>UV₁₀₀₀/H₂O₂</th>
<th>O₃₁.₅ mgL</th>
<th>O₃₆ mgL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetaminophen</td>
<td>0%</td>
<td>70%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>76%</td>
<td>74%</td>
</tr>
<tr>
<td>Caffeine</td>
<td>6%</td>
<td>5%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>39%</td>
<td>72%</td>
</tr>
<tr>
<td>Trimethoprim</td>
<td>30%</td>
<td>100%</td>
<td>0%</td>
<td>15%</td>
<td>6%</td>
<td>40%</td>
<td>72%</td>
<td>100%</td>
</tr>
<tr>
<td>Sucralose</td>
<td>21%</td>
<td>49%</td>
<td>13%</td>
<td>42%</td>
<td>35%</td>
<td>20%</td>
<td>22%</td>
<td>40%</td>
</tr>
<tr>
<td>Primidone</td>
<td>9%</td>
<td>13%</td>
<td>2%</td>
<td>10%</td>
<td>2%</td>
<td>36%</td>
<td>29%</td>
<td>76%</td>
</tr>
<tr>
<td>Sulfamethoxazole</td>
<td>17%</td>
<td>100%</td>
<td>58%</td>
<td>93%</td>
<td>62%</td>
<td>95%</td>
<td>79%</td>
<td>99%</td>
</tr>
<tr>
<td>Diphenylhydramine</td>
<td>32%</td>
<td>99%</td>
<td>7%</td>
<td>33%</td>
<td>25%</td>
<td>61%</td>
<td>60%</td>
<td>99%</td>
</tr>
<tr>
<td>Simazine</td>
<td>0%</td>
<td>8%</td>
<td>28%</td>
<td>68%</td>
<td>26%</td>
<td>71%</td>
<td>20%</td>
<td>60%</td>
</tr>
<tr>
<td>Ditiazem</td>
<td>77%</td>
<td>100%</td>
<td>60%</td>
<td>97%</td>
<td>72%</td>
<td>99%</td>
<td>92%</td>
<td>100%</td>
</tr>
<tr>
<td>Carbamezamine</td>
<td>3%</td>
<td>15%</td>
<td>7%</td>
<td>25%</td>
<td>20%</td>
<td>55%</td>
<td>62%</td>
<td>100%</td>
</tr>
<tr>
<td>TCEP</td>
<td>2%</td>
<td>2%</td>
<td>10%</td>
<td>0%</td>
<td>0%</td>
<td>5%</td>
<td>21%</td>
<td>6%</td>
</tr>
</tbody>
</table>
10.4.2 Baseline toxicity revealed by Microtox assay

Microtox is one of the most widely used techniques to water toxicity assessment (indicated baseline toxicity) (Escher et al., 2010, 2011, Dizer et al, 2002). Different toxicity values are observed for original raw effluents and samples subjected to different disinfection or oxidative processes. Toxic unit (100%/EC50) (Farr et al., 2001) was used for exhibition of Microtox results (Figure 10.2). Compared with the raw effluents, all disinfection processes exhibit the ability of toxicity removal, except for chlorination processes. The chlorination process led to significantly increased toxicity level with TI50 value, which is consistent with previous reports (Watson et al., 2012; Neal et al., 2012). The ozonation processes (both low and high dose scenarios) seem to be more effective in toxicity reductions that lower the baseline toxicity level by around 30%. The UV processes, either with or without combination with H2O2 addition, seemed to only slightly reduce the baseline toxicity level of the raw effluent with comparable or slightly decreased TU values. These observations show that

<table>
<thead>
<tr>
<th></th>
<th>0%</th>
<th>0%</th>
<th>0%</th>
<th>0%</th>
<th>0%</th>
<th>0%</th>
<th>0%</th>
<th>0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>PFBS</td>
<td>15%</td>
<td>19%</td>
<td>0%</td>
<td>0%</td>
<td>10%</td>
<td>12%</td>
<td>20%</td>
<td>9%</td>
</tr>
<tr>
<td>Fluoxetine</td>
<td>16%</td>
<td>25%</td>
<td>35%</td>
<td>79%</td>
<td>43%</td>
<td>84%</td>
<td>45%</td>
<td>100%</td>
</tr>
<tr>
<td>DEET</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>6%</td>
<td>1%</td>
<td>34%</td>
<td>30%</td>
</tr>
<tr>
<td>Naproxene</td>
<td>12%</td>
<td>100%</td>
<td>19%</td>
<td>53%</td>
<td>24%</td>
<td>73%</td>
<td>69%</td>
<td>100%</td>
</tr>
<tr>
<td>TCPP</td>
<td>10%</td>
<td>9%</td>
<td>10%</td>
<td>8%</td>
<td>0%</td>
<td>16%</td>
<td>1%</td>
<td>4%</td>
</tr>
<tr>
<td>Diclofenac</td>
<td>14%</td>
<td>100%</td>
<td>87%</td>
<td>99%</td>
<td>88%</td>
<td>100%</td>
<td>90%</td>
<td>99%</td>
</tr>
<tr>
<td>Ibuprofen</td>
<td>12%</td>
<td>0%</td>
<td>0%</td>
<td>14%</td>
<td>0%</td>
<td>48%</td>
<td>58%</td>
<td>90%</td>
</tr>
<tr>
<td>PFOA</td>
<td>4%</td>
<td>4%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>1%</td>
<td>0%</td>
</tr>
<tr>
<td>Gemfibrozil</td>
<td>0%</td>
<td>72%</td>
<td>0%</td>
<td>17%</td>
<td>12%</td>
<td>51%</td>
<td>43%</td>
<td>99%</td>
</tr>
<tr>
<td>Triclocarban</td>
<td>5%</td>
<td>100%</td>
<td>56%</td>
<td>95%</td>
<td>64%</td>
<td>97%</td>
<td>52%</td>
<td>100%</td>
</tr>
<tr>
<td>Triclosan</td>
<td>100%</td>
<td>100%</td>
<td>59%</td>
<td>100%</td>
<td>58%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>PFOS</td>
<td>15%</td>
<td>19%</td>
<td>0%</td>
<td>0%</td>
<td>10%</td>
<td>12%</td>
<td>20%</td>
<td>9%</td>
</tr>
</tbody>
</table>
different disinfection and oxidation processes indeed have varying effects on wastewater effluent baseline toxicity levels. Persoone et al. (Persoone et al., 2005) proposed that the lowest percentage effect considered to have a significant toxic impact corresponds to TU equal to 0.04 for effluents discharged into the aquatic environment. Following this criterion, all samples tested in this study, including raw effluents as well as disinfected effluents, are considered as “no acute toxicity”.

Figure 10.2 Comparison of toxicity level among effluents treated with different disinfection processes. Y axis: TU, toxic unit that calculated as 100%/EC50 based on V. fisheri inhibition tests, with higher TU indicating more toxicity. X-axis: blank control (Blank), raw effluents before disinfection (Raw) and effluents disinfected with different processes.
10.4.3 Overall effluent toxicity revealed by quantitative toxicogenomics assay

We evaluated the overall effluent toxicity among various disinfection technologies using our quantitative stress response pathways ensemble-base toxicogenomics assay that include over 112 biomarkers and covers 5 stress response categories including DNA stress (DNA damage repair), oxidative stress (including ROS defense and redox balancing), protein stress, membrane (lipid stress and membrane transporter), and general stress (other stress as pH, temperature, et al. See Table 2.1). As described in methods section, the quantitative toxicogenomics assay measures the temporal gene expression levels of over 112 biomarkers indicative of various cellular stress responses, damage repair pathways. The overall molecular disturbance was quantified by TELI, which calculates the cumulative effects evoked upon toxicants exposure. At the original effluent concentration level (REF=1), the relative and comparative toxicity levels among the effluent showed some generally consistent trends as those revealed with Microtox assay. Both assay indicated that ozonation led to reduction in overall toxicity compared to the raw secondary effluent. UV-treatments, with or without H2O2 led to no significant changes in effluent toxicity. The blank control exhibits no significantly toxicity compare to assay detection limit (TELI=1.5). All the effluent samples except chlorination showed significantly elevated toxicity compared to blank control.

The toxicogenomics assay detected elevated toxicity in the effluents after chloramination treatment or UV-treatment at lower dose. Macova et al. (Macova et al., 2011) also observed an increase in toxicity after chloramination, which is
consistent with the result in this study. Ozone is reported that can remove baseline toxicity, estrogenicity and genotoxicity effectively (Misik et al., 2011, Macova et al., 2010; Escher et al., 2009), explaining the reduction in toxicity after ozonation observed in this study. However, it should be noted that oxidation by-products formed during ozonation are often more hydrophilic (e.g. NDMA) than their parent compound and therefore they might not be extracted efficiently by SPE [Macova et al., 2010; Staiter et al., 2011; Benner et al., 2009]. A noticeable point is that sample treated by chloramination exhibits the highest TELI_total value while the one treated by HOCl seemed to yield lowest TELI value, implying that NH2Cl is likely to induce substances that activate high magnitude of stress responses (demonstrated by TELI value) whereas HOCl may reduce them. The results correspond to our toxicity test that chlorination exhibits the better ability on toxicity removal.

In order to further evaluate the toxicity characteristics of various disinfected effluents and increase the probability to detect differences among them, concentrated samples are generally applied for toxicity assays (Stalter et al., 2010; Cao et al., 2009; Crosby et al., 2005). In this study, toxicogenomics evaluation is also performed with a more concentrated sample at REF=10, which yielded distinct results from those at original concentration, as shown in Figure 10.3. The toxicity of 10-times concentrated blank control still exhibits marginally detectable toxicity, suggesting again that the extraction procedure did not induce noticeable toxicity to samples. Consistent with results at native concentration, ozonation and chlorination processes remove the toxic substances in the effluents, resulting in
lower toxicity than raw effluents. The most pronounced observation with samples at REF of 10 is that NH2Cl introduces a significant elevation on TELI_total value. Both chlorine and chloramine oxidize other compounds in water and form similar DBPs, including THMs and HAAs. It is usually known that chloramination forms significant lower amount of THMs and HAAs than chlorination due to its relatively lower redox potential. Thus, chloramination is used as an alternative disinfectant to chlorine. However, in this case NH2Cl seems to generate transformation products, which are quite toxic that exert higher toxicity than those with HOCl treatment. As mentioned above, chloramination process is associated with the formation of multiple toxic emerging DBPs (iodinated THMs, HANs, haloketones, chloral hyrate, cyanogens chloride, chloropicrin and NDMA) (Yang et al., 2008) besides regulated ones. Many of those emerging DBPs are more toxic than regulated ones. For example, iodinated THMs have been predicted to be more toxic than chlorinated and brominated THMs (Richardson et al., 2007). HANs and other nitrogenous DBPs (e.g. NDMA) are also more toxic than regulated carbon-based DBPs, such as HAAs (Muellner et al., 2007; Bond et al., 2011). Therefore, it makes sense that the effluent sample treated by NH2Cl is more toxic than the one treated by chlorine.
Figure 10.3 Total TELI values for wastewater secondary effluent samples treated by various disinfection processes when REF=1 and REF=10. Y axis: TELI_{total}, toxicity index that calculated over entire stress response library, with higher value indicating more toxicity. X-axis: blank control (Blank), raw effluents before disinfection (Raw) and effluents disinfected with different processes. Blank control is the extraction control with DI water goes through SPE process, and raw effluent is the original secondary effluent.
UV is a common disinfection process used in wastewater treatment application as it’s highly effective at inactivating protozoans and could reduce the concentration of regulated DBPs (THMs and HAAs). Similar as REF=1, no significantly changes in toxicity were found in all UV-treated effluents, comparing with native effluents, suggests that UV process can produce transformation products that exhibit toxicity while the reduction of parent compounds. It’s different from the results observed in other researches, which show UV can reduce chronic cytotoxicity (Plewa et al., 2012) or acute toxicity (Escher et al., 2012). However, Liu et al. (Liu et al., 2002) reported that UV could induce the formation of aldehydes and carboxylic acids under certain conditions. So the formation of those emerging DBPs may lead to elevated toxicity.

10.4.4 Quantitative toxicogenomics assay revealed distinct toxicity profile and nature among effluents of different disinfection processes

Disinfection treatment technology-specific effluent toxicity fingerprints (pattern) were observed (Figure 10.4, Figure 10.5). Both UV treatments, at low or high doses, exhibited similar toxicity profiles, indicating similar effluent composition and matrices resulted from UV treatment. UV/H2O2, also yield high resemblance in their effluent toxicity fingerprints that are different from those with UV treatment alone, suggesting that addition of H2O2 affect the effluent toxicity, Compare with raw effluents, most of genes related to oxidative stress for effluent treated by NH2Cl are differentially expressed, leading to the oxidative stress response caused by NH2Cl is the strongest among all samples. Contrary, HOCl
and ozone reduce toxicity in raw effluent effectively. Especially, ozonation, both low and high dose, can remove toxicity overall stress categories, particularly for DNA damage category (i.e. genotoxicity). The ability of ozonation on genotoxicity removal is also verified by other previous resport (Misik et al., 2011; Escher et al., 2009). Comparing the expression profiles of genes related to DNA damage, 15 genes which are indicated as up-regulated in raw effluents switch to “not differentially expressed” genes ((TELI gene <1.5), indicating less DNA stress response. The lowest removal efficiency of ozonation happens to category of oxidative stress. Experimental evidence indicates that bromate, one major class of DBPs induced by ozonation, mediates toxicological effects via the induction of oxidative stress (Escher et al., 2009).

UV and UV based technology appeared to be most efficacious for attenuation of general stress and DNA stress. Previous reports.(Monarca et al., 2000; Haider et al., 2002) also observed lower genotoxicity after UV disinfection. Comparing the toxicity of UV and UV/H2O2 disinfected effluents, it could be found that the addition of H2O2 led to contrast effects on toxicity at different UV intensity. At lower level, toxicity increased for most of the stress categories, while toxicity reduced with higher UV intensity. Toor and Mohseni (Toor and Mohseni , 2007) observed that UV/H2O2 was effective in reducing DBP formation potential only when UV dose is higher than 1000 mJ/cm2. The raise of toxicity with the addition of H2O2 at UV 250 mJ/cm2 may because the photo-catalysis effect rather than photo- oxidation effect produced more hydrophobic/non-polar organics as the precursor of DBP formation (Liang et al., 2003)
HOCl appeared to be the most efficient treatment technology on toxicity removal, however deviates from other reported conclusions that chlorination increased acute toxicity and genotoxicity instead (Cao et al., 2009). Conversely, NH2Cl had less reduction effects on most of the stress categories, some of the activities did not decreased, even increased after the treatment, such as oxidative stress and protein stress, indicating limited removal efficiency or production of reactive transformation by-products during the disinfection process, which is consistent with the pattern of gene expression profiling. Holder et al. (Holder et al., 2013) used DNA microarray to investigate the global gene expression of E.coli cells exposed to monochloramine and got similar results as we did in this study. They found that genes induced by monochloramine were associated with several stress response functions, including oxidative stress, DNA damage, multidrug efflux, biofilm formation, antibiotic resistance and cell wall repair.

The TELI value reduction/increase during the treatment, which indicates the treatment efficacy on toxicity removal, was also visualized as a heatmap (Figure 10. 4). The similarity and correlations of stress response fingerprints among effluents with different disinfection processes were characterized by hierarchical clustering. Overall, water treatment efficiently eliminated part of the activity, although this depended both on treatment and stress categories. The results of hierarchical clustering (Figure 10.4) showed samples treated by UV or UV/ H2O2 clustering together and distributing far from other samples. The closest similarity existed between ozone 6.0 – ozone 1.5, UV 250 – UV250/H2O2, and NH2Cl and raw effluent, while blank control is further away in hierarchy distance.
Figure 10.4 Hierarchical clustering of toxicogenomics test results expressed as TELI for all water samples at REF=1. More red means more over expression on molecular stress response. Blank control is the extraction control with DI water goes through SPE process, and raw effluent is the original secondary effluent.

Figure 10.5 presents TELI values (TELI\textsubscript{pathway}) of each stress category for disinfected samples at REF=1 and REF=10. Similar as the results obtained from gene expression profiles. Effluent treated by NH\textsubscript{2}Cl shows strongest responses over all stress categories at both concentrations, especially oxidative stress. Ozonation processes, both low and high dose, demonstrate good toxicity removal efficiency (40%~80%) overall stress categories, particularly for DNA stress at REF=10 (i.e. genotoxicity). The lowest removal efficiency happens to category of oxidative stress. Chlorination exhibits similar toxicity removal efficiencies as ozonation. For all UV samples, the main elevated stress category at REF=10 is DNA stress compared to raw effluent. And addition of H\textsubscript{2}O\textsubscript{2} seems to affect the
removal of DNA stress. The potential inducers for the elevated DNA stress might 
some transformation byproducts, which be further oxidized or degraded to 
nontoxic products by UV/H$_2$O$_2$ but not by UV alone. This conjecture is also 
supported by other studies on comparison of UV and UV/H$_2$O$_2$ but from the 
perspective of chemical analysis (Kim et al., 2009; Yuan et al., 2011). Advanced 
oxidation process combined UV with H$_2$O$_2$ can reduce a wide range of organic 
contaminants in water by photolysis and oxidation (Pereira et al., 2007). Several 
studies (Thomson et al., 2004; Kleiser et al., 2000) reported that UV/ H$_2$O$_2$ can 
degrade aromatic NOM to smaller molecular size compounds, which are more 
biodegradable and less hydrophobic. The reduction in aromatic NOM and the 
entailed abatement in DBP formation potential may lead to the weaker DNA 
stress response than effluents disinfected by UV alone.
Figure 10.5 TELI values of each stress category for wastewater secondary effluent samples treated by various disinfection processes. Y-axis: TELI\textsubscript{pathway}, toxicity index that calculated for a single stress category, with higher value indicating more toxicity. X-axis: blank control (Blank), raw effluents before disinfection (Raw) and effluents disinfected with different processes.
10.5 Conclusion

In this study, we selected 8 different disinfected effluents, covering a broad range of different subsequent water treatment methods. All water extracts were screened with Microtox test as well as a novel quantitative toxicogenomics test. Previous works proved that for most of the chemicals, the effect concentration can be as low as nanomolar concentration or less, which indicate good sensitivity and make them plausible to be applied in environmental water samples where most compounds occurs at trace level. The result of this study showed that molecular toxicity with different stress responses was observed in all the water samples and the applied technologies could not completely remove all the stress responses. Some treatment, such as NH2Cl lead to increasing of toxicity after treatment, which may indicate a need of further investigation of possible transformation products. The derived toxicogenomics index TELI is a good indicator for water toxicity assessment. Therefore, the information this study has provided might be used to establish the toxicogenomics platform and approach that is suited and relevant for water quality assessment, and guides further toxicogenomics approach in water monitoring. The conclusion of this study could be used to guide the future selection of wastewater and drinking water treatment technologies.
Chapter 11

Conclusions and Future Work

11.1 Conclusions

The recognized and unknown health risks and the harmful environmental impacts associated with the ever-increasing number of emerging pollutants in our water presents a serious threat to us all. The current toxicological approach that rely on resource-intensive and laborious toxicity testing using laboratory animals can not meet the demands for assessing and quantifying the toxicity exerted by these pollutants, individually or as mixtures in various environment matrices. This poses a pressing need for a breakthrough in toxicity-assessment technology because the available methods are neither feasible nor sufficient to provide the timely information needed for regulatory decision making to eliminate these threats.

To contribute to the efforts in filling in the knowledge gap and addressing some key aspects of the discussed challenges in toxicity and hazards assessment, as well as in environmental monitoring, this study developed a novel, fast, cost-effective yet powerful quantitative toxicogenomics-based toxicity assessment platform for high-throughput and effective chemical hazardous identification and environmental toxicity monitoring. We systematically optimized the assay platform, evaluated its robustness and performance, validated the assay output and demonstrated its wide applications for environmental hazards assessment, water quality monitoring and remediation technology evaluation.
The main findings and novel contributions of this study are highlighted below:

1. This study have developed, validated a quantitative toxicogenomics technology platform that improved upon existing approaches and technologies in its sensitivity, feasibility, cost-effective and reliability. This technology platform can be employed for toxicity screening, prioritization, environmental sample toxicity assessment, technology effectiveness assessment, and environmental monitoring.

2. This study have proposed and developed a new molecular toxicity quantifier named Transcriptional Effect Level Index (TELI) to convert the information-rich toxicogenomic data into integrated and quantitative endpoints. This, for the first time, quantitated molecular response at individual gene, pathway or network level. The quantitative molecular endpoints allow the establishment of preliminary framework for molecular toxicity quantification, modeling and predictions.

3. This study explored and proposed a molecular mixture toxicity model framework and demonstrated its validity with a number of binary mixtures that exhibited additive, synergistic and antagonistic effects. This pioneered quantitative molecule toxicity modeling within the context of toxicogenomics paved the road for further toxicity identification and prediction.

4. This study applied the developed quantitative toxicogenomics approach for mechanistic toxicity evaluation, screening, and classification of CECs. The results provided an initial database of comprehensive molecular toxicity information for a large number and diverse categories of CECs, which can assist future QSAR and predictive toxicity model development.
5. This study proposed an approach for Mechanistic Toxicity Assessment of Nanomaterials using Whole-cell-array Stress Genes Expression Analysis. We demonstrated the application for toxicity assessment of nAg, nTIO2 and SWCNT. The results led to a better understanding of the mode of action (MOA) of metal and metal oxide nanomaterials.

6. This study performed detailed toxicity assessment of TiO2 nanoparticles anatase using global transcriptomics and network analysis. The information-rich toxicogenomics studies provide fundamental understanding and insights into the mechanisms involved in nanotoxicity exerted by nTiO2.

7. The study showed correlations of the molecular response endpoints derived from the developed quantitative toxicogenomics approach with conventional regulatory-relevant phenotypic endpoints in multiple species. The presence of “phenotype anchoring” provides evidence important for qualitative and quantitative model prediction of adverse outcomes that result from cellular disturbance measured via stress response pathways.

8. This study linked the derived toxicogenomics endpoints with the standard endpoints for water toxicity assessment. The developed quantitative toxicogenomics approach therefore can be adopted as an alternative standard method for water toxicity assessment and environmental monitoring. In addition, the toxicity results can be extrapolated or translated to other single cell organism in the ecological community.

9. This study demonstrated that the developed assay can be applied for revealing the toxicity dynamics and evolution during treatment processes and how different
remediation technology may impact whole effluent toxicity as results of targeted pollutants removal and risk-based technology efficacy assessment for risk reduction and minimization. The quantitative toxicogenomics assay may serve as a useful tool for remediation technology efficacy assessment, and provide guidance on process design and optimization for desired toxicity elimination and risk reduction.

10. This study demonstrated the application of the developed quantitative toxicogenomics approach to evaluate the changes in effluent toxicity level and profile through unit processes in a wastewater treatment plant. The resulted demonstrated that the developed quantitative toxicogenomics approach could serve as a cost-effective yet informative alternative method for water toxicity assessment and water quality monitoring.
11.2 Recommendations for Further Studies

- **Technology platform can be extended to other higher-level organisms**

In this study, we chose prokaryotic whole-cell-array-based toxicogenomic technique for environmental applications due to several advantages: (1) faster, easier, and less labor-intensive procedure for monitoring gene response in situ with GFP-infused E. coli strains that eliminate mRNA extraction, PCR amplification, labeling, or hybridization as needed on other transcriptomics technologies. (2) Most of the genome of E. coli has been well studied, and there is substantial information available on the functions of the genes, which allows for understanding, mapping, and visualizing systematic cellular response pathways and molecular events occurring as a response to chemical exposure. (3) Bacteria can proliferate rapidly under either aerobic or anaerobic conditions in very inexpensive media. (4) Finally, because stabilization and reusability are important factors for practical use, methods available for stabilizing, immobilizing the bacteria cells, and/or storing the freeze-dried biosensor enable practical application of this technique.

With the rapid advances in molecular biology, “omics” technology and data science, more GFP or other reporter cells become available, along with the growing genomic information and understanding of cellular responses networks. These advances will enable us to extend the similar toxicity assessment platform developed in this study to other organisms and cell lines. Particularly, we have focused on selected stress response genes that are present and highly conserved in most cell types of metazoans, therefore extension to other organisms or cell lines is reasonable.

- **Extrapolate the Principles and approach for molecular toxicity quantification and characterization to other “omics” technologies**
One of the most challenging aspects of implementing toxicogenomics-based toxicity assessment in regulation and in risk assessment involves establishing the theoretical framework and comprehensive knowledge base required to obtain quantitative toxicogenomics assay endpoints. It is mostly unclear and still under active research whether the molecular level disturbance by stressors can be quantified. We have proposed a new concept and computational method for converting the information-rich toxicogenomics data into integrated and quantitative endpoints-TELI (transcriptional effect level index) with prokaryotic E coli reporter cells. This study pioneered and demonstrated the validity and potential application of this concept for quantifying molecular perturbation in prokaryotic cells and linking it with phenotypic endpoints. The methodology and principal developed in this study can be expanded to other high-throughput approach (proteomics, metabolomics, battery of bioassays etc.) for more broader environmental samples assessment, such as soil, sediments, and air samples.

- **Further enlarge molecular toxicological database**

  This study initiated a database of comparable molecular toxicological data for 30 CECs of variety categories. The molecular toxicity database can be further expanded with more chemicals. A large toxicogenomics database, in combination of bioinformatics and statistical tools, will allow for further exploration and investigation on chemical toxicity identification and characterization, mechanism-based chemical classification, and possibly toxicity prediction for new chemicals.

- **Application of the quantitative HTS (you never used this term else where, please use the same term) for routine water quality monitoring**
We have demonstrated that the developed quantitative toxicity assay was able to provide effective monitoring of the overall toxicity of effluent and reveal the toxicity nature and dynamics along the wastewater treatment processes. Potential application of the assay for other water quality monitoring and warning including drinking water, reuse water, on-site water supply etc. can be further investigated. The low cost and high throughput nature of the assay makes it possible to be used for larger-scale water quality survey across geographical locations to generate mixture toxicity index-based evaluation and mapping of various water sources.

- **Application of the developed quantitative toxicogenomics approach as risk-based evaluation tool for technology efficacy assessment and optimization**

We demonstrate the feasibility and applicability of the developed quantitative toxicogenomics approach for evaluation of the toxicity evolution and nature along remediation process as well as for revealing how different disinfection technologies can impact effluent quality and toxicity, which could not be achieved with current conventional toxicity evaluation tools. The new, cost-effective and feasible bioavailability quantitative toxicogenomics approach will help facilitate design, optimization and implementation of remediation strategies and, therefore, improve science-based decision making for site management, priority-setting, and remedy selection.

- **Correlate the derived toxicogenomics endpoints to more regulatory relevant and phenotypic endpoints**

Our quantitative toxicogenomics approach integrates changes in gene or protein expression over time and magnitude to derive quantitative endpoints based on dose
response relationships, and connects them to standard water toxicity assessment endpoints. Correlation of molecular endpoints with regulatory relevant phenotypic endpoints can be further explored and evaluated with more organism and more different or specific toxicity endpoints (i.e. genotoxicity, baseline toxicity, oxidative stress) beyond cytotoxicity EC50.

Future attempts should be made to incorporate the toxicogenomics information from our assay into national programme (EPA ToxCast, CDC Human Exposure to Environmental Chemicals) to allow more systematics, cross-scale comparison, correlation and integrations analysis to assist in the realization of future Tox21 vision in implementing in vitro mechanic assay into water quality monitoring and environmental risk assessment framework.
REFERENCE


Adverse Outcome Pathway Wiki,


British Columbia Center for Disease Control, British Columbia Annual Summary of Antibiotics Utilization 2010, , August 2012.


Chhabra, R. S., Huff, J. E., Schwetz, B. S., & Selkirk, J. (1990). An overview of prechronic and chronic toxicity/carcinogenicity experimental study designs and criteria used by the National Toxicology Program. Environmental Health Perspectives, 86, 313.

Chinese Chemical Inventory of Existing Chemical Substances (IECSC). 2013


Das, A., Chakrabarty, S., Choudhury, D., & Chakrabarti, G. (2010). 1, 4-Benzoquinone (PBQ) induced toxicity in lung epithelial cells is mediated by the disruption of the microtubule network and activation of caspase-3. Chemical research in toxicology, 23(6), 1054-1066.


Global Ob-Gyn Group ,Global Ob-Gyn Group Urges Greater Efforts to Prevent Toxic Chemical Exposure.


Hansen, L. H., & Sørensen, S. J. (2001). The use of whole-cell biosensors to detect and quantify compounds or conditions affecting biological systems. Microbial ecology, 42(4), 483-494.


Korea Toxic Chemicals Control Act, 1991


Liu, K. J., & Shi, X. (2001). In vivo reduction of chromium (VI) and its related free radical generation. In Molecular Mechanisms of Metal Toxicity and Carcinogenesis (pp. 41-47). Springer US.


Ludewig, G., Dogra, S., & Glatt, H. (1989). Genotoxicity of 1, 4-benzoquinone and 1, 4-naphthoquinone in relation to effects on glutathione and NAD (P) H levels in V79 cells. Environmental health perspectives, 82, 223.


Madrigal, A. (2009). Humans have made, found or used over 50 million unique chemicals.


Microtox SOLO Manual” N.D. Retrieved 23 May 2014,

Microtox for sediment testing.” SDIX. 2010. Retrieved 23 May 2014,

Microtox 500: Industry-leading toxicity testing”

Microtox Toxicity Testing” Leeder Consulting. N.D. Retrieved 28 May 2014


NRC, TOXICITY TESTING, supra note 1.


Public Law 92-500 (1972), The Clean Water Act (CWA) United States Federal Water Pollution Control Act Amendments,


Safe, S. (1990). Polychlorinated biphenyls (PCBs), dibenzo-p-dioxins (PCDDs), dibenzofurans (PCDFs), and related compounds: environmental and mechanistic considerations which support the development of toxic equivalency factors (TEFs). Critical reviews in toxicology, 21(1), 51-88.


SDIX, How to use the Microtox Acute Toxicity Test to perform an In-House Toxicity Reduction Evaluation (TRE)”. 2006 Retrieved 23 May 2014


USEPA, Drinking Water Contaminant Candidate List (CCL) and Regulatory Determination, available at: http://www.epa.gov/ccl


USEPA, Method for Aquatic Toxicity Identification Evaluation, 1991

USEPA, Methods for Aquatic Toxicity Identification Evaluation (TIE), Washington DC, 1991


USEPA, TOXIC SUBSTANCES CONTROL ACT, Washington, DC, 2002


We Interact with 100,000+ Chemicals, and the Dangers Are Barely Understood, http://www.alternet.org/story/143130/we_interact_with_100,000%2B_chemicals,_and_the_dangers_are_barely_understood


## APPENDIX

### Appendix 4.1 Detailed resources and purity information for tested chemical

<table>
<thead>
<tr>
<th>Group</th>
<th>Chemical</th>
<th>Vendor</th>
<th>Purity/Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heavy metal</td>
<td>Arsenic (VI), reference standard solution</td>
<td>Fisher Chemical</td>
<td>1000 ppm</td>
</tr>
<tr>
<td></td>
<td>Chromium (VI), reference standard solution</td>
<td>Fisher Chemical</td>
<td>1000 ppm</td>
</tr>
<tr>
<td></td>
<td>Lead, reference standard solution</td>
<td>Fisher Chemical</td>
<td>1000 ppm</td>
</tr>
<tr>
<td>Antibiotics</td>
<td>Ciprofloxacin HCl</td>
<td>Mediatech, Inc</td>
<td>&gt;98%</td>
</tr>
<tr>
<td></td>
<td>Erythromycin</td>
<td>MP BIOMEDICALS</td>
<td>~98%</td>
</tr>
<tr>
<td></td>
<td>Mitomycin C</td>
<td>Fisher BioReagents</td>
<td>&gt;98%</td>
</tr>
<tr>
<td>Pharmaceuticals</td>
<td>Atovastatin</td>
<td>TOCRIS</td>
<td>&gt;99%</td>
</tr>
<tr>
<td></td>
<td>Ibuprofen</td>
<td>Acros organics</td>
<td>99%</td>
</tr>
<tr>
<td></td>
<td>Metoprolol tartrate</td>
<td>MP BIOMEDICALS</td>
<td>99.5%</td>
</tr>
<tr>
<td>Drinking water byproduct</td>
<td>N-Nitrosodimethylamine</td>
<td>SPEX CertiPrep™</td>
<td>1000 µg/mL</td>
</tr>
<tr>
<td></td>
<td>Bromodichloromethane</td>
<td>Acros organics</td>
<td>98%</td>
</tr>
<tr>
<td></td>
<td>Trihalomethane mix (Bromodichloromethane, Bromoform, Chloroform, Dibromochloromethane)</td>
<td>Restek</td>
<td>2,000 µg/mL each in P&amp;T methanol</td>
</tr>
<tr>
<td></td>
<td>Nitrosamine mix (N-Nitrosodiethylamine, N-Nitrosodimethylamine, N-Nitrosodi-n-butylamine, N-Nitrosodi-n-propylamine, N-Nitrosomethylphethylamine, N-Nitrosopiperidine, N-Nitrosopyrrolidine)</td>
<td>Restek</td>
<td>1,000 µg/mL each in methylene chloride</td>
</tr>
<tr>
<td>Industrial and food additive</td>
<td>Bisphenol A</td>
<td>Sigma-Aldrich</td>
<td>&gt;99%</td>
</tr>
<tr>
<td></td>
<td>Surcolase</td>
<td>AK Scientific, Inc.</td>
<td>98%</td>
</tr>
<tr>
<td></td>
<td>4-n-Nonylphenol</td>
<td>Sigma-Aldrich</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bis(2-ethylhexyl) phthalate</td>
<td>Restek</td>
<td>1000 mg/mL in Methylene Chloride</td>
</tr>
<tr>
<td></td>
<td>Caffeine</td>
<td>Alfa Aesar</td>
<td>99.7%</td>
</tr>
<tr>
<td>Herbicide and insecticide</td>
<td>5-CHLORO-2-(2,4-DICHLOROPHENOXY)PHENOL</td>
<td>Calbiochem</td>
<td>&gt;97%</td>
</tr>
<tr>
<td></td>
<td>Carbaryl</td>
<td>Sigma-Aldrich</td>
<td>100 ng/µL in</td>
</tr>
<tr>
<td></td>
<td>Compound</td>
<td>Supplier</td>
<td>Purity/Other Details</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------</td>
<td>--------------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>2-chloro-4-(ethylamino)-6-(isopropylamino)-s-triazine</td>
<td>SPEX CertiPrep™</td>
<td>1000µg/mL in Acetone</td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>Hydrogen peroxide</td>
<td>Fisher Chemical</td>
<td>30%</td>
</tr>
<tr>
<td></td>
<td>Trichloroacetic acid</td>
<td>Fisher BioReagents</td>
<td>&gt;99%</td>
</tr>
<tr>
<td></td>
<td>17-beta-estradiol</td>
<td>MP BIOMEDICALS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Phenol</td>
<td>Acros organics</td>
<td>99%</td>
</tr>
<tr>
<td></td>
<td>BAP+S9</td>
<td>Sigma-Aldrich</td>
<td>&gt;96%</td>
</tr>
<tr>
<td>Nanomaterials</td>
<td>nAg</td>
<td>NanoDynamics Inc.,</td>
<td>~60nm</td>
</tr>
<tr>
<td></td>
<td>nTiO2_a</td>
<td>NanoStructured &amp; Amorphous Materials</td>
<td>~10nm</td>
</tr>
<tr>
<td></td>
<td>nTiO2_r</td>
<td>Sigma-Aldrich</td>
<td>~10 nm thick, 40 nm laterally,</td>
</tr>
<tr>
<td></td>
<td>Fullerene soot</td>
<td>M.E.R.Co</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 4.2 Freely dissolved aqueous concentrations and concentrations in cells/tissues for selected chemicals.

<table>
<thead>
<tr>
<th>Group</th>
<th>Chemical</th>
<th>Designed C. (M)</th>
<th>Freely dissolved C. (M)</th>
<th>C. for cells/tissues (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antibiotics</strong></td>
<td>Ciprofloxacin HCl</td>
<td>3.02*10^{-13}</td>
<td>2.33*10^{-13}</td>
<td>5.25*10^{-14}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.02*10^{-12}</td>
<td>2.33*10^{-12}</td>
<td>5.25*10^{-13}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.02*10^{-11}</td>
<td>2.33*10^{-11}</td>
<td>5.25*10^{-12}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.02*10^{-10}</td>
<td>2.33*10^{-10}</td>
<td>5.25*10^{-11}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.02*10^{-9}</td>
<td>2.33*10^{-9}</td>
<td>5.25*10^{-10}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.02*10^{-8}</td>
<td>2.33*10^{-8}</td>
<td>5.25*10^{-9}</td>
</tr>
<tr>
<td><strong>Mitomycin C</strong></td>
<td></td>
<td>1.50*10^{-11}</td>
<td>1.48*10^{-11}</td>
<td>3.19*10^{-12}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.50*10^{-10}</td>
<td>1.48*10^{-10}</td>
<td>3.19*10^{-14}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.50*10^{-9}</td>
<td>1.48*10^{-9}</td>
<td>3.19*10^{-13}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.50*10^{-8}</td>
<td>1.48*10^{-8}</td>
<td>3.19*10^{-12}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.50*10^{-7}</td>
<td>1.48*10^{-7}</td>
<td>3.19*10^{-11}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.50*10^{-6}</td>
<td>1.48*10^{-6}</td>
<td>3.19*10^{-10}</td>
</tr>
<tr>
<td><strong>Ibuprofen</strong></td>
<td></td>
<td>4.85*10^{-9}</td>
<td>3.10*10^{-12}</td>
<td>8.05*10^{-9}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.85*10^{-8}</td>
<td>3.10*10^{-11}</td>
<td>8.05*10^{-8}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.85*10^{-7}</td>
<td>3.10*10^{-10}</td>
<td>8.05*10^{-7}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.85*10^{-6}</td>
<td>3.10*10^{-9}</td>
<td>8.05*10^{-6}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.85*10^{-5}</td>
<td>3.10*10^{-8}</td>
<td>8.05*10^{-5}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.85*10^{-4}</td>
<td>3.10*10^{-7}</td>
<td>8.05*10^{-4}</td>
</tr>
<tr>
<td><strong>Drinking water byproduct</strong></td>
<td>N-Nitrosodimethylamine</td>
<td>1.89*10^{-8}</td>
<td>1.83*10^{-8}</td>
<td>1.02*10^{-9}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.89*10^{-7}</td>
<td>1.83*10^{-7}</td>
<td>1.02*10^{-8}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.89*10^{-6}</td>
<td>1.83*10^{-6}</td>
<td>1.02*10^{-7}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.89*10^{-5}</td>
<td>1.83*10^{-5}</td>
<td>1.02*10^{-6}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.89*10^{-4}</td>
<td>1.83*10^{-4}</td>
<td>1.02*10^{-5}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.89*10^{-3}</td>
<td>1.83*10^{-3}</td>
<td>1.02*10^{-4}</td>
</tr>
<tr>
<td><strong>Bromodichloromethane</strong></td>
<td></td>
<td>8.55*10^{-8}</td>
<td>1.06*10^{-8}</td>
<td>1.23*10^{-7}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.55*10^{-7}</td>
<td>1.06*10^{-7}</td>
<td>1.23*10^{-6}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.55*10^{-6}</td>
<td>1.06*10^{-6}</td>
<td>1.23*10^{-5}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.55*10^{-5}</td>
<td>1.06*10^{-5}</td>
<td>1.23*10^{-4}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.55*10^{-4}</td>
<td>1.06*10^{-4}</td>
<td>1.23*10^{-3}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.55*10^{-3}</td>
<td>1.06*10^{-3}</td>
<td>1.23*10^{-2}</td>
</tr>
<tr>
<td><strong>Industrial and food additive</strong></td>
<td>Bisphenol A</td>
<td>4.38*10^{-9}</td>
<td>9.14*10^{-12}</td>
<td>7.27*10^{-9}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.38*10^{-8}</td>
<td>9.14*10^{-11}</td>
<td>7.27*10^{-8}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.38*10^{-7}</td>
<td>9.14*10^{-10}</td>
<td>7.27*10^{-7}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.38*10^{-6}</td>
<td>9.14*10^{-9}</td>
<td>7.27*10^{-6}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.38*10^{-5}</td>
<td>9.14*10^{-8}</td>
<td>7.27*10^{-5}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.38*10^{-4}</td>
<td>9.14*10^{-7}</td>
<td>7.27*10^{-4}</td>
</tr>
<tr>
<td><strong>Surcolase</strong></td>
<td></td>
<td>8.76*10^{-7}</td>
<td>7.04*10^{-10}</td>
<td>1.46*10^{-6}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.76*10^{-6}</td>
<td>7.04*10^{-9}</td>
<td>1.46*10^{-5}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.76*10^{-5}</td>
<td>7.04*10^{-8}</td>
<td>1.46*10^{-4}</td>
</tr>
<tr>
<td>Chemical</td>
<td>10^-2</td>
<td>10^-1</td>
<td>10^0</td>
<td></td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td><strong>8.76</strong></td>
<td>7.04</td>
<td>1.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-4</strong></td>
<td>7.04</td>
<td>1.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-3</strong></td>
<td>7.04</td>
<td>1.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-2</strong></td>
<td>7.04</td>
<td>1.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>4-n-Nonylphenol</strong></td>
<td>1.27</td>
<td>2.76</td>
<td>2.11</td>
<td></td>
</tr>
<tr>
<td><strong>10^-9</strong></td>
<td>2.76</td>
<td>2.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-8</strong></td>
<td>2.76</td>
<td>2.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-7</strong></td>
<td>2.76</td>
<td>2.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-6</strong></td>
<td>2.76</td>
<td>2.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-5</strong></td>
<td>2.76</td>
<td>2.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-4</strong></td>
<td>2.76</td>
<td>2.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Caffeine</strong></td>
<td>1.03</td>
<td>8.73</td>
<td>2.56</td>
<td></td>
</tr>
<tr>
<td><strong>10^-8</strong></td>
<td>8.73</td>
<td>2.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-7</strong></td>
<td>8.73</td>
<td>2.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-6</strong></td>
<td>8.73</td>
<td>2.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-5</strong></td>
<td>8.73</td>
<td>2.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-4</strong></td>
<td>8.73</td>
<td>2.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Herbicide and insecticide</strong></td>
<td>3.47</td>
<td>2.41</td>
<td>5.78</td>
<td></td>
</tr>
<tr>
<td><strong>5-CHLORO-2-(2,4-DICHLOROPHENOX Y)PHENOL</strong></td>
<td>2.41</td>
<td>5.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-8</strong></td>
<td>2.41</td>
<td>5.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-7</strong></td>
<td>2.41</td>
<td>5.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-6</strong></td>
<td>2.41</td>
<td>5.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-5</strong></td>
<td>2.41</td>
<td>5.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-4</strong></td>
<td>2.41</td>
<td>5.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-3</strong></td>
<td>2.41</td>
<td>5.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Carbaryl</strong></td>
<td>4.97</td>
<td>1.10</td>
<td>8.05</td>
<td></td>
</tr>
<tr>
<td><strong>10^-10</strong></td>
<td>1.10</td>
<td>8.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-9</strong></td>
<td>1.10</td>
<td>8.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-8</strong></td>
<td>1.10</td>
<td>8.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-7</strong></td>
<td>1.10</td>
<td>8.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-6</strong></td>
<td>1.10</td>
<td>8.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>2-chloro-4-(ethylamino)-6-(isopropylamino)s-triazine</strong></td>
<td>7.02</td>
<td>9.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-12</strong></td>
<td>7.02</td>
<td>9.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-11</strong></td>
<td>7.02</td>
<td>9.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-10</strong></td>
<td>7.02</td>
<td>9.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-9</strong></td>
<td>7.02</td>
<td>9.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-8</strong></td>
<td>7.02</td>
<td>9.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-7</strong></td>
<td>7.02</td>
<td>9.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Trichloroacetic acid</strong></td>
<td>4.94</td>
<td>8.68</td>
<td>6.71</td>
<td></td>
</tr>
<tr>
<td><strong>10^-7</strong></td>
<td>8.68</td>
<td>6.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-6</strong></td>
<td>8.68</td>
<td>6.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-5</strong></td>
<td>8.68</td>
<td>6.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-4</strong></td>
<td>8.68</td>
<td>6.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-3</strong></td>
<td>8.68</td>
<td>6.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-2</strong></td>
<td>8.68</td>
<td>6.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>17-beta-estradiol</strong></td>
<td>3.67</td>
<td>1.61</td>
<td>6.11</td>
<td></td>
</tr>
<tr>
<td><strong>10^-9</strong></td>
<td>1.61</td>
<td>6.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-8</strong></td>
<td>1.61</td>
<td>6.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-7</strong></td>
<td>1.61</td>
<td>6.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-6</strong></td>
<td>1.61</td>
<td>6.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-5</strong></td>
<td>1.61</td>
<td>6.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>10^-4</strong></td>
<td>1.61</td>
<td>6.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenol</td>
<td>(1.49 \times 10^{-6})</td>
<td>(2.51 \times 10^{-7})</td>
<td>(2.03 \times 10^{-5})</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.49 \times 10^{-5})</td>
<td>(2.51 \times 10^{-6})</td>
<td>(2.03 \times 10^{-3})</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.49 \times 10^{-4})</td>
<td>(2.51 \times 10^{-5})</td>
<td>(2.03 \times 10^{-4})</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.49 \times 10^{-3})</td>
<td>(2.51 \times 10^{-4})</td>
<td>(2.03 \times 10^{-3})</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.49 \times 10^{-2})</td>
<td>(2.51 \times 10^{-3})</td>
<td>(2.03 \times 10^{-2})</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.49 \times 10^{-1})</td>
<td>(2.51 \times 10^{-2})</td>
<td>(2.03 \times 10^{-1})</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 4.3 Toxicity profile changes based on the stress response pathway ensemble for the tested chemicals with different concentrations. X-axis: list of stress response categories and selected genes (see STable1). Altered gene expression changes in relative to the untreated control were indicated as the TELI value for each individual gene and color-coded with the scale 0-5 (Red spectrum colors indicate the magnitude of the altered gene expression, while black indicates no change). Y-axis left: exposure concentrations in mg/L. Those stress response categories highlighted with blue boxes are significantly (p<0.05) affected, with green boxes are induced (0.5<p<0.2) based on the gene enrichment analysis.

**Arsenic**

<table>
<thead>
<tr>
<th>Concentration (µM)</th>
<th>Oxidative Stress</th>
<th>Protein Stress</th>
<th>Membrane Stress</th>
<th>General Stress</th>
<th>DNA Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.87*10^-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.87*10^-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.87*10^-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>187.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Chromium**

<table>
<thead>
<tr>
<th>Concentration (µM)</th>
<th>Oxidative Stress</th>
<th>Protein Stress</th>
<th>Membrane Stress</th>
<th>General Stress</th>
<th>DNA Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.69*10^-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.69*10^-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.69*10^-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26.92</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>269.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Lead

<table>
<thead>
<tr>
<th>Concentration (µM)</th>
<th>Oxidative Stress</th>
<th>Protein Stress</th>
<th>Membrane Stress</th>
<th>General Stress</th>
<th>DNA Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.03*10⁻⁴</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.03*10⁻³</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.03*10⁻²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.03*10⁻¹</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Ciprofloxacin

<table>
<thead>
<tr>
<th>Concentration (µM)</th>
<th>Oxidative Stress</th>
<th>Protein Stress</th>
<th>Membrane Stress</th>
<th>General Stress</th>
<th>DNA Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.02*10⁻⁷</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.02*10⁻⁶</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.02*10⁻⁵</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.02*10⁻⁴</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.02*10⁻³</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.02*10⁻²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Erythromycin

<table>
<thead>
<tr>
<th>Concentration (µM)</th>
<th>Oxidative Stress</th>
<th>Protein Stress</th>
<th>Membrane Stress</th>
<th>General Stress</th>
<th>DNA Stress</th>
</tr>
</thead>
</table>

---

**Note:** The tables and diagrams represent the concentration of various substances across different stress categories, indicating levels of stress for Oxidative, Membrane, Protein, General, and DNA stress at different concentrations.
<table>
<thead>
<tr>
<th>Concentration (µM)</th>
<th>Oxidative Stress</th>
<th>Protein Stress</th>
<th>Membrane Stress</th>
<th>General Stress</th>
<th>DNA Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.79*10^{-3}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.79*10^{-2}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.79*10^{-1}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17.92</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>179.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1792.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Ibuprofen**

<table>
<thead>
<tr>
<th>Concentration (µM)</th>
<th>Oxidative Stress</th>
<th>Protein Stress</th>
<th>Membrane Stress</th>
<th>General Stress</th>
<th>DNA Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.85*10^{-3}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.85*10^{-2}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.85*10^{-1}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.85</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>485.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Metoprolol**
### N-Nitrosodimethylamine

<table>
<thead>
<tr>
<th>Concentration (µM)</th>
<th>Oxidative Stress</th>
<th>Protein Stress</th>
<th>Membrane Stress</th>
<th>General Stress</th>
<th>DNA Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.74*10^{-3}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.74*10^{-2}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.74*10^{-1}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>374.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Bromodichloromethane

<table>
<thead>
<tr>
<th>Concentration (µM)</th>
<th>Oxidative Stress</th>
<th>Protein Stress</th>
<th>Membrane Stress</th>
<th>General Stress</th>
<th>DNA Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.91*10^{-2}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.91*10^{-1}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>191.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1918.92</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Trihalomethane mix**

**Nitrosamine**

**Bisphenol A**
Sucralose

4-n-Nonylphenol
### Bis(2-ethylhexyl) phthalate

<table>
<thead>
<tr>
<th>Concentration (µM)</th>
<th>Oxidative Stress</th>
<th>Protein Stress</th>
<th>Membrane Stress</th>
<th>General Stress</th>
<th>DNA Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.27*10^{-3}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.27*10^{-2}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.27*10^{-1}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.72</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>127.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Caffeine

<table>
<thead>
<tr>
<th>Concentration (µM)</th>
<th>Oxidative Stress</th>
<th>Protein Stress</th>
<th>Membrane Stress</th>
<th>General Stress</th>
<th>DNA Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.56*10^{-2}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.56*10^{-1}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>256.41</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2564.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5-CHLORO-2-(2,4-DICHLOROPHENOXY)PHENOL

Concentration (µM)

Oxidative Stress  Protein Stress  Membrane Stress  General Stress  DNA Stress

1.03*10^{-2}  
1.03*10^{-1}  
1.03  
10.30  
103.09  
1030.92  

Carbaryl
<table>
<thead>
<tr>
<th>Concentration (µM)</th>
<th>Oxidative Stress</th>
<th>Protein Stress</th>
<th>Membrane Stress</th>
<th>General Stress</th>
<th>DNA Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.97*10^-4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.97*10^-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.97*10^-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.97*10^-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Atrazine**

<table>
<thead>
<tr>
<th>Concentration (µM)</th>
<th>Oxidative Stress</th>
<th>Protein Stress</th>
<th>Membrane Stress</th>
<th>General Stress</th>
<th>DNA Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.81*10^-6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.81*10^-5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.81*10^-4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.81*10^-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.81*10^-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.81*10^-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Hydrogen peroxide**
**Trichloroacetic acid**

<table>
<thead>
<tr>
<th>Concentration (µM)</th>
<th>Oxidative Stress</th>
<th>Protein Stress</th>
<th>Membrane Stress</th>
<th>General Stress</th>
<th>DNA Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>149</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>296</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>593</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1186</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2352</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**17-beta-estradiol**

<table>
<thead>
<tr>
<th>Concentration (µM)</th>
<th>Oxidative Stress</th>
<th>Protein Stress</th>
<th>Membrane Stress</th>
<th>General Stress</th>
<th>DNA Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.90*10^-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>490.79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4907.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49079.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concentration (µM)</td>
<td>Oxidative Stress</td>
<td>Protein Stress</td>
<td>Membrane Stress</td>
<td>General Stress</td>
<td>DNA Stress</td>
</tr>
<tr>
<td>--------------------</td>
<td>------------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
<td>-----------</td>
</tr>
<tr>
<td>3.67*10^-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.67*10^-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.67*10^-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.67</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>367.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Phenol**

<table>
<thead>
<tr>
<th>Concentration (µM)</th>
<th>Oxidative Stress</th>
<th>Protein Stress</th>
<th>Membrane Stress</th>
<th>General Stress</th>
<th>DNA Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>148.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1489.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14893.61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>148936.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Benz(a)pyrene**
Appendix 4.4 Selected (>4 out 6) overrepresented \((p<0.05)\) GO biological process annotations for tested CECs among 6 concentrations.

<table>
<thead>
<tr>
<th>Group</th>
<th>Chemical</th>
<th>Identified Gene Ontology</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>GO term</td>
</tr>
<tr>
<td>Heavy metal</td>
<td>Arsenic (VI)</td>
<td>GO: 302</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 303</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 305</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>GO: 6139</td>
<td>nucleic acid metabolic process</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DNA metabolic process</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DNA repair</td>
<td></td>
</tr>
<tr>
<td>GO: 6259</td>
<td>Base-excision repair</td>
<td></td>
</tr>
<tr>
<td>GO: 6281</td>
<td>Base-excision repair, AP site formation</td>
<td></td>
</tr>
<tr>
<td>GO: 6284</td>
<td>Nucleotide-excision repair</td>
<td></td>
</tr>
<tr>
<td>GO: 6285</td>
<td>Transcription</td>
<td></td>
</tr>
<tr>
<td>GO: 6289</td>
<td>Transcription, DNA-dependent</td>
<td></td>
</tr>
<tr>
<td>GO: 6350</td>
<td>Regulation of transcription, DNA-dependent</td>
<td></td>
</tr>
<tr>
<td>GO: 6351</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 6355</td>
<td>Oxygen and reactive oxygen species metabolic process</td>
<td></td>
</tr>
<tr>
<td>GO: 6800</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 6950</td>
<td>Response to stress</td>
<td></td>
</tr>
<tr>
<td>GO: 6974</td>
<td>Response to DNA damage stimulus</td>
<td></td>
</tr>
<tr>
<td>GO: 6979</td>
<td>Response to oxidative stress</td>
<td></td>
</tr>
<tr>
<td>GO: 6986</td>
<td>Response to unfolded protein</td>
<td></td>
</tr>
<tr>
<td>GO: 7154</td>
<td>Cell communication</td>
<td></td>
</tr>
<tr>
<td>GO: 9059</td>
<td>Macromolecule biosynthetic process</td>
<td></td>
</tr>
<tr>
<td>GO: 9266</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 9408</td>
<td>Response to temperature stimulus</td>
<td></td>
</tr>
<tr>
<td>GO: 9432</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 9605</td>
<td>Response to heat</td>
<td></td>
</tr>
<tr>
<td>GO: 9628</td>
<td>SOS response</td>
<td></td>
</tr>
<tr>
<td>GO: 9889</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 9991</td>
<td>Regulation of gene expression</td>
<td></td>
</tr>
<tr>
<td>GO: 10035</td>
<td>Regulation of macromolecular biosynthetic process</td>
<td></td>
</tr>
<tr>
<td>GO: 10468</td>
<td>RNA metabolic process</td>
<td></td>
</tr>
<tr>
<td>GO: 10556</td>
<td>Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic</td>
<td></td>
</tr>
<tr>
<td></td>
<td>process</td>
<td></td>
</tr>
<tr>
<td>GO: 16070</td>
<td>Regulation of metabolic process</td>
<td></td>
</tr>
<tr>
<td>GO: 19219</td>
<td>Regulation of cellular metabolic process</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Regulation of cellular biosynthetic process</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cellular response to external stimulus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RNA biosynthetic process</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Response to hydrogen peroxide</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cellular response to stress</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cellular response to oxidative stress</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cellular response to reactive oxygen species</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cellular macromolecule biosynthetic process</td>
<td></td>
</tr>
<tr>
<td>GO:</td>
<td>Response to chemical stimulus</td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-------------------------------</td>
<td></td>
</tr>
<tr>
<td>31326</td>
<td>Response to drug</td>
<td></td>
</tr>
<tr>
<td>GO:</td>
<td>Macromolecular metabolic process</td>
<td></td>
</tr>
<tr>
<td>31668</td>
<td>Protein unfolding</td>
<td></td>
</tr>
<tr>
<td>GO:</td>
<td>Cellular macromolecule metabolic process</td>
<td></td>
</tr>
<tr>
<td>32774</td>
<td>DNA replication and proofreading</td>
<td></td>
</tr>
<tr>
<td>GO:</td>
<td>Maintenance of fidelity involved in DNA dependent DNA replication</td>
<td></td>
</tr>
<tr>
<td>33194</td>
<td>Regulation of transcription</td>
<td></td>
</tr>
<tr>
<td>33554</td>
<td>Cell redox homeostasis</td>
<td></td>
</tr>
<tr>
<td>GO:</td>
<td>Regulation of biological process</td>
<td></td>
</tr>
<tr>
<td>34599</td>
<td>Regulation of cellular process</td>
<td></td>
</tr>
<tr>
<td>GO:</td>
<td>Response to stimulus</td>
<td></td>
</tr>
<tr>
<td>34614</td>
<td>Regulation of nitrogen compound metabolic process</td>
<td></td>
</tr>
<tr>
<td>GO:</td>
<td>Regulation of RNA metabolic process</td>
<td></td>
</tr>
<tr>
<td>34645</td>
<td>Cellular response to stimulus</td>
<td></td>
</tr>
<tr>
<td>GO:</td>
<td>Response to protein stimulus</td>
<td></td>
</tr>
<tr>
<td>42221</td>
<td>Regulation of macromolecule metabolic process</td>
<td></td>
</tr>
<tr>
<td>42493</td>
<td>Biological regulation</td>
<td></td>
</tr>
<tr>
<td>GO:</td>
<td>Cellular response to oxygen radical</td>
<td></td>
</tr>
<tr>
<td>43170</td>
<td>Cellular response to external stimulus</td>
<td></td>
</tr>
<tr>
<td>GO:</td>
<td>Regulation of primary metabolic process</td>
<td></td>
</tr>
<tr>
<td>43335</td>
<td>Nucleic acid metabolic process</td>
<td></td>
</tr>
<tr>
<td>44260</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>45004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>45005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>45449</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>45454</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50789</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50794</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50896</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>51171</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 51252</td>
<td>Response to reactive oxygen species</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------------------------</td>
<td></td>
</tr>
<tr>
<td>GO: 51716</td>
<td>Response to superoxide</td>
<td></td>
</tr>
<tr>
<td>GO: 51789</td>
<td>Response to oxygen radical</td>
<td></td>
</tr>
<tr>
<td>GO: 60255</td>
<td>Nucleobase, nucleoside, nucleotide and nucleic acid metabolic</td>
<td></td>
</tr>
<tr>
<td>GO: 65007</td>
<td>DNA metabolic process</td>
<td></td>
</tr>
<tr>
<td>GO: 71450</td>
<td>DNA repair</td>
<td></td>
</tr>
<tr>
<td>GO: 71496</td>
<td>Base-excision repair</td>
<td></td>
</tr>
<tr>
<td>GO: 80090</td>
<td>Transcription</td>
<td></td>
</tr>
<tr>
<td>GO: 90304</td>
<td>Transcription, DNA-dependent</td>
<td></td>
</tr>
<tr>
<td>Chromium (VI)</td>
<td>Regulation of transcription, DNA-dependent</td>
<td></td>
</tr>
<tr>
<td>GO: 302</td>
<td>Oxygen and reactive oxygen species metabolic process</td>
<td></td>
</tr>
<tr>
<td>GO: 303</td>
<td>Cellular aldehyde metabolic process</td>
<td></td>
</tr>
<tr>
<td>GO: 305</td>
<td>Nitrogen compound metabolic process</td>
<td></td>
</tr>
<tr>
<td>GO: 6139</td>
<td>Response to stress</td>
<td></td>
</tr>
<tr>
<td>GO: 6259</td>
<td>Response to DNA damage stimulus</td>
<td></td>
</tr>
<tr>
<td>GO: 6281</td>
<td>Response to oxidative stress</td>
<td></td>
</tr>
<tr>
<td>GO: 6284</td>
<td>Cell communication</td>
<td></td>
</tr>
<tr>
<td>GO: 6350</td>
<td>Response to temperature stimulus</td>
<td></td>
</tr>
<tr>
<td>GO: 6351</td>
<td>SOS response</td>
<td></td>
</tr>
<tr>
<td>GO: 6355</td>
<td>Response to external stimulus</td>
<td></td>
</tr>
<tr>
<td>GO: 6800</td>
<td>Regulation of biosynthetic process</td>
<td></td>
</tr>
<tr>
<td>GO: 6801</td>
<td>Response to extracellular stimulus</td>
<td></td>
</tr>
<tr>
<td>GO: 6807</td>
<td>Regulation of gene expression</td>
<td></td>
</tr>
<tr>
<td>GO: 6950</td>
<td>Regulation of macromolecular biosynthetic process</td>
<td></td>
</tr>
<tr>
<td>GO: 10556</td>
<td>Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>GO: 19219</td>
<td>Regulation of metabolic process</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Removal of superoxide radicals</td>
<td></td>
</tr>
<tr>
<td>GO: 19222</td>
<td>Regulation of cellular biosynthetic process</td>
<td></td>
</tr>
<tr>
<td>GO: 19430</td>
<td>Cellular response to external stimulus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RNA biosynthetic process</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cellular response to stress</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cellular response to oxidative stress</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cellular response to reactive oxygen species</td>
<td></td>
</tr>
<tr>
<td>GO: 31326</td>
<td>Response to drug</td>
<td></td>
</tr>
<tr>
<td>GO: 31668</td>
<td>Macromolecular metabolic process</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cellular macromolecule metabolic process</td>
<td></td>
</tr>
<tr>
<td>GO: 32774</td>
<td>Regulation of transcription</td>
<td></td>
</tr>
<tr>
<td>GO: 33554</td>
<td>Regulation of biological process</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Regulation of cellular process</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Response to stimulus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Regulation of nitrogen compound metabolic process</td>
<td></td>
</tr>
<tr>
<td>GO: 34614</td>
<td>Regulation of RNA metabolic process</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cellular response to stimulus</td>
<td></td>
</tr>
<tr>
<td>GO: 42493</td>
<td>Regulation of macromolecule metabolic process</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Biological regulation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cellular response to chemical stimulus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cellular response to oxygen radical</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cellular response to superoxide</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cellular response to external stimulus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Regulation of primary metabolic process</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nucleic acid metabolic process</td>
<td></td>
</tr>
</tbody>
</table>

| GO: 50789 |
| GO: 50794 |
| GO: 50896 |
| GO: 51171 |

<p>| GO: 51252 |
| GO:       |
| Lead | GO: 302 | Response to reactive oxygen species |
|      | GO: 6800 | Oxygen and reactive oxygen species metabolic process |
|      | GO: 6950 | Response to stress |
|      | GO: 6970 | Response to osmotic stress |
|      | GO: 6974 | Response to DNA damage stimulus |
|      | GO: 6979 | Response to oxidative stress |
|      | GO: 9628 | Response to abiotic stimulus |
|      | GO: 10035 | Response to inorganic stimulus |
|      | GO: 33554 | Cellular response to stress |
|      | GO: 34599 | Cellular response to oxidative stress |
|      | GO: 34614 | Cellular response to reactive oxygen species |
|      | GO: 42221 | Cellular response to chemical |
|      | GO: 42743 | Hydrogen peroxide metabolic process |
|      | GO: 42744 | Hydrogen peroxide catabolic process |
|      | GO: 50896 | Response to stimulus |
|      | GO: 51716 | Cellular response to stimulus |
|      | GO: 90304 | Cellular response to chemical stimulus |</p>
<table>
<thead>
<tr>
<th>Antibiotics</th>
<th>Ciprofloxacin</th>
<th>GO: 302</th>
<th>Response to reactive oxygen species</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>GO: 303</td>
<td>Response to superoxide</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 305</td>
<td>Response to oxygen radical</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 6139</td>
<td>Nucleobase, nucleoside, nucleotide and nucleic acid metabolic process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 6259</td>
<td>DNA metabolic process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 6260</td>
<td>DNA replication</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 6281</td>
<td>Transcription</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 6350</td>
<td>Transcription, DNA-dependent</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 6351</td>
<td>DNA template transcription, initiation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 6352</td>
<td>Regulation of transcription, DNA-dependent</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 6355</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 6800</td>
<td>Oxygen and reactive oxygen species metabolic process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 6801</td>
<td>Cellular aldehyde metabolic process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 6950</td>
<td>Response to stress</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 6974</td>
<td>Response to DNA damage stimulus</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 6979</td>
<td>Response to oxidative stress</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 7154</td>
<td>Cell communication</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 9059</td>
<td>Macromolecule biosynthetic process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 9266</td>
<td>Response to temperature stimulus</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 9432</td>
<td>SOS response</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 9605</td>
<td>Response to abiotic stimulus</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 9628</td>
<td>Regulation of biosynthetic process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 9889</td>
<td>Response to extracellular stimulus</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 9991</td>
<td>Regulation of gene expression</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 10468</td>
<td>Regulation of macromolecular biosynthetic process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 10556</td>
<td>Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 19219</td>
<td>Regulation of metabolic process</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Regulation of cellular metabolic process</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Regulation of cellular biosynthetic process</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cellular response to external stimulus</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 19222</td>
<td>RNA biosynthetic process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 31323</td>
<td>Cellular response to stress</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 31326</td>
<td>Cellular response to reactive oxygen species</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 31326</td>
<td>Cellular macromolecule biosynthetic process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 31668</td>
<td>Response to chemical stimulus</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Macromolecular metabolic process</td>
</tr>
<tr>
<td>GO: 32774</td>
<td>Cellular macromolecule metabolic process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 33554</td>
<td>Regulation of transcription</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 34614</td>
<td>Cell redox homeostasis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 34645</td>
<td>Regulation of biological process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 42221</td>
<td>Regulation of cellular process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 43170</td>
<td>Response to stimulus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 44260</td>
<td>Regulation of nitrogen compound metabolic process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 45449</td>
<td>Cellular response to stimulus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 45454</td>
<td>Regulation of RNA metabolic process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 50789</td>
<td>Regulation of macromolecule metabolic process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 50794</td>
<td>Biological regulation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 50896</td>
<td>Cellular response to chemical stimulus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 51171</td>
<td>Cellular response to oxygen radical</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 51252</td>
<td>Cellular response to superoxide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 51716</td>
<td>Cellular response to external stimulus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 60255</td>
<td>Regulation of primary metabolic process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 65007</td>
<td>Nucleic acid metabolic process</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Erythromycin | 80090  
|             | GO: 302  
|             | GO: 303  
|             | GO: 305  
|             | GO: 6259  
|             | GO: 6281  
|             | GO: 6284  
|             | GO: 6289  
|             | GO: 6800  
|             | GO: 6801  
|             | GO: 6950  
|             | GO: 6974  
|             | GO: 6979  
|             | GO: 7154  
|             | GO: 9432  
|             | GO: 9605  
|             | GO: 9991  
|             | GO: 19430  
|             | GO: 31668  
|             | GO: 33554  
|             | GO: 34599  
|             | GO: 34614  
|             | GO: 42221  
|             | GO: 43170  
|             | GO: 44260  
|             | GO: 50896  
|             | GO: 51716  
|             | GO: 71450  
|             | GO: 71451  
|             | GO: 71496  

Response to reactive oxygen species  
Response to superoxide  
Response to oxygen radical  
DNA metabolic process  
DNA repair  
Base-excision repair  
Nucleotide-excision repair  
Oxygen and reactive oxygen species metabolic process  
Superoxide metabolic process  
Response to stress  
Response to DNA damage stimulus  
Response to oxidative stress  
Cell communication  
SOS response  
Response to external stimulus  
Response to extracellular stimulus  
Removal of superoxide radicals  
Cellular response to external stimulus  
Cellular response to stress  
Cellular response to oxidative stress  
Cellular response to reactive oxygen species  
Response to chemical stimulus  
Macromolecular metabolic process  
Cellular macromolecule metabolic process  
Response to stimulus  
Cellular response to stimulus  
Cellular response to oxygen radical  
Cellular response to superoxide  
Cellular response to external stimulus
<p>| Mitomycin C | GO: 18 | Regulation of DNA recombination |
| | GO: 302 | Response to reactive oxygen species |
| | GO: 303 | Response to superoxide |
| | GO: 6139 | Nucleobase, nucleoside, nucleotide and nucleic acid metabolic process |
| | GO: 6259 | DNA metabolic process |
| | GO: 6260 | DNA replication |
| | GO: 6261 | DNA repair |
| | GO: 6281 | Base-excision repair |
| | GO: 6284 | Base-excision repair, AP site formation |
| | GO: 6285 | Nucleotide-excision repair |
| | GO: 6289 | Mismatch repair |
| | GO: 6298 | Regulation of transcription, DNA-dependent |
| | GO: 6355 | Oxygen and reactive oxygen species metabolic process |
| | GO: 6800 | Nitrogen compound metabolic process |
| | GO: 6807 | Response to stress |
| | GO: 6950 | Response to DNA damage stimulus |
| | GO: 6974 | Response to oxidative stress |
| | GO: 6979 | Cell communication |
| | GO: 7154 | Response to temperature stimulus |
| | GO: 9266 | SOS response |
| | GO: 9432 | Response to external stimulus |
| | GO: 9605 | Regulation of biosynthetic process |
| | GO: 9889 | Cellular process |
| | GO: 9987 | Response to extracellular stimulus |
| | GO: 9991 | Response to inorganic stimulus |
| | GO: 10035 | Regulation of macromolecular biosynthetic process |
| | GO: 10556 | Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process |
| | GO: 19219 | Regulation of metabolic process |
| | GO: 19222 | Regulation of cellular metabolic process |
| | GO: 31323 | Regulation of cellular biosynthetic process |
| | GO: 31326 | Cellular response to external stimulus |
| | GO: 31668 | DNA geometric change |
| | GO: 31668 | DNA duplex unwinding |
| | GO: 31668 | Cellular response to stress |
| | GO: 31668 | Cellular response to oxidative stress |
| | GO: 31668 | Cellular response to reactive oxygen species |
| | GO: 31668 | Response to chemical stimulus |
| | GO: 31668 | Hydrogen peroxide metabolic process |</p>
<table>
<thead>
<tr>
<th>GO:</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>32392</td>
<td>Hydrogen peroxide catabolic process</td>
</tr>
<tr>
<td>32508</td>
<td>Macromolecular metabolic process</td>
</tr>
<tr>
<td>33554</td>
<td>Protein unfolding</td>
</tr>
<tr>
<td>34599</td>
<td>Cellular macromolecule metabolic process</td>
</tr>
<tr>
<td>34614</td>
<td>Regulation of transcription</td>
</tr>
<tr>
<td>42221</td>
<td>Regulation of biological process</td>
</tr>
<tr>
<td>42743</td>
<td>Regulation of cellular process</td>
</tr>
<tr>
<td>42744</td>
<td>Response to stimulus</td>
</tr>
<tr>
<td>43170</td>
<td>Regulation of DNA metabolic process</td>
</tr>
<tr>
<td>43335</td>
<td>Regulation of nitrogen compound metabolic process</td>
</tr>
<tr>
<td>44260</td>
<td>Regulation of macromolecule metabolic process</td>
</tr>
<tr>
<td>45449</td>
<td>Biological regulation</td>
</tr>
<tr>
<td>50789</td>
<td>Cellular response to chemical stimulus</td>
</tr>
<tr>
<td>50794</td>
<td>Cellular response to antibiotics</td>
</tr>
<tr>
<td>50896</td>
<td>Cellular response to external stimulus</td>
</tr>
<tr>
<td>51052</td>
<td>Regulation of primary metabolic process</td>
</tr>
<tr>
<td>51171</td>
<td>Nucleic acid metabolic process</td>
</tr>
<tr>
<td>51252</td>
<td>Nucleic acid phosphodiester bond hydrolysis</td>
</tr>
<tr>
<td>51716</td>
<td></td>
</tr>
<tr>
<td>60255</td>
<td></td>
</tr>
<tr>
<td>65007</td>
<td></td>
</tr>
<tr>
<td>65007</td>
<td></td>
</tr>
<tr>
<td>65007</td>
<td></td>
</tr>
<tr>
<td>65007</td>
<td></td>
</tr>
<tr>
<td>Pharmaceuticals</td>
<td>Atorvastatin</td>
</tr>
<tr>
<td>------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Ibuprofen</td>
<td>GO: 302</td>
</tr>
<tr>
<td>GO: 6800</td>
<td>Oxygen and reactive oxygen species metabolic process</td>
</tr>
<tr>
<td>----------</td>
<td>------------------------------------------------------</td>
</tr>
<tr>
<td>GO: 6950</td>
<td>Response to stress</td>
</tr>
<tr>
<td>GO: 6974</td>
<td>Response to DNA damage stimulus</td>
</tr>
<tr>
<td>GO: 6979</td>
<td>Response to oxidative stress</td>
</tr>
<tr>
<td>GO: 7154</td>
<td>Cell communication</td>
</tr>
<tr>
<td>GO: 9059</td>
<td>Macromolecule biosynthetic process</td>
</tr>
<tr>
<td>GO: 9266</td>
<td>Response to temperature stimulus</td>
</tr>
<tr>
<td>GO: 9408</td>
<td>Response to heat</td>
</tr>
<tr>
<td>GO: 9432</td>
<td>SOS response</td>
</tr>
<tr>
<td>GO: 9605</td>
<td>Response to abiotic stimulus</td>
</tr>
<tr>
<td>GO: 9628</td>
<td>Response to extracellular stimulus</td>
</tr>
<tr>
<td>GO: 9991</td>
<td>Response to inorganic stimulus</td>
</tr>
<tr>
<td>GO: 10035</td>
<td>Translesion synthesis</td>
</tr>
<tr>
<td>GO: 19985</td>
<td>Cellular response to external stimulus</td>
</tr>
<tr>
<td>GO: 31668</td>
<td>Response to hydrogen peroxide</td>
</tr>
<tr>
<td>GO: 33194</td>
<td>Cellular response to stress</td>
</tr>
<tr>
<td>GO: 33554</td>
<td>Cellular response to reactive oxygen species</td>
</tr>
<tr>
<td>GO: 34614</td>
<td>Cellular macromolecule biosynthetic process</td>
</tr>
<tr>
<td>GO: 34645</td>
<td>Response to chemical stimulus</td>
</tr>
<tr>
<td>GO: 42221</td>
<td>Response to drug</td>
</tr>
<tr>
<td>GO: 42493</td>
<td>Protein unfolding</td>
</tr>
<tr>
<td>GO: 43335</td>
<td>Cell redox homeostasis</td>
</tr>
<tr>
<td>GO: 45454</td>
<td>Response to antibiotic</td>
</tr>
<tr>
<td>GO: 46677</td>
<td>Regulation of biological process</td>
</tr>
<tr>
<td>GO: 50789</td>
<td>Regulation of cellular process</td>
</tr>
<tr>
<td>GO: 50794</td>
<td>Response to stimulus</td>
</tr>
<tr>
<td>GO: 50896</td>
<td>Cellular response to stimulus</td>
</tr>
<tr>
<td>GO: 51716</td>
<td>Biological regulation</td>
</tr>
<tr>
<td>GO:</td>
<td>Regulation of biological quality</td>
</tr>
<tr>
<td>GO:</td>
<td>Cellular response to chemical stimulus</td>
</tr>
<tr>
<td>GO:</td>
<td>Cellular response to external stimulus</td>
</tr>
<tr>
<td>GO: 65007</td>
<td>Response to reactive oxygen species</td>
</tr>
<tr>
<td>GO: 65008</td>
<td>Response to superoxide</td>
</tr>
<tr>
<td>GO: 70887</td>
<td>Response to oxygen radical</td>
</tr>
<tr>
<td>GO: 71496</td>
<td>DNA repair</td>
</tr>
<tr>
<td>GO: 302</td>
<td>Oxygen and reactive oxygen species</td>
</tr>
<tr>
<td>GO: 303</td>
<td>metabolic process</td>
</tr>
<tr>
<td>GO: 305</td>
<td>Superoxide metabolic process</td>
</tr>
<tr>
<td>GO: 6281</td>
<td>Response to stress</td>
</tr>
<tr>
<td>GO: 6800</td>
<td>Response to oxidative stress</td>
</tr>
<tr>
<td>GO: 6801</td>
<td>Response to temperature stimulus</td>
</tr>
<tr>
<td>GO: 6950</td>
<td>SOS response</td>
</tr>
<tr>
<td>GO: 6979</td>
<td>Response to external stimulus</td>
</tr>
<tr>
<td>GO: 9266</td>
<td>Response to abiotic stimulus</td>
</tr>
<tr>
<td>GO: 9432</td>
<td>Response to extracellular stimulus</td>
</tr>
<tr>
<td>GO: 9605</td>
<td>Response to inorganic stimulus</td>
</tr>
<tr>
<td>GO: 9628</td>
<td>Removal of superoxide radicals</td>
</tr>
<tr>
<td>GO: 9991</td>
<td>Response to hydrogen peroxide</td>
</tr>
<tr>
<td>GO: 10035</td>
<td>Cellular response to stress</td>
</tr>
<tr>
<td>GO: 19430</td>
<td>Cellular response to oxidative stress</td>
</tr>
<tr>
<td>GO: 33194</td>
<td>Cellular response to reactive oxygen</td>
</tr>
<tr>
<td>GO: 33554</td>
<td>species</td>
</tr>
<tr>
<td>GO: 34599</td>
<td>Response to chemical stimulus</td>
</tr>
<tr>
<td>GO: 34614</td>
<td>Response to drug</td>
</tr>
<tr>
<td>GO: 34614</td>
<td>Hydrogen peroxide metabolic process</td>
</tr>
<tr>
<td>GO: 34614</td>
<td>Hydrogen peroxide catabolic process</td>
</tr>
<tr>
<td>GO: 34614</td>
<td>Regulation of transcription</td>
</tr>
<tr>
<td>GO: 42221</td>
<td>Response to antibiotic</td>
</tr>
<tr>
<td>GO: 42493</td>
<td>Regulation of biological process</td>
</tr>
<tr>
<td>GO: 42743</td>
<td>Regulation of cellular process</td>
</tr>
<tr>
<td>GO: 42743</td>
<td>Response to stimulus</td>
</tr>
<tr>
<td>GO: 42744</td>
<td>Biological regulation</td>
</tr>
<tr>
<td>GO: 45449</td>
<td>Regulation of biological quality</td>
</tr>
<tr>
<td>GO: 46677</td>
<td>Cellular response to chemical stimulus</td>
</tr>
<tr>
<td>GO:</td>
<td>Cellular response to oxygen radical</td>
</tr>
<tr>
<td>GO:</td>
<td>Cellular response to superoxide</td>
</tr>
<tr>
<td>Drinking water byproduct</td>
<td>N-Nitrosodimethylamine</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>GO: 42221</td>
<td>Regulation of transcription</td>
</tr>
<tr>
<td>GO: 43170</td>
<td>Regulation of biological process</td>
</tr>
<tr>
<td>GO: 44260</td>
<td>Regulation of cellular process</td>
</tr>
<tr>
<td>GO: 45004</td>
<td>Response to stimulus</td>
</tr>
<tr>
<td>GO: 45005</td>
<td>Regulation of nitrogen compound metabolic process</td>
</tr>
<tr>
<td>GO: 45449</td>
<td>Regulation of RNA metabolic process</td>
</tr>
<tr>
<td>GO: 50789</td>
<td>Cellular response to stimulus</td>
</tr>
<tr>
<td>GO: 50794</td>
<td>Regulation of macromolecule metabolic process</td>
</tr>
<tr>
<td>GO: 50896</td>
<td>Biological regulation</td>
</tr>
<tr>
<td>GO: 51171</td>
<td>Cellular response to oxygen radical</td>
</tr>
<tr>
<td>GO: 51252</td>
<td>Cellular response to external stimulus</td>
</tr>
<tr>
<td>GO: 51716</td>
<td>Regulation of primary metabolic process</td>
</tr>
<tr>
<td>GO: 60255</td>
<td>Nucleic acid metabolic process</td>
</tr>
<tr>
<td>GO: 65007</td>
<td>Nucleic acid phosphodiester bond hydrolysis</td>
</tr>
<tr>
<td>GO: 71450</td>
<td></td>
</tr>
<tr>
<td>GO: 71496</td>
<td></td>
</tr>
<tr>
<td>GO: 80090</td>
<td></td>
</tr>
<tr>
<td>GO: 90304</td>
<td></td>
</tr>
<tr>
<td>GO: 90305</td>
<td></td>
</tr>
</tbody>
</table>

<p>| Bromodichloromethane | GO: 18 | Regulation of DNA recombination |
| GO: 6259 | DNA metabolic process |
| GO: 6281 | DNA repair |
| GO: 6350 | Transcription |
| GO: 6351 | Transcription, DNA-dependent |
| GO: 6355 | Regulation of transcription, DNA-dependent |
| GO: 6950 | Response to stress |
| GO: 6974 | Response to DNA damage stimulus |
| GO: 6979 | Response to oxidative stress |
| GO: 9266 | Response to temperature stimulus |
| GO: 9991 | Response to extracellular stimulus |
| 19219 | Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process |
| 19222 | Regulation of metabolic process |
| GO: | Regulation of cellular metabolic process |
| 31323 | Cellular response to external stimulus |
| GO: | RNA biosynthetic process |
| 31668 | Cellular response to stress |
| GO: | Response to chemical stimulus |
| 32774 | Cellular macromolecule metabolic process |
| GO: | Cell redox homeostasis |
| 33554 | Regulation of biological process |
| GO: | Regulation of cellular process |
| 42221 | Response to stimulus |
| GO: | Regulation of nitrogen compound metabolic process |
| 44260 | Regulation of RNA metabolic process |
| GO: | Cellular response to stimulus |
| 45454 | Regulation of macromolecule metabolic process |
| 50789 | Biological regulation |
| GO: | Cellular response to external stimulus |
| 50794 | Regulation of primary metabolic process |
| 50896 | Nucleic acid metabolic process |
| GO: | |
| 51171 | |
| GO: | |
| 51252 | |
| GO: | |
| 51716 | |
| GO: | |
| 60255 | |
| GO: | |
| 65007 | |</p>
<table>
<thead>
<tr>
<th>Compound</th>
<th>GO IDs</th>
<th>Processes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trihalomethane mix</td>
<td>GO: 71496</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GO: 80090</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GO: 90304</td>
<td></td>
</tr>
<tr>
<td>Nitrosamine mix</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industrial and food additive</td>
<td>GO: 302</td>
<td>Response to reactive oxygen species</td>
</tr>
<tr>
<td></td>
<td>GO: 303</td>
<td>Response to superoxide</td>
</tr>
<tr>
<td></td>
<td>GO: 305</td>
<td>Response to oxygen radical</td>
</tr>
<tr>
<td></td>
<td>GO: 6800</td>
<td>Oxygen and reactive oxygen species metabolic process</td>
</tr>
<tr>
<td></td>
<td>GO: 6801</td>
<td>Superoxide metabolic process</td>
</tr>
<tr>
<td></td>
<td>GO: 6950</td>
<td>Response to stress</td>
</tr>
<tr>
<td></td>
<td>GO: 6979</td>
<td>Response to oxidative stress</td>
</tr>
<tr>
<td></td>
<td>GO: 10035</td>
<td>Response to inorganic stimulus</td>
</tr>
<tr>
<td></td>
<td>GO: 19219</td>
<td>Regulation to nucleobase, nucleoside, nucleotide and nucleic acid metabolic process</td>
</tr>
<tr>
<td></td>
<td>GO: 19430</td>
<td>Removal of superoxide radicals</td>
</tr>
<tr>
<td></td>
<td>GO: 19725</td>
<td>Cellular homeostasis</td>
</tr>
<tr>
<td></td>
<td>GO: 33194</td>
<td>Response to hydrogen peroxide</td>
</tr>
<tr>
<td></td>
<td>GO: 33554</td>
<td>Cellular response to stress</td>
</tr>
<tr>
<td></td>
<td>GO: 34614</td>
<td>Cellular response to reactive oxygen species</td>
</tr>
<tr>
<td></td>
<td>GO: 42221</td>
<td>Response to chemical stimulus</td>
</tr>
<tr>
<td></td>
<td>GO: 42493</td>
<td>Response to drug</td>
</tr>
<tr>
<td></td>
<td>GO: 42592</td>
<td>Homeostatic process</td>
</tr>
<tr>
<td></td>
<td>GO: 45454</td>
<td>Cell redox homeostasis</td>
</tr>
<tr>
<td></td>
<td>GO: 50789</td>
<td>Regulation of biological process</td>
</tr>
<tr>
<td></td>
<td>GO: 50794</td>
<td>Regulation of cellular process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Response to stimulus</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cellular response to stimulus</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Biological regulation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Regulation of biological quality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cellular response to chemical stimulus</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cellular response to oxygen radical</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cellular response to superoxide</td>
</tr>
<tr>
<td>GO: 50896</td>
<td>GO: 51716</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>GO: 65007</td>
<td>GO: 65008</td>
<td></td>
</tr>
<tr>
<td>GO: 70887</td>
<td>GO: 71450</td>
<td></td>
</tr>
<tr>
<td>GO: 71451</td>
<td>GO: 6350</td>
<td></td>
</tr>
<tr>
<td>GO: 6351</td>
<td>GO: 6355</td>
<td></td>
</tr>
<tr>
<td>GO: 6800</td>
<td>GO: 6801</td>
<td></td>
</tr>
<tr>
<td>GO: 6950</td>
<td>GO: 6979</td>
<td></td>
</tr>
<tr>
<td>GO: 9059</td>
<td>GO: 9628</td>
<td></td>
</tr>
<tr>
<td>GO: 9889</td>
<td>GO: 10468</td>
<td></td>
</tr>
<tr>
<td>GO: 10556</td>
<td>GO: 19219</td>
<td></td>
</tr>
<tr>
<td>GO: 19222</td>
<td>GO: 19430</td>
<td></td>
</tr>
<tr>
<td>GO: 19725</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Surcolase**

<p>| GO: 302  | Response to reactive oxygen species |
| GO: 303  | Response to superoxide |
| GO: 305  | Response to oxygen radical |
| GO: 6350 | Transcription |
| GO: 6351 | Transcription, DNA-dependent |
| GO: 6355 | Regulation of transcription, DNA-dependent |
| GO: 6800 | Oxygen and reactive oxygen species metabolic process |
| GO: 6801 | Superoxide metabolic process |
| GO: 6950 | Response to stress |
| GO: 6979 | Response to oxidative stress |
| GO: 9059 | Macromolecule biosynthetic process |
| GO: 9628 | Response to abiotic stimulus |
| GO: 9889 | Regulation of biosynthetic process |
| GO: 10468 | Regulation of gene expression |
| GO: 10556 | Regulation of macromolecular biosynthetic process |
| GO: 19219 | Regulation to nucleobase, nucleoside, nucleotide and nucleic acid metabolic process |
| GO: 19222 | Regulation of metabolic process |
| GO: 19430 | Removal of superoxide radicals |
| GO: 19725 | Cellular homeostasis |
| GO:      | Regulation of anatomical structure morphogenesis |
| GO:      | Regulation of cellular metabolic process |
| GO:      | Regulation of cellular biosynthetic process |
| GO:      | RNA biosynthetic process |
| GO:      | Cellular response to stress |
| GO:      | Cellular response to oxidative stress |
| GO:      | Cellular response to reactive oxygen species |</p>
<table>
<thead>
<tr>
<th>GeneID</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>22603</td>
<td>Response to chemical stimulus</td>
</tr>
<tr>
<td></td>
<td>Response to drug</td>
</tr>
<tr>
<td></td>
<td>Homeostatic process</td>
</tr>
<tr>
<td>31323</td>
<td>Regulation of transcription</td>
</tr>
<tr>
<td>31326</td>
<td>Cell redox homeostasis</td>
</tr>
<tr>
<td></td>
<td>Response to antibiotic</td>
</tr>
<tr>
<td>32774</td>
<td>Regulation of biological process</td>
</tr>
<tr>
<td></td>
<td>Regulation of cellular process</td>
</tr>
<tr>
<td>33554</td>
<td>Response to stimulus</td>
</tr>
<tr>
<td></td>
<td>Regulation of nitrogen compound metabolic process</td>
</tr>
<tr>
<td>34599</td>
<td>Regulation of RNA metabolic process</td>
</tr>
<tr>
<td></td>
<td>Cellular response to stimulus</td>
</tr>
<tr>
<td>34614</td>
<td>Biological regulation</td>
</tr>
<tr>
<td>42221</td>
<td>Regulation of biological quality</td>
</tr>
<tr>
<td></td>
<td>Cellular response to chemical stimulus</td>
</tr>
<tr>
<td>42493</td>
<td>Cellular response to oxygen radical</td>
</tr>
<tr>
<td></td>
<td>Cellular response to superoxide</td>
</tr>
<tr>
<td>42592</td>
<td>Regulation of primary metabolic process</td>
</tr>
<tr>
<td>45449</td>
<td></td>
</tr>
<tr>
<td>45454</td>
<td></td>
</tr>
<tr>
<td>46677</td>
<td></td>
</tr>
<tr>
<td>50789</td>
<td></td>
</tr>
<tr>
<td>50794</td>
<td></td>
</tr>
<tr>
<td>50896</td>
<td></td>
</tr>
<tr>
<td>51171</td>
<td></td>
</tr>
<tr>
<td>51252</td>
<td></td>
</tr>
<tr>
<td>51716</td>
<td></td>
</tr>
<tr>
<td>65007</td>
<td></td>
</tr>
<tr>
<td>65008</td>
<td></td>
</tr>
<tr>
<td>70887</td>
<td></td>
</tr>
<tr>
<td>4-n-Nonylphenol</td>
<td>GO: 303</td>
</tr>
<tr>
<td></td>
<td>GO: 305</td>
</tr>
<tr>
<td></td>
<td>GO: 6259</td>
</tr>
<tr>
<td></td>
<td>GO: 6281</td>
</tr>
<tr>
<td></td>
<td>GO: 6350</td>
</tr>
<tr>
<td></td>
<td>GO: 6351</td>
</tr>
<tr>
<td></td>
<td>GO: 6800</td>
</tr>
<tr>
<td></td>
<td>GO: 6950</td>
</tr>
<tr>
<td></td>
<td>GO: 9266</td>
</tr>
<tr>
<td></td>
<td>GO: 9432</td>
</tr>
<tr>
<td></td>
<td>GO: 9605</td>
</tr>
<tr>
<td></td>
<td>GO: 9628</td>
</tr>
<tr>
<td></td>
<td>GO: 9889</td>
</tr>
<tr>
<td></td>
<td>GO: 9991</td>
</tr>
<tr>
<td></td>
<td>GO: 10556</td>
</tr>
<tr>
<td></td>
<td>GO: 19219</td>
</tr>
<tr>
<td></td>
<td>GO: 19222</td>
</tr>
<tr>
<td></td>
<td>GO: 31323</td>
</tr>
<tr>
<td></td>
<td>GO: 31326</td>
</tr>
<tr>
<td></td>
<td>GO: 31668</td>
</tr>
<tr>
<td></td>
<td>GO: 32774</td>
</tr>
<tr>
<td></td>
<td>GO: 33554</td>
</tr>
<tr>
<td></td>
<td>GO: 34599</td>
</tr>
<tr>
<td></td>
<td>GO: 34599</td>
</tr>
<tr>
<td></td>
<td>GO: 34599</td>
</tr>
<tr>
<td></td>
<td>GO: 34599</td>
</tr>
<tr>
<td></td>
<td>GO: 34599</td>
</tr>
<tr>
<td></td>
<td>GO: 34599</td>
</tr>
<tr>
<td></td>
<td>GO: 34599</td>
</tr>
<tr>
<td></td>
<td>GO: 34599</td>
</tr>
<tr>
<td></td>
<td>GO: 34599</td>
</tr>
<tr>
<td>GO:</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>34614</td>
<td>Regulation of RNA metabolic process</td>
</tr>
<tr>
<td>42221</td>
<td>Cellular response to stimulus</td>
</tr>
<tr>
<td>43335</td>
<td>Regulation of macromolecule metabolic process</td>
</tr>
<tr>
<td>45449</td>
<td>Biological regulation</td>
</tr>
<tr>
<td>50789</td>
<td>Cellular response to oxygen radical</td>
</tr>
<tr>
<td>50794</td>
<td>Cellular response to superoxide</td>
</tr>
<tr>
<td>50896</td>
<td>Cellular response to external stimulus</td>
</tr>
<tr>
<td>51171</td>
<td>Regulation of primary metabolic process</td>
</tr>
<tr>
<td>51252</td>
<td>Nucleic acid metabolic process</td>
</tr>
<tr>
<td>51716</td>
<td></td>
</tr>
<tr>
<td>60255</td>
<td></td>
</tr>
<tr>
<td>65007</td>
<td></td>
</tr>
<tr>
<td>71450</td>
<td>DNA metabolic process</td>
</tr>
<tr>
<td>71451</td>
<td>DNA repair</td>
</tr>
<tr>
<td>71496</td>
<td>Base-excision repair</td>
</tr>
<tr>
<td>80090</td>
<td>Oxygen and reactive oxygen species metabolic process</td>
</tr>
<tr>
<td>90304</td>
<td>Response to stress</td>
</tr>
<tr>
<td>6259</td>
<td>DNA metabolic process</td>
</tr>
<tr>
<td>6281</td>
<td>DNA repair</td>
</tr>
<tr>
<td>6284</td>
<td>Base-excision repair</td>
</tr>
<tr>
<td>6800</td>
<td>Oxygen and reactive oxygen species metabolic process</td>
</tr>
<tr>
<td>6950</td>
<td>Response to stress</td>
</tr>
<tr>
<td>6974</td>
<td>Cell communication</td>
</tr>
<tr>
<td>7154</td>
<td>Response to heat</td>
</tr>
<tr>
<td>Bis(2-ethylhexyl) phthalate</td>
<td></td>
</tr>
<tr>
<td>GO: 6259</td>
<td>DNA metabolic process</td>
</tr>
<tr>
<td>GO: 6281</td>
<td>DNA repair</td>
</tr>
<tr>
<td>GO: 6284</td>
<td>Base-excision repair</td>
</tr>
<tr>
<td>GO: 6800</td>
<td>Oxygen and reactive oxygen species metabolic process</td>
</tr>
<tr>
<td>GO: 6950</td>
<td>Response to stress</td>
</tr>
<tr>
<td>GO: 6974</td>
<td>Response to DNA damage stimulus</td>
</tr>
<tr>
<td>GO: 7154</td>
<td>Cell communication</td>
</tr>
<tr>
<td></td>
<td>Response to heat</td>
</tr>
<tr>
<td>GO: 9408</td>
<td>SOS response</td>
</tr>
<tr>
<td>GO: 9432</td>
<td>Response to external stimulus</td>
</tr>
<tr>
<td>GO: 9605</td>
<td>Response to extracellular stimulus</td>
</tr>
<tr>
<td>GO: 9991</td>
<td>Cellular response to external stimulus</td>
</tr>
<tr>
<td>GO: 31668</td>
<td>Cellular response to stress</td>
</tr>
<tr>
<td>GO: 33554</td>
<td>Response to chemical stimulus</td>
</tr>
<tr>
<td>GO: 42221</td>
<td>Macromolecular metabolic process</td>
</tr>
<tr>
<td>GO: 43170</td>
<td>Cellular macromolecule metabolic process</td>
</tr>
<tr>
<td>GO: 44260</td>
<td>Response to stimulus</td>
</tr>
<tr>
<td>GO: 50896</td>
<td>Cellular response to stimulus</td>
</tr>
<tr>
<td>GO: 51716</td>
<td>Cellular response to external stimulus</td>
</tr>
</tbody>
</table>

**GO: 6139**

| GO: 6259 | Nucleobase, nucleoside, nucleotide and nucleic acid metabolic process |
| GO: 6261 | DNA metabolic process |
| GO: 6281 | DNA-dependent DNA replication |
| GO: 6284 | DNA repair |
| GO: 6285 | Base-excision repair |
| GO: 6289 | Base-excision repair, AP site formation |
| GO: 6301 | Nucleotide-excision repair |
| GO: 6355 | Postreplication repair |
| GO: 6950 | Regulation of transcription, DNA-dependent |
| GO: 6974 | Response to stress |
| GO: 7154 | Response to DNA damage stimulus |
| GO: 9432 | Cell communication |
| GO: 9889 | SOS response |
| GO: 9991 | Regulation of biosynthetic process |

**GO: 10468**

| GO: 10556 | Response to extracellular stimulus |
| GO: 19219 | Regulation of gene expression |
| GO: 10556 | Regulation of macromolecular biosynthetic process |

<p>| GO: 19219 | Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process |
| GO: 19219 | Regulation of metabolic process |
| GO: 19219 | Translesion synthesis |
| GO: 19219 | Regulation of cellular metabolic process |</p>
<table>
<thead>
<tr>
<th>Herbicide and insecticide</th>
<th>5-CHLORO-2-(2,4-DICHLOROPHENOXY)PHENOL</th>
<th>GO: 302</th>
<th>Response to reactive oxygen species</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>GO: 303</td>
<td>Response to superoxide</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 305</td>
<td>Response to oxygen radical</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 6350</td>
<td>Transcription</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO: 6351</td>
<td>Oxygen and reactive oxygen species</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO</td>
<td>Description</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 6800</td>
<td>metabolic process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 6801</td>
<td>Superoxide metabolic process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 6950</td>
<td>Response to stress</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 6979</td>
<td>Response to oxidative stress</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 9266</td>
<td>Response to temperature stimulus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 9408</td>
<td>Response to heat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 9628</td>
<td>Response to abiotic stimulus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 10035</td>
<td>Response to inorganic stimulus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 10556</td>
<td>Regulation of macromolecular biosynthetic process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 19219</td>
<td>Regulation to nucleobase, nucleoside, nucleotide and nucleic acid metabolic process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 19430</td>
<td>Removal of superoxide radicals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 32774</td>
<td>RNA biosynthetic process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 34599</td>
<td>Cellular response to oxidative stress</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 34614</td>
<td>Cellular response to reactive oxygen species</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 35886</td>
<td>Response to chemical stimulus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 36777</td>
<td>Response to drug</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 42221</td>
<td>Regulation of transcription</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 42493</td>
<td>Response to antibiotic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 45449</td>
<td>Regulation of biological process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 50896</td>
<td>Regulation of cellular process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 50789</td>
<td>Response to stimulus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 50794</td>
<td>Response to stimulus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 51171</td>
<td>Regulation of nitrogen compound metabolic process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 51171</td>
<td>Regulation of RNA metabolic process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 51171</td>
<td>Cellular response to stimulus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 51171</td>
<td>Biological regulation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 51171</td>
<td>Regulation of biological quality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 51171</td>
<td>Cellular response to chemical stimulus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 51171</td>
<td>Cellular response to oxygen radical</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO: 51171</td>
<td>Cellular response to superoxide</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| GO: 51252 | **Carbaryl** | GO: 302 Response to reactive oxygen species  
GO: 303 Response to superoxide  
GO: 305 Response to oxygen radical  
GO: 6281 DNA repair  
GO: 6350 Transcription  
GO: 6351 Transcription, DNA-dependent  
GO: 6800 Oxygen and reactive oxygen species metabolic process  
GO: 6801 Superoxide metabolic process  
GO: 6950 Response to stress  
GO: 6974 Response to DNA damage stimulus  
GO: 6979 Response to oxidative stress  
GO: 6986 Response to unfolded protein  
GO: 7154 Cell communication  
GO: 9266 Response to temperature stimulus  
GO: 9432 SOS response  
GO: 9605 Response to external stimulus  
GO: 9628 Response to abiotic stimulus  
GO: 9991 Response to extracellular stimulus  
GO: 10035 Response to inorganic stimulus  
GO: 10556 Regulation of macromolecular biosynthetic process  
GO: 19219 Regulation to nucleobase, nucleoside, nucleotide and nucleic acid metabolic process  
GO: 19430 Removal of superoxide radicals  
GO: 19725 Cellular homeostasis |
| GO:        | Macromolecular metabolic process  |
| GO:        | Protein unfolding                |
| GO:        | Regulation of transcription      |
| GO:        | Cell redox homeostasis           |
| GO:        | Response to antibiotic           |
| GO:        | Regulation of biological process |
| GO:        | Regulation of cellular process   |
| GO:        | Response to stimulus             |
| GO:        | Regulation of nitrogen compound  |
| GO:        | Cellular response to stimulus    |
| GO:        | Response to protein stimulus     |
| GO:        | Biological regulation            |
| GO:        | Regulation of biological quality |
| GO:        | Cellular response to chemical stimulus |
| GO:        | Cellular response to oxygen radical |
| GO:        | Cellular response to superoxide  |
| GO:        | Cellular response to external stimulus |
| 2-chloro-4- (ethylamino)-6- (isopropylamino)-s-triazine | GO: 6800  
71451  
GO: 71496 | Oxygen and reactive oxygen species metabolic process  
Response to stress  
Response to oxidative stress  
Response to temperature stimulus  
Response to abiotic stimulus  
Response to inorganic stimulus  
Removal of superoxide radicals  
Cellular homeostasis  
Response to chemical stimulus  
Cell redox homeostasis  
Regulation of biological process  
Regulation of cellular process  
Response to stimulus  
Biological regulation  
Regulation of biological quality |
| Others | Hydrogen peroxide | GO: 302  
GO: 303  
GO: 305  
GO: 6139  
GO: 6259  
GO: 6281  
GO: 6350  
GO: 6351  
GO: 6355  
GO: 6800  
GO: 6801  
GO: 6950  
GO: 6974 | Response to reactive oxygen species metabolic process  
Response to superoxide  
Response to oxygen radical  
Nucleobase-containing compound metabolic process  
DNA metabolic process  
DNA repair  
Transcription  
Transcription, DNA-dependent  
Regulation of transcription, DNA-dependent  
Oxygen and reactive oxygen species metabolic process  
Superoxide metabolic process  
Response to stress  
Response to DNA damage stimulus  
Response to oxidative stress |
| GO: 6979  | Cell communication                        |
| GO: 7154  | Macromolecule biosynthetic process        |
| GO: 9059  | Response to temperature stimulus          |
| GO: 9266  | Response to heat                          |
| GO: 9408  | SOS response                              |
| GO: 9432  | Response to external stimulus             |
| GO: 9605  | Response to abiotic stimulus              |
| GO: 9628  | Regulation of biosynthetic process        |
| GO: 9889  | Response to extracellular stimulus        |
| GO: 9991  | Response to inorganic stimulus            |
| GO: 10035 | Regulation of gene expression             |
| GO: 10468 | Regulation of macromolecular biosynthetic process |
| GO: 10556 | Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process |
| GO: 19219 | Regulation of metabolic process           |
| GO: 19222 | Removal of superoxide radical             |
| GO: 19430 | Regulation of cellular metabolic process  |
| GO: 31323 | Regulation of cellular biosynthetic process |
| GO: 31326 | Cellular response to external stimulus    |
| GO: 31668 | Cellular response to stress               |
| GO: 33554 | Cellular response to oxidative stress     |
| GO: 34599 | Cellular response to reactive oxygen species |
| GO: 34614 | Cellular macromolecule metabolic process  |
| GO: 34645 | Regulation of transcription               |
| GO: 34684 | Regulation of cellular process            |
| GO: 42221 | Response to chemical stimulus             |
| GO: 43170 | Regulation of nitrogen compound metabolic process |
| GO: 43335 | Cellular response to stimulus             |
| GO:       | Regulation of primary metabolic process   |
| GO:       | Nucleic acid metabolic process            |
|-------------------------------|---------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| 17-beta-estradiol | GO: 303 | Response to superoxide |
| | GO: 305 | Response to oxygen radical |
| | GO: 6139 | Nucleobase, nucleoside, nucleotide and nucleic acid metabolic process |
| | GO: 6259 | DNA metabolic process |
| | GO: 6260 | DNA replication |
| | GO: 6261 | DNA-dependent DNA replication |
| | GO: 6281 | DNA repair |
| | GO: 6284 | Base-excision repair |
| | GO: 6289 | Nucleotide-excision repair |
| | GO: 6301 | Postreplication repair |
| | GO: 6801 | Superoxide metabolic process |
| | GO: 6807 | Nitrogen compound metabolic process |
| | GO: 6950 | Response to stress |
| | GO: 6974 | Response to DNA damage stimulus |
| | GO: 6979 | Response to oxidative stress |
| | GO: 6979 | Response to unfolded protein |
| | GO: 6986 | Cell communication |
| | GO: 7154 | Macromolecule biosynthetic process |
| | GO: 9059 | Response to temperature stimulus |
| | GO: 9266 | Response to heat |
| | GO: 9408 | SOS response |
| | GO: 9432 | Response to external stimulus |
| | GO: 9605 | Response to abiotic stimulus |
| | GO: 9628 | Response to extracellular stimulus |
| | GO: 9991 | Regulation to nucleobase, nucleoside, nucleotide and nucleic acid metabolic process |
| | GO: 19219 | Removal of superoxide radicals |
| | GO: 19430 | Translesion synthesis |
| | GO: 19985 | Cellular response to external stimulus |
| | GO: 31668 | Cellular response to stress |
| | GO: 33554 | Cellular response to reactive oxygen species |
| | GO: 34614 | Cellular macromolecule biosynthetic process |
| | GO: 34645 | Response to chemical stimulus |
| | | Macromolecular metabolic process |
| | | Protein unfolding |
| GO: 42221 | Cellular macromolecule metabolic process |
| GO: 43170 | Response to antibiotic |
| GO: 43335 | Regulation of biological process |
| GO: 46677 | Regulation of cellular process |
| GO: 50789 | Response to stimulus |
| GO: 50794 | Regulation of nitrogen compound metabolic process |
| GO: 51171 | Cellular response to stimulus |
| GO: 51716 | Response to protein stimulus |
| GO: 51789 | Biological regulation |
| GO: 50896 | Regulation of biological quality |
| GO: 50789 | Cellular response to chemical stimulus |
| GO: 50794 | Cellular response to oxygen radical |
| GO: 51789 | Cellular response to superoxide |
| GO: 50896 | Cellular response to external stimulus |
| GO: 51171 | Nucleic acid metabolic process |
| GO: 51789 | Nucleic acid phosphodiester bond hydrolysis |

Phenol liquor

<p>| GO: 6139 | Nucleobase, nucleoside, nucleotide and nucleic acid metabolic process |
| GO: 6259 | DNA metabolic process |
| GO: 6281 | DNA repair |
| GO: 6289 | Nucleotide-excision repair |
| GO: 6355 | Regulation of transcription, DNA-dependent |
| GO: 6950 | Response to stress |
| GO: 6974 | Response to DNA damage stimulus |
| GO: 7154 | Cell communication |
| GO: 9432 | SOS response |
| GO: 9991 | Response to extracellular stimulus |
| GO: 10468 | Regulation of gene expression |
| GO: 10556 | Regulation of macromolecular biosynthetic process |
| GO: 19219 | Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process |
| GO: 19222 | Translesion synthesis |
| GO: 19985 | Cellular response to external stimulus |
| GO: 50896 | Cellular response to stress |
| GO: 31668 | Response to chemical stimulus |
| GO: 19985 | Macromolecular metabolic process |
| GO: 33554 | Cellular macromolecule metabolic process |
| GO: 42221 | Regulation of transcription |
| GO: 43170 | Regulation of biological process |
| GO: 44260 | Regulation of cellular process |
| GO: 45449 | Response to stimulus |
| GO: 45449 | Regulation of nitrogen compound metabolic process |
| GO: 50789 | Regulation of RNA metabolic process |
| GO: 51171 | Biological regulation |
| GO: 51252 | Cellular response to external stimulus |
| GO: 50794 | Nucleic acid metabolic process |
| GO: 65007 | Nucleic acid phosphodiester bond hydrolysis |
| GO: 71496 | Regulation of transcription |
| BAP+S9  | GO: 18  | Regulation of DNA recombination |
|        | GO: 303 | Response to superoxide          |
|        | GO: 6139| Nucleobase, nucleoside, nucleotide and nucleic acid metabolic process |
|        | GO: 6200| ATP catabolic process           |
|        | GO: 6259| DNA metabolic process           |
|        | GO: 6281| DNA repair                      |
|        | GO: 6351| Transcription, DNA-dependent    |
|        | GO: 6800| Oxygen and reactive oxygen species metabolic process |
|        | GO: 6950| Response to stress              |
|        | GO: 6974| Response to DNA damage stimulus |
|        | GO: 6979| Response to oxidative stress    |
|        | GO: 7154| Cell communication              |
|        | GO: 9166| Nucleotide catabolic process    |
|        | GO: 9266| Response to temperature stimulus|
|        | GO: 9408| Response to heat                |
|        | GO: 9432| SOS response                    |
|        | GO: 9605| Response to external stimulus   |
|        | GO: 9628| Response to abiotic stimulus    |
|        | GO: 9991| Response to extracellular stimulus |
|        | GO: 19219| Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process |
|        | GO: 19222| Regulation of metabolic process |
|        | GO: 19222| Regulation of cellular metabolic process |
|        | GO: 31323| Cellular response to external stimulus |
|        | GO: 31668| Cellular response to stress     |
|        | GO: 33554| Cellular response to reactive oxygen species |
|        | GO: 34614| Response to chemical stimulus   |
|        | GO: 34614| Macromolecular metabolic process |
|        | GO: 34614| Cellular macromolecule metabolic process |
|        | GO: 34614| Regulation of transcription     |
|        | GO: 34614| Regulation of biological process |
|        | GO: 34614| Regulation of cellular process  |
|        | GO: 34614| Response to stimulus            |
|        | GO: 42221| Regulation of DNA metabolic process |
|        | GO: 43170| Regulation of nitrogen compound metabolic process |
|        | GO: 44260| Regulation of RNA metabolic process |
|        | GO: 44260| Cellular response to stimulus   |</p>
<table>
<thead>
<tr>
<th>Nanomaterials</th>
<th>nAg</th>
</tr>
</thead>
<tbody>
<tr>
<td>nTiO2_a</td>
<td></td>
</tr>
<tr>
<td>nTiO2_r</td>
<td></td>
</tr>
<tr>
<td>Fullerene soot</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 4.6 Description of mass balance model

The mass balance model and corresponding idea are all derived from Armitage’s work “Application of Mass Balance Models and the Chemical Activity Concept to Facilitate the Use of in Vitro Toxicity Data for Risk Assessment” published on Environmental Science and Technology1.

The model assumes instantaneous equilibrium partitioning between various phase present in the in vitro test system.

First, the freely dissolved aqueous concentration in the test system ($C_w$, in mol L$^{-1}$, M) is calculated as

$$C_w = \frac{M_T}{K_{AW}V_A + V_W + K_{Saw}V_{Sa} + K_{SIW}V_{SI} + K_{DW}V_D + K_{CW}V_C}$$

- $M_T$ – the total amount of chemical calculated from the nominal test concentration, mol.
- $K_{AW}$ – the air-water partition coefficient
- $V_A$ – the volume of head space
- $K_{Saw}$ – the serum albumin-water partition coefficient, calculated as
  $$\log K_{Saw} = 0.71 * \log K_{OW} + 0.42$$
- $V_{Sa}$ – the volume of serum albumin
- $K_{SIW}$ – the serum lipid-water partition coefficient, calculated as
  $$\log K_{SIW} = \log K_{OW}$$
- $V_{SI}$ – the volume of serum lipid
- $K_{DW}$ – the dissolved organic matter (DOM)-water partition coefficient, calculated as
  $$K_{DW} = \rho K_{OW}$$
  $\rho$ is a proportionality constant relating the sorption capacity of DOM to octanol, and is set as 0.05 in this study.
- $V_D$ – the volume of DOM
- $K_{CW}$ – the cell/tissue-water partition coefficient
  $$K_{CW} = f_L K_{OW}$$
  $f_L$ is the total lipid equivalent content of cells/tissue, and is set as 1 in this study.
- $V_C$ – the volume of cells/tissue.

Then, the cell/tissue concentration $C_C$ (M) is calculated as

$$C_C = C_w (or S_w, if C_w > S_w) * K_{CW}$$

$S_w$ is the water solubility limit of chemical.

*Noted that, all the input parameters have been adjusted with experiment temperature ($37^\circ$C) and ionic strength of working medium (0.105 for M9 minimal medium).

Reference:
Appendix 7.1 Real-time (temporal) gene expression profiles of 91 stress genes in *E.coli* in exposure to nTiO$_2$ a 10mg/L (left) and 50mg/L (right). X-axis top: natural log of induction factor (lnI). (Red spectrum colors indicate up-regulation, green spectrum colors indicate down-regulation) and time in minutes (the first data point shown is at 15 minutes after exposure due to moving average), Y-axis left: clustering of the profiles, Y-axis right: list of genes color-coded based on functional categorization.
Appendix 7.2 Temporal gene expression patterns observed upon exposure to nAg 10mg/L (up) and nTiO$_2$-a 10mg/L (down), determined using Self Organizing Map (SOM). X-axis: time in minutes, Y-axis left: Dotted lines represent the average natural log value of gene expression induction factor (lnI) for all the genes in each cluster with standard deviation shown as vertical bars. Values: 0, neutral; >0, up-regulated; <0, down-regulated. Right table lists all the genes in each cluster.
Appendix 7.3 Exemplary concentrations-dependent temporal gene expression profiles for representative genes upon exposure to nAg and nTiO$_2$-a exposure at various concentrations. $oxyR$, $cls$, and $cspB$ are the potential bio-markers for nAg exposure (right) and $mutT$, $sodB$ $pbpG$ are the three potential bio-markers for n_TiO$_2$-a exposure (left).
Appendix 9.1 Experiment setup for electron-fenton process

Anode

H₂O

O₂

H₂O₂

Fe²⁺●OH

CECs

Aromatics

Organic acids

CO₂

Cathode

TI/MMO
Appendix 9.2 Intermediate compounds identified during CECs transformation at 15 minutes

<table>
<thead>
<tr>
<th>Compounds</th>
<th>Chemical structure</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bisphenol A</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1,4-Benzoquinone</td>
<td><img src="" alt="Chemical structure" /></td>
<td>HPLC</td>
</tr>
<tr>
<td>Hydroquinone</td>
<td><img src="" alt="Chemical structure" /></td>
<td>HPLC</td>
</tr>
<tr>
<td>4-Isopropenylphenol</td>
<td><img src="" alt="Chemical structure" /></td>
<td>GC-MS</td>
</tr>
<tr>
<td>Bisphenol A</td>
<td><img src="" alt="Chemical structure" /></td>
<td>HPLC and GC-MS</td>
</tr>
<tr>
<td><strong>Triclosan</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-Chlorocatechol</td>
<td><img src="" alt="Chemical structure" /></td>
<td>HPLC</td>
</tr>
<tr>
<td>2,4-Dichlorophenol</td>
<td><img src="" alt="Chemical structure" /></td>
<td>HPLC and GC-MS</td>
</tr>
<tr>
<td>Hydroxylated triclosan</td>
<td><img src="" alt="Chemical structure" /></td>
<td>Logical speculation</td>
</tr>
<tr>
<td>Triclosan</td>
<td><img src="" alt="Chemical structure" /></td>
<td>HPLC and GC-MS</td>
</tr>
<tr>
<td><strong>Ibuprofen</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-Isobutylacetophenone</td>
<td><img src="" alt="Chemical structure" /></td>
<td>HPLC and GC-MS</td>
</tr>
<tr>
<td>Ibuprofen</td>
<td><img src="image" alt="Chemical Structure of Ibuprofen" /></td>
<td>HPLC and GC-MS</td>
</tr>
</tbody>
</table>
Appendix 9.3 Toxicty changes during the electro-Fenton oxidative transformation of CECs based on the Microtox results. EC50 refers to the concentration (reported as relative enrichment factor (REF)) that caused a 50% bioluminescence inhibition effect.

![Graph showing EC50 values over time for Bisphenol A, Triclosan, and Ibuprofen.](image)

Appendix 9.4 Changes in total organic carbon (TOC) values during the electro-Fenton transformation of three CECs.

![Graph showing TOC values over time for Bisphenol A, Triclosan, and Ibuprofen.](image)