

**Transforming chemical safety assessment using transcriptomics to
define toxicological potency and common mechanisms of BPA
replacements**

Gerónimo Parodi-Matteo

Thesis submitted to the University of Ottawa
in partial Fulfillment of the requirements for the
Doctorate in Philosophy Biology and Specialization in Bioinformatics

Department of Biology
Faculty of Science
University of Ottawa

© Gerónimo Parodi-Matteo, Ottawa, Canada, 2025

Abstract

Humans are exposed to chemicals that disrupt the function of endocrine tissues through the use of everyday consumer products and other environmental exposures. Bisphenol A (BPA) is used to manufacture polycarbonate plastics and epoxies and is a developing agent in thermal paper. BPA is an endocrine disruptor that acts on several nuclear hormone receptors, including the estrogen receptor. Most people have measurable levels of BPA in their urine, and exposure is associated with several negative health outcomes. In the late 2000s, BPA was regulated in Canada and other jurisdictions worldwide. This has led to the proliferation of chemical alternatives to BPA, sometimes in products labelled 'BPA-free'. Thus, humans are exposed to BPA alternatives, often as mixtures. Unfortunately, most BPA alternatives are data-poor, limiting regulatory toxicology assessment. The goal of my thesis was to use high-throughput transcriptomics to evaluate the potency of data-poor BPA alternative chemicals and characterize mechanisms of action *in vitro*. I produced global gene expression profiles to derive measures of biological activity and potency of 20 individual BPA alternatives and mixtures of these chemicals. Gene set enrichment analysis was used to identify perturbed biological pathways and upstream regulators. Transcriptomic biomarkers were applied to identify chemicals that activated the estrogen receptor alpha (ER α) and cellular stress response pathways. Bisphenol AF (BPAF) was the most potent chemical tested overall, activating the ER α at the lowest concentration, followed by BPA. Most BPA alternatives activated ER α and related pathways and enriched similar gene sets associated with dysregulated cell cycle control and cancer-related outcomes *in vitro*. Next, I used a concentration response model to test whether mixtures of BPA alternative chemicals have additive effects in breast cancer cells. Both equimolar and complex mixtures were

additive, activated the ER α , and had similar gene expression profiles as their individual components. Lastly, I sought to test BPA alternative chemicals in a more physiologically relevant model. Human mammary epithelial cells (HMECs) from multiple donors were exposed to a subset of BPA alternatives chemicals. These chemicals had similar potency rankings in MCF-7 and HMEC cells, with BPAF exhibiting the highest potency. Some of the BPA alternatives activated cancer-related pathways, while others inhibited the same pathways in HMECs. Overall, my thesis demonstrates that many BPA alternatives have comparable toxicological potencies to BPA, behave additively in mixtures, and have similar mechanisms of action *in vitro*. Importantly, this research identified BPAF as toxicologically potent BPA alternative chemical. Finally, these data demonstrate a pragmatic approach to the implementation of transcriptomics in regulatory toxicology and decision making.

Acknowledgements

To my doctoral supervisors, Dr. Carole Yauk and Dr. Ella Atlas, thank you for your mentorship and encouragement. I am especially grateful for your patience in showing me how to become a better communicator, scientist, and person. Thank you for believing in me and for giving me every opportunity to succeed. I've learned more than I could have ever imagined over these four years, and I am eternally grateful. To my thesis advisory committee, Dr. Jason O'Brien and Dr. Marc Ekker, thank you for your steady support and guidance. To my mentors, Dr. Peter Buist, Dr. Jeffrey M. Manthorpe, and Dr. J. David Miller, thank you for your continued guidance and inspiration.

To my GReAT lab, Atlas lab, and Health Canada colleagues, thank you for supporting me both inside and outside the lab. To David Schuster, Emmanuelle Monniez, and David Eickmeyer, thank you for all the laughs; here's to many more. To Jennifer Crosswait, Alice Kawata, Dr. Gregory Addicks, Dr. Lauren Bradford, Dr. Matthew Meier, Andrew Williams, Dr. Eunnara Cho, Andrea Rowan-Carrol, and Dr. Christopher Corton, thank you for your guidance and mentorship. A special thank you to Marc Rigden, I truly could not have done this without you.

To my friends and loved ones, thank you for your unwavering love and support. To my brothers Andrew McNaughton, Benjamin Fleishman, Lucas Solenthaler, Marcus Vaz, Mutasem Abu Hammad, Chawki-Jason Huard Berro, and Jethro G. Constant, you are the foundation of my well-being, and I'm endlessly grateful. To Michael Taylor and Fr. Harry McNeil, thank you for being with me every step of the way.

Finally, to the funding agencies that supported this work, thank you for investing in critical science to protect the health of Canadians.

Table of Contents

Chapter 1: Introduction	1
1.1 Environmental pollutants	1
1.1.1 Canadian legislation on environmental pollutants	1
1.1.2 Endocrine disrupting chemicals	2
1.1.3 Estrogen mimicking chemicals.....	3
1.2 Assessing chemicals for estrogen-like effects.....	4
1.2.1 New approach methodologies	5
1.3 Estrogens	6
1.3.1 Estrogen receptor	7
1.3.2 Mammary gland	8
1.3.3 Breast cancer cell lines	9
1.4 Bisphenol A.....	10
1.4.1 Bisphenol A regulations in Canada.....	11
1.4.2 Bisphenol A alternative chemicals	11
1.4.3 Mixtures of BPA alternative chemicals	13
1.5 Using transcriptomics to study BPA alternative chemicals.....	15
1.5.1 Estrogen receptor alpha transcriptomic biomarker	16
1.6 Thesis Overview	16
1.6.1 Rationale.....	17
1.6.2 Chapter Objectives	18
Chapter 2: In vitro transcriptomic analyses reveal pathway perturbations, estrogenic activities, and potencies of data-poor BPA alternative chemicals.....	23
2.1 Abstract.....	23
2.2 Introduction.....	24
2.3 Methods	28
2.3.1 Chemicals	28
2.3.2 Cell culture	28

2.3.3	Cell viability assay.....	29
2.3.4	Cell proliferation assay	30
2.3.5	TempO-Seq library building and next-generation sequencing.....	30
2.3.6	Data processing and quality control.....	32
2.3.7	Gene expression biomarker analysis.....	34
2.3.8	Points of departure analysis	35
2.3.9	Pathway and upstream regulator enrichment analysis	37
2.4	Results	38
2.4.1	Cell viability assay.....	38
2.4.2	Cell proliferation assay	38
2.4.3	General HTTr data summary	39
2.4.4	Stress response biomarker activation.....	39
2.4.5	Final sample inclusion and exclusion for ER α activity, potency comparison, and enrichment analysis	40
2.4.6	Identification of ER α receptor agonists and potency ranking for ER α effects 41	
2.4.7	Potency ranking using BMC analysis	42
2.4.8	Pathway and Upstream Regulator Enrichment analysis.....	44
2.5	Discussion.....	46
2.6	Tables and Figures	52
Chapter 3: High-throughput transcriptomics analysis of equipotent and human relevant mixtures of BPA alternatives reveal additive effects in vitro		
3.1	Abstract.....	57
3.2	Introduction.....	58
3.3	Materials and methods.....	61
3.3.1	Chemicals and mixtures.....	61
3.3.2	Cell culture	62
3.3.3	Cell viability assay.....	63
3.3.4	TempO-Seq library building and next-generation sequencing.....	64
3.3.5	Data processing and quality control.....	65

3.3.6	Gene expression biomarker analysis.....	67
3.3.7	Transcriptomic point of departure analysis	68
3.3.8	Mixture modelling.....	70
3.3.9	Ingenuity Pathway Analysis	71
3.4	Results	72
3.4.1	Cell viability	72
3.4.2	General transcriptomic data summary	72
3.4.3	Stress response biomarker activation.....	73
3.4.4	ER α biomarker activation.....	74
3.4.5	Transcriptomic Point of Departure (tPOD) derivation	75
3.4.6	Mixture modelling.....	77
3.4.7	Ingenuity Pathway Analysis	78
3.5	Discussion.....	80
3.6	Tables and Figures	86
Chapter 4: Transcriptomic analysis of BPA alternative chemicals in human mammary epithelial cells		93
4.1	Abstract.....	93
4.2	Introduction.....	94
4.3	Materials and methods.....	97
4.3.1	Chemicals	97
4.3.2	Cell culture	98
4.3.3	Cell viability assay.....	99
4.3.4	TempO-Seq library building and next-generation sequencing.....	99
4.3.5	Data processing.....	101
4.3.6	Quality control	101
4.3.7	Transcriptomic biomarkers.....	102
4.3.8	Transcriptomic point of departure analysis	104
4.3.9	Average donor transcriptomic point of departure.....	105
4.3.10	Ingenuity Pathway Analysis	106

4.4	Results	107
4.4.1	Only the highest concentration of BPAF reduced cell viability in HMECs .	107
4.4.2	Most BPA alternatives induced broad transcriptional changes in donors..	107
4.4.3	Transcriptomic biomarker analysis reveals cellular stress response and no ER α activation	109
4.4.4	BPA alternatives had similar toxicological potencies in HMECs.....	110
4.4.5	BPA alternatives oppositely enriched gene sets associated with cancer related outcomes.....	111
4.5	Discussion.....	113
4.6	Tables and Figures	119
	Chapter 5: Final Discussion and Conclusion	125
5.1	Summary of Outcomes.....	125
5.2	Contributions to Scientific Knowledge.....	132
5.3	Future Directions	134
5.4	Concluding Remarks	140
	Appendix A: A 50-gene biomarker identifies estrogen receptor-modulating chemicals in a microarray compendium	157
A.1	Abstract	157
A.2	Introduction	158
A.3	Methods.....	162
A.4	Results.....	173
A.4.1.	Assembly and characterization of a 50-gene ER biomarker.....	173
A.4.2	Comparison of the biomarker to individual biosets used to create the biomarker 174	
A.4.3	Functional classification of the biomarker genes	175
A.4.4	The biomarker identifies ER modulators in ER-positive breast cancer cell lines but not other cell lines	176
A.4.5	The biomarker identifies genetic perturbations of the ESR1 gene	179
A.4.6	Determination of the predictive accuracy of the ER biomarker.....	180

A.4.7 Application of the biomarker to predictions in HTTr data: identification and ranking of BPA alternatives	185
A.5 Discussion	186
A.6 Tables and Figures.....	194
Appendix B: From vision toward best practices: Evaluating <i>in vitro</i> transcriptomic points of departure for application in risk assessment using a uniform workflow	203
B.1 Abstract	203
B.2 Introduction	205
B.3 Materials and methods	208
B.3.1 Study design	208
B.3.2 Datasets and exposure conditions.....	209
B.3.3 Data processing and handling	211
B.3.4 Derivation of <i>in vitro</i> points of departure using transcriptomic concentration-response data	213
B.3.4.1 The fifth percentile	215
B.3.4.2 The first mode	215
B.3.4.3 The 25th lowest ranked gene BMC	216
B.3.4.4 The LCRD	216
B.3.4.5 The lowest gene set.....	217
B.3.5 Conversion of transcriptomic points of departure to administered equivalent doses	218
B.3.6 Comparison of administered equivalent dose to corresponding <i>in vivo</i> data..	219
B.4 Results and discussion	220
B.4.1 A transcriptomic workflow for a diverse chemical space	220
B.4.2 A general comparison of approaches	222
B.4.3 Conversion to human relevant exposure values.....	227
B.4.4 Comparing <i>in vitro</i> to <i>in vivo</i> derived data.....	231
B.4.5 Identifying outliers using ratios	234
B.4.6 Examining design considerations and uncertainty when applying transcriptomic data for risk assessment	236

B.4.6.1 Study design	236
B.4.6.2 Uncertainty relating to apical endpoints	237
B.4.6.3 Uncertainty relating to <i>in vitro</i> models to derive AEDs	239
B.5 Conclusion.....	242
B.6 Tables and Figures.....	243
Appendix C: Supplementary Material	261
C.1 Chapter 2.....	261
C.2 Chapter 3.....	265
C.3 Chapter 4.....	273

List of Tables

Table 2.1: List of chemicals and suppliers.

Table 3.1: List of chemicals and suppliers.

Table 3.2: List of mixtures and relative proportions.

Table 3.3: Comparison of transcriptomic points of departure in breast cancer cells exposed to mixtures of BPA alternatives and their individual components.

Table 4.1: List of chemicals and suppliers.

Table 4.2: Comparison of transcriptomic points of departure in human mammary epithelial cells from three biological donors exposed to BPA and four alternatives.

List of Figures

Figure 2.1: Comparison of ER α bioactivity and transcriptomic points of departure (tPODs) derived from whole transcriptome analysis in MCF-7 cells ($n = 3 - 4$ per concentration) exposed to BPA and 15 alternative chemicals at a range of concentrations (0.0005 – 100 μ M) for 48 hours.

Figure 2.2: Gene accumulation plot of MCF-7 cells ($n = 3 - 4$ per concentration) exposed to BPA and 15 alternative chemicals (0.0005 – 100 μ M) for 48 hours.

Figure 2.3: Ingenuity Pathway Analysis of genes fitting BMC models in the MCF-7 cells ($n = 3 - 4$ per concentration) exposed to bisphenol A (BPA) and 15 alternative chemicals for 48 hours, as well as positive control 17 β -estradiol.

Figure 3.1: Chemical structure of bisphenol A (BPA) and alternative chemicals.

Figure 3.2: Gene accumulation plot of MCF-7 cells ($n = 2 - 4$ per concentration) exposed to seven mixtures of bisphenol A (BPA) alternatives and 12 individual chemicals at a range of concentrations (0.001 – 50 μ M) for 48 h.

Figure 3.3: Comparison of transcriptomic points of departure (tPODs) derived from whole transcriptome analysis of MCF-7 cells ($n = 2 - 4$ per concentration) that were exposed to seven mixtures of bisphenol A (BPA) alternatives and 12 individual chemicals at a range of concentrations (0.001 – 50 μ M) for 48 h.

Figure 3.4: Ingenuity Pathway Analysis of all genes fitting benchmark concentration models in MCF-7 cells ($n = 2 - 4$ per concentration) that were exposed to seven mixtures of bisphenol A (BPA) alternatives and 12 individual chemicals at a range of concentrations (0.001 – 50 μ M) for 48 h.

Figure 4.1: Chemical structures of bisphenol A (BPA) alternative chemicals.

Figure 4.2: Gene accumulation plot of human mammary epithelial cells ($n = 2 - 4$ per concentration) for the first 100 genes fitting benchmark concentrations (BMCs) from three biological donors exposed to bisphenol A (BPA) and four alternatives at a range of concentrations (0.001 – 50 μ M) for 48 h.

Figure 4.3: Comparison of transcriptomic points of departure (tPODs) derived from whole transcriptome analysis of human mammary epithelial cells ($n = 2 - 4$ per concentration) from three biological donors exposed to bisphenol A (BPA) and four alternatives at a range of concentrations (0.001 – 50 μ M) for 48 h.

Figure 4.4: Ingenuity Pathway Analysis of all genes fitting benchmark concentration models in human mammary epithelial cells ($n = 2 - 4$ per concentration) from three biological donors exposed to bisphenol A (BPA) and four alternatives at a range of concentrations (0.001 – 50 μ M) and 17 β -estradiol (E2; 1 and 10 nM) for 48 h.

List of Abbreviations

BMC	Benchmark concentration
BPA	Bisphenol A
BPAF	Bisphenol AF
BPC	Bisphenol C
BPF	4,4'-bisphenol F
BPS	4,4'-bisphenol S
TMBPF	Tetramethyl 4,4'-bisphenol F
DO	Detector oligonucleotide
EDC	Endocrine disrupting chemical
ER α	Estrogen receptor alpha
HTTr	High-throughput transcriptomics
HMEC	Human mammary epithelial cells
IPA	Ingenuity Pathway Analysis
MCF-7	Michigan Cancer Foundation Cells Passage 7
TempO-Seq	Templated Oligonucleotide Sequencing
tPOD	Transcriptomic point of departure

Statement of Contributions

Chapter 2: *In vitro* transcriptomic analyses reveal pathway perturbations, estrogenic activities, and potencies of data-poor BPA alternative chemicals

Published in: *Toxicological Sciences*, 191(2):266-275, 17 Feb 2023

DOI: 10.1093/toxsci/kfac127

Authors: Geronimo Matteo^{1,2}, Karen Leingartner², Andrea Rowan-Carroll², Matthew Meier², Andrew Williams², Marc A Beal³, Matthew Gagné⁴, Reza Farmahin⁴, Shamkia Wickramasuriya⁴, Anthony JF Reardon⁴, Tara Barton-Maclaren⁴, J. Christopher Corton⁵, Carole L Yauk^{1,2}, Ella Atlas^{2,6}

Affiliations:

¹Department of Biology, University of Ottawa

²Environmental Health Science and Research Bureau, Health Canada

³Bureau of Chemical Safety, Health Canada

⁴Existing Substances Risk Assessment Bureau, Health Canada

⁵Center for Computational Toxicology and Exposure, US Environmental Protection Agency

⁶Department of Biochemistry, University of Ottawa

Experimental design:	Carole Yauk, Ella Atlas
Manuscript writing:	Geronimo Matteo
Cell culture and chemical exposure:	Karen Leingartner
TempO-Seq library building:	Karen Leingartner
Next-Generation Sequencing:	Andrea Rowan-Carroll
Bioinformatic data processing:	Matthew Meier
Statistical analysis:	Andrew Williams
Data analysis and interpretation:	Geronimo Matteo
Transcriptomic biomarker analysis:	Christopher Corton
Regulatory input for application:	Marc Beal, Matthew Gagné, Reza Farmahin, Shamkia Wickramasuriya, Anthony Reardon, Tara Barton-Maclaren
Manuscript review and approval:	Carole Yauk, Ella Atlas, Karen Leingartner, Andrea Rowan-Carroll, Matthew Meier, Andrew Williams, Marc Beal, Matthew Gagné, Reza

Farmahin, Shamkia Wickramasuriya, Anthony Reardon, Tara Barton-Maclaren, Christopher Corton

Chapter 3: High-throughput transcriptomics analysis of equipotent and human relevant mixtures of BPA alternatives reveal additive effects *in vitro*

Published in: *Archives of Toxicology*, 99(9):3707-3719, 18 June 2025

DOI: 10.1007/s00204-025-04110-3

Authors: Geronimo Matteo^{1,2}, Eunnara Cho¹, Marc Rigden¹, David C. Eickmeyer², Lauren M. Bradford¹, Matthew J. Meier¹, Andrew Williams¹, J. Christopher Corton³, Carole L. Yauk², Ella Atlas^{1,4}

Affiliations:

¹Environmental Health Science and Research Bureau, Health Canada

²Department of Biology, University of Ottawa

³Center for Computational Toxicology and Exposure, US Environmental Protection Agency

⁴Department of Biochemistry, University of Ottawa

Experimental design:	Geronimo Matteo, Carole Yauk, Ella Atlas
Manuscript writing:	Geronimo Matteo
Cell culture and chemical exposure:	Geronimo Matteo, Marc Rigden
TempO-Seq library building:	Geronimo Matteo, Eunnara Cho
Next-Generation Sequencing:	David Eickmeyer
Bioinformatic data processing:	Lauren Bradford, Matthew Meier
Statistical analysis:	Andrew Williams
Data analysis and interpretation:	Geronimo Matteo
Transcriptomic biomarker analysis:	Christopher Corton, Andrew Williams
Manuscript review and approval:	Carole Yauk, Ella Atlas, Eunnara Cho, Marc Rigden, David Eickmeyer, Lauren Bradford, Matthew Meier, Andrew Williams ¹ Christopher Corton

Chapter 4: Transcriptomic analysis of BPA alternative chemicals in human mammary epithelial cells

Intended journal: Cell Biology and Toxicology

Authors: Geronimo Matteo^{1,2}, Andrea Rowan-Carrol¹, Lauren M. Bradford¹, Matthew J. Meier¹, Andrew Williams¹, J. Christopher Corton³, Carole L. Yauk² Ella Atlas^{1,4}

Affiliations:

¹Environmental Health Science and Research Bureau, Health Canada

²Department of Biology, University of Ottawa

³Center for Computational Toxicology and Exposure, US Environmental Protection Agency

⁴Department of Biochemistry, University of Ottawa.

Experimental design:	Geronimo Matteo, Carole Yauk, Ella Atlas
Manuscript writing:	Geronimo Matteo
Cell culture and chemical exposure:	Geronimo Matteo
TempO-Seq library building:	Geronimo Matteo
Next-Generation Sequencing:	Andrea Rowan-Carrol
Bioinformatic data processing:	Lauren Bradford, Matthew Meier
Statistical analysis:	Andrew Williams
Data analysis and interpretation:	Geronimo Matteo
Transcriptomic biomarker analysis:	Christopher Corton, Andrew Williams
Manuscript review and approval:	Carole Yauk, Ella Atlas, Andrea Rowan-Carrol, Lauren Bradford, Matthew Meier, Andrew Williams, Christopher Corton

Appendix A: A 50-gene biomarker identifies estrogen receptor-modulating chemicals in a microarray compendium

Published in: *Chemico-Biological Interactions*, 394:110952, 1 May 2024

DOI: 10.1016/j.cbi.2024.110952

Authors: J. Christopher Corton¹, Geronimo Matteo^{2,3}, Brian Chorley¹, Jie Liu¹, Beena Vallanat¹, Logan Everett¹, Ella Atlas², Matthew J. Meier², Andrew Williams², Carole Lyn Yauk³

Affiliations:

¹ Center for Computational Toxicology and Exposure, US Environmental Protection Agency, Research Triangle Park, NC, 27711, USA

² Environmental Health Science and Research Bureau, Health Canada, Ottawa, ON, K1A 0K9, Canada

³ Department of Biology, University of Ottawa, Ottawa, ON, K1N 6N5, Canada

Manuscript writing: Christopher Corton

Cell culture and chemical exposure: Geronimo Matteo

Data analysis: Geronimo Matteo, Brian Chorley, Jie Liu, Beena Vallanat, Logan Everett, Ella Atlas, Matthew Meier

Statistical analysis: Andrew Williams

Conceptualization: Carole Yauk, Christopher Corton

Manuscript Review and Approval: Geronimo Matteo, Brian Chorley, Jie Liu, Beena Vallanat, Logan Everett, Ella Atlas, Matthew Meier, Andrew Williams, Carole Yauk

Appendix B: From vision toward best practices: Evaluating in vitro transcriptomic points of departure for application in risk assessment using a uniform workflow

Published in: *Frontiers of Toxicology*, 5:1194895, 23 May 2023

DOI: 10.3389/ftox.2023.1194895

Authors: Anthony JF Reardon¹, Reza Farmahin¹, Andrew Williams², Matthew J Meier², Gregory C Addicks², Carole L Yauk³, Geronimo Matteo³, Ella Atlas^{2,4}, Joshua Harrill⁵, Logan J Everett⁵, Imran Shah⁵, Richard Judson⁵, Sreenivasa Ramaiahgari⁶, Stephen S Ferguson⁶, Tara S Barton-Maclaren¹

Affiliations:

¹Existing Substances Risk Assessment Bureau, Healthy Environments and Consumer Safety Branch, Health Canada, Ottawa, ON, Canada.

²Environmental Health Science and Research Bureau, Healthy Environments and Consumer Safety Branch, Health Canada, Ottawa, ON, Canada.

³Department of Biology, University of Ottawa, Ottawa, ON, Canada.

⁴Department of Biochemistry, University of Ottawa, Ottawa, ON, Canada.

⁵Center for Computational Toxicology and Exposure, US Environmental Protection Agency, Durham, NC, United States.

⁶Division of Translational Toxicology, Mechanistic Toxicology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Durham, NC, United States.

Manuscript writing: Anthony Reardon, Reza Farmahin, Tara Barton-Maclaren

Bioinformatics: Andrew Williams, Matthew Meier

Data analysis and interpretation: Carole Yauk, Geronimo Matteo, Ella Atlas, Joshua Harrill, Logan Everett, Imran Shah, Richard Judson, Sreenivasa Ramaiahgari, Stephen Ferguson

Manuscript Review and Approval: Reza Farmahin, Andrew Williams, Matthew Meier, Gregory Addicks, Carole Yauk, Geronimo Matteo, Ella Atlas, Joshua Harrill, Logan Everett, Imran Shah, Richard Judson, Sreenivasa Ramaiahgari, Stephen Ferguson, Tara S Barton-Maclaren

Chapter 1: Introduction

1.1 Environmental pollutants

Environmental pollution, driven largely by industrial activity, contributes to ecosystem disruption and human health risks on a global scale. Substances (e.g., chemicals, materials) produced by industry enter air, soil, and/or water through their use and disposal, leading to widespread contamination of ecosystems. Some substances are chemically transformed in the environment, but others persist almost indefinitely (e.g., per- and polyfluorinated alkyl substances). Humans have measurable levels of environmental pollutants in their bodies, and their exposure is associated with negative health outcomes. Coordinated efforts are necessary to monitor and reduce pollution, identify chemical hazards, and manage exposure and environmental release.

1.1.1 Canadian legislation on environmental pollutants

The Canadian Environmental Protection Act, 1999 (CEPA, 1999) aims to reduce the negative impacts of substances on human and environmental health. It provides the Canadian federal government with a broad mandate to manage pollution and its sources, assess, regulate, and phase out harmful substances by labeling them as toxic, and facilitate cooperation with stakeholders across sectors. The Domestic Substances List (DSL) catalogues all commercial substances manufactured in or imported into Canada

and currently contains approximately 28,000 substances. An additional 400–500 new substances are added each year, and CEPA mandates that these undergo assessments for human and environmental health risks. Hazard and exposure data inform chemical risk assessments and the development of risk management strategies to reduce potential harms. To address growing public concerns about chemical safety, the Chemicals Management Plan (CMP) was launched in 2006 to systematically evaluate and manage substances in Canada that may pose risks to human or environmental health. The CMP prioritizes and reviews substances on the DSL, engages stakeholders at multiple levels, and oversees ongoing environmental and human biomonitoring programs.

1.1.2 Endocrine disrupting chemicals

The endocrine system is composed of a network of glands that synthesize, store, and secrete hormones into the bloodstream to regulate vital functions in the body. These hormones primarily belong to the estrogen, androgen, thyroid, and steroidogenesis pathways. Endocrine disrupting chemicals (EDCs) are those perturb any aspect of endocrine system function like receptor function, hormone synthesis, transport, metabolism, epigenetics, cellular fate, among others (La Merrill *et al.*, 2020). Most of these chemicals are exogenous and synthetically manufactured, though EDCs of natural origin exist. Humans are generally exposed to EDCs through food, contact with commercial products (e.g., cosmetic products) and their environment (e.g., freshwater). While some EDCs can bioaccumulate in tissues for years (e.g., polychlorinated biphenyls), others are

metabolized and excreted within hours (e.g., phthalates). Due to their widespread use and environmental contamination (Y. Chen *et al.*, 2022), the general population is chronically exposed to low doses of EDCs. Given the negative health outcomes associated with EDC exposure, many jurisdictions worldwide have regulated EDCs to protect public health. Substances with potential endocrine-related effects are among those prioritized by CEPA for regulatory toxicological review.

1.1.3 Estrogen mimicking chemicals

Diethylstilbestrol (DES) is a synthetic estrogen-like medication prescribed between the mid 1940's to the late 1970's to prevent miscarriages, despite evidence that it was ineffective for this purpose (Al Jishi and Sergi, 2017). Follow-up studies of mothers prescribed DES revealed that they were more likely to experience early menopause, infertility, cancer, and pregnancy complications. In addition, *in utero* DES exposure adversely impacted the reproductive health of both male and female offspring, including an increased risk of rare cancers and structural abnormalities. Due to growing evidence of its risks, the U.S. Food and Drug Administration recommended that doctors no longer prescribe DES to pregnant mothers in 1971. By then, millions of women had been exposed. The DES story raised public awareness of the long-term risks posed by estrogen-like chemicals and contributed to the regulation of EDCs. However, human

exposure to estrogen-like chemicals through consumer products and their environments remains a concern to this day.

1.2 Assessing chemicals for estrogen-like effects

Regulatory screening of potential estrogen-like chemicals typically relies on tiered testing strategies to identify hazards. For example, the US Environmental Protection Agency's Endocrine Disruptor Screening Program uses a two-stage system in which chemicals are screened for endocrine activity and then evaluated for adverse effects (US Environmental Protection Agency, 2012b). The European Union incorporates similar approaches under frameworks like REACH, and Canada is developing comparable strategies through the CMP. To support these efforts, the Organisation for Economic Co-operation and Development (OECD) convenes expert panels to validate *in vitro* and *in vivo* assays, establishing international standards for regulatory toxicology.

In the first stage of chemical screening, *in silico* methods like Quantitative Structure-Activity Relationships (QSAR) are used to predict the ability of chemicals to interact with the estrogen receptor (ER) based on their chemical structure. This is followed by validated *in vitro* binding assays (e.g., OECD TG 455, 457, 493) that are used to test the ability of the chemical to agonize or antagonize the ER (OECD, 2021, 2012, 2024), along with assays that measure cellular proliferation in estrogen-sensitive cell lines (e.g., E-SCREEN; Soto et al., 1995). For second stage, *in vivo* assays are traditionally used to

determine adverse effects in estrogen-sensitive tissues. For example, a uterotrophic assay (OECD TG 440) can be used to determine the effects of short-term chemical exposure on uterine size in immature or ovariectomized female rodents (OECD, 2007). Longer term *in vivo* toxicity assays (e.g., OEC TG 443) can be leveraged to determine the effects of parental chemical exposure on multiple endpoints like reproductive organ development across multiple generations (OECD, 2018). However, conventional *in vitro* assays are generally low-throughput, and *in vivo* assays are time-consuming and resource-intensive. Thus, novel approaches are needed to meet growing demands for screening new and data-poor substances (Krewski *et al.*, 2010).

1.2.1 New approach methodologies

New approach methodologies (NAMs) have been developed to support chemical hazard assessment using high-throughput *in vitro*, *ex vivo*, *in silico*, and *in chemico* methods—significantly reducing reliance on conventional whole-animal testing assays (Stucki *et al.*, 2022). Recent amendments to CEPA reinforce this approach by promoting reduced vertebrate animal use and prioritizing NAMs in risk assessment (Health Canada and Environment and Climate Change Canada, 2023). Among these, high-throughput transcriptomics (HTTr) is a powerful tool to screen chemicals of interest (i.e., environmental chemicals, new or data-poor chemicals, etc.) for possible adverse biological responses (Jennings *et al.*, 2013).

HTTr has been increasingly used to identify hazards and establish toxicological points of departure (POD) for regulatory toxicology (Thomaset al., 2013, 2019). One important approach to of establishing a POD is to conduct benchmark concentration (BMC) analysis. Specifically, the BMC refers to a concentration that produces a predetermined increase (benchmark response) in a given parameter like a gene or a biological pathway (Crump, 1984). This approach has several advantages compared to traditional methods to derive PODs such as the no- or lowest observed adverse effect level (NOAEL/LOAEL) (National Toxicology Program, 2018). For example, the BMC modelling is less sensitive to study design (e.g., sample size and concentration selection), incorporates information from the entire concentration-response curve (i.e., as opposed to a single point), and produces confidence intervals that facilitate cross-study comparisons (Hardy *et al.*, 2017; US Environmental Protection Agency, 2012a). Integrating of NAMs like HTTr into tiered testing strategies will enable regulatory toxicology agencies to more efficiently screen substances for estrogen-like effects than traditional assays alone.

1.3 Estrogens

Estrogens interact with nuclear estrogen receptors (ERs) to mediate the menstrual cycle, reproduction, lactation, sexual maturation, and breast development, as well as playing roles in bone density, brain function, cholesterol mobilization, and the immune system in biological females (Fuentes and Silveyra, 2019). They are a class of 18 carbon

(C₁₈H₂₄O₂) steroid hormones derived from cholesterol including estrone, estradiol, estriol, and estetrol. Estradiol, the primary estrogen in biological females, is synthesized from estrone by 17 β -hydroxysteroid dehydrogenase and is secreted by granulosa cells of the ovarian follicles and the corpus lutea, as well as the adrenal glands and adipose tissue to a lesser extent (Fuentes and Silveyra, 2019). Estradiol promotes epithelial cell proliferation in the uterine endometrium and mammary glands during puberty. Throughout pregnancy, estradiol supports the development of the placenta and lactation. In menopause, ovarian function decreases with senescence and estradiol levels can decline (Camon *et al.*, 2024).

1.3.1 Estrogen receptor

The estrogen receptor alpha (ER α) was the first ER cloned from Michigan Cancer Foundation Passage 7 (MCF-7) cells in mid 1980s and the ER beta (ER β) was discovered several years later (P. Chen *et al.*, 2022). The ER α is encoded by ESR1 on chromosome 6 locus 6q25.1 and has multiple isoforms made from alternative start codons and splicing. The ER β is encoded by ESR2 on chromosome 14 locus 14q23-24 and has multiple isoforms. Both receptors have the same principle functional domains including A/B, C, D, and E/F. The A/B region is the amino-terminal domain (N terminal domain; NTD) that mediates gene transcription and transactivation via a zinc-finger that binds to target sequences (P. Chen *et al.*, 2022). The C region is the DNA binding domain (DBD), that

facilitates estrogen receptor homodimerization and interacting with specific sequences of chromatin, known as estrogen response elements (EREs), near promoters and/or 3'-untranslated regions of target genes to affect gene expression. The D domain connects the C and E domains, associates with chaperone proteins, and contains the nuclear localization signal that is unmasked when estrogen binds. The E/F region is at the carboxy-terminal and contains the ligand binding domain as well as binding sites for coactivators and corepressors. Two additional regulators of ER transcriptional activity are activation function 1 and 2, which are located within the NTD and DBD, respectively (P. Chen *et al.*, 2022).

1.3.2 Mammary gland

The mammary glands are hormonally regulated organs that develop through distinct life stages and are sensitive to endocrine disruption. Breast tissue forms *in utero* from the ectodermal germ layer and develops further in puberty as part of sexual maturation (Biswas *et al.*, 2022). During pregnancy, the mammary gland undergoes significant remodelling to produce milk that is partially reversed following weaning. Adult mammary glands are composed of two compartments: the epithelium and the mesenchyme/stroma. The former contains luminal epithelial cells surrounded by myoepithelial cells that form the ducts and alveoli necessary for lactation. The latter structure is composed of various cell types (e.g., adipocytes, immune cells, mesenchymal

stem cells, fibroblasts) and connective tissue that support the epithelium. Endocrine hormones like estrogens, progesterone, androgens, and prolactin, along with their respective enzymes and receptors, are crucial for the development, maintenance, and function of the mammary glands throughout life (Hannan *et al.*, 2023). The mammary glands are considered an EDC sensitive organs as exposure can negatively impact their development and function (Vandenberg, 2021).

1.3.3 Breast cancer cell lines

Breast cancer is the most common cancer diagnosed and a major cause of cancer-related death in biological women (Clusan *et al.*, 2023). While of diverse etiology, most breast cancers test positive for the ER α , while only a subpopulation of non-cancerous breast epithelial cells express the receptor (Hua *et al.*, 2018). By contrast, the ER β is expressed to a lesser extent in breast cancers, and its activation is generally anti-proliferative in breast cancer cells (Hua *et al.*, 2018). Estradiol and chemicals that activate the ER α generally cause proliferative effects in ER α positive breast cancer cells (Darbre, 2021). These data support that overexpression and activation of the ER α in breast epithelial cells is associated with cancer-related outcomes, but many of the underlying mechanisms remain unknown. To this end, several immortalized breast cell lines have been established. For example, MCF-7 breast cancer cells that overexpress the ER α were central to the discovery of anti-cancer drugs like the selective ER modulator

tamoxifen (Lee *et al.*, 2015). An important application of MCF-7 cells in regulatory toxicology has been to screen chemicals for their ability to perturb the estrogen system and cause endocrine dysfunction.

1.4 Bisphenol A

Bisphenol A (BPA) is a chemical used to manufacture polycarbonate plastics and epoxy resins used in a variety of consumer products such as plastic containers, can lining, dental materials, thermal paper, and more (Manzoor *et al.*, 2022). Stressors like mechanical abrasion, exposure to acids, detergents, and/or heat can release individual BPA molecules from polycarbonate plastic (Khalili Sadrabad *et al.*, 2023). Due to its widespread use, BPA is a global environmental contaminant (Corrales *et al.*, 2015) and is detectable in urine of most humans throughout the world (Colorado-Yohar *et al.*, 2021). Diet is the primary route of exposure (Geens *et al.*, 2012) as BPA is detected in food products such as beverages (Mandrah *et al.*, 2017), canned food (Cunha *et al.*, 2012), packaged meals (Regueiro and Wenzl, 2015), baby food (García-Córcoles *et al.*, 2018), infant formula (Cunha *et al.*, 2011), cooking oil (Xian *et al.*, 2017), and animal products (Deceuninck *et al.*, 2019). BPA is considered an EDC (Diamanti-Kandarakis *et al.*, 2009) since it affects hormonal signalling through interactions with the estrogen, androgen, thyroid, and glucocorticoid receptors (Cimmino *et al.*, 2020). BPA exposure is also associated with inflammation, oxidative stress, epigenetic changes, and possibly DNA

damage (Ma *et al.*, 2019). Exposure to BPA is associated with negative health outcomes including metabolic and endocrine dysfunction, reproductive and developmental disorders, and hormone-related cancers (Ma *et al.*, 2019).

1.4.1 Bisphenol A regulations in Canada

Due to public health concerns, Health Canada announced a chemical assessment on BPA in 2008. Later that year, Canada became the first country to label BPA as toxic through CEPA. In 2010, BPA was banned in baby bottles in Canada, and many jurisdictions worldwide took similar regulatory action. Regulatory restrictions on BPA spurred the use of alternative chemicals, frequently marketed in products as ‘BPA free’. Unfortunately, many of these chemicals are data-poor, limiting regulatory toxicology assessment. In 2020, the Government of Canada issued a notice under CEPA to collect information on 188 BPA alternative chemicals to inform future risk assessments.

1.4.2 Bisphenol A alternative chemicals

A review of BPA alternative chemicals found that exposure to several bisphenols was associated with endocrine disruption, reproductive toxicity, and carcinogenicity in some cases (den Braver-Sewradj *et al.*, 2020). Several BPA structural analogue chemicals have been detected in individuals, including bisphenol S (BPS), 4,4'-bisphenol F (BPF),

bisphenol AF (BPAF), bisphenol AP (BPAP), bisphenol B (BPB), bisphenol C (BPC), bisphenol E (BPE), and others (Pelch *et al.*, 2019; Kojima *et al.*, 2019; Keminer *et al.*, 2020). The most commonly detected BPA alternatives are BPS and 4,4'-BPF and a systematic review highlighted their ability to act as endocrine disruptors at similar concentrations to BPA (Rochester and Bolden, 2015). These data are supported by a growing body of evidence that associates exposure to BPA alternatives with negative health outcomes. As a result, there is a clear need to identify hazards of BPA alternatives and determine relative potencies to BPA. Unfortunately, while some BPA alternative chemicals like BPS and 4,4'-BPF are better studied, many others remain data-poor (den Braver-Sewradj *et al.*, 2020).

An important approach to screening BPA alternatives is to specifically compare potencies in interacting with nuclear hormone receptors. For example, one study tested the effects of exposure to several BPA alternatives (0.01 μM – 100 μM) for 24 hours using MCF-7 cells transfected with an ERE luciferase reporter. Interestingly, some BPA alternatives, including BPAF and BPB, activated the ER α at lower concentrations than BPA, whereas others like BPS and 4,4'-BPF required higher concentrations (Kitamura *et al.*, 2005). This is in line with another study using a similar method to measure ER α activation using MCF-7 cells and T47D-KBluc cells exposed to 6 BPA alternatives (0.1 nM - 100 μM) for 48 hours (Mesnage *et al.*, 2017). Indeed, this study highlighted the ability of BPA alternatives including BPAF, bisphenol Z (BPZ), BPB to cause cell proliferation

and stimulate the ER at lower concentrations than BPA, with ER α activation corroborated by co-exposure to an ER α antagonist (Mesnage *et al.*, 2017). These data, along with others (Rosenmai *et al.*, 2014), collectively suggest that exposure to BPA alternatives have similar endocrine disrupting effects as BPA *in vitro*.

1.4.3 Mixtures of BPA alternative chemicals

Epidemiological studies indicate that most individuals are exposed to mixtures of BPA alternative chemicals. Indeed, several studies highlight that individuals have urinary levels of BPA structural analogues that are comparable to BPA (Liu *et al.*, 2019; Lucarini *et al.*, 2020). In particular, BPS and 4,4'-BPF are routinely detected in the general population, and this is in line with their increased use in commercial products as BPA is phased-out in some industries (Chen *et al.*, 2016). Currently, it is assumed that exposure to mixtures of BPA alternative effects cause additive effects, however there is little data to support this. Therefore, there is an urgent need for regulators to understand interactions between bisphenols in mixtures and establish their relative potencies.

Few studies have focused on the effects of exposure to mixtures of bisphenols. One group tested binary mixtures of bisphenols including BPA, BPC, BPE, 4,4'-BPF, BPG, BPM, BPP, BPS, BPZ, and BPFL using the Microtox and XenoScreen yeast estrogen screen (YES) and yeast androgen screen (YAS) assay (Owczarek *et al.*, 2018). Individually, 7 of the 10 BPA alternative chemicals tested caused toxicity in the Microtox

assay with EC50s within 3 – 60 μ M, and 6 of the 10 were endocrine active in the XenoScreen YES/YAS assay with a wide range of micromolar LOECs. Interestingly, BPS showed synergistic activity with 4 bisphenols: BPA, BPC, BPE, and 4,4'-BPF. These data must be interpreted with caution due to the limitations of the assays used and warrant further investigation using *in vitro* models. Another study exposed HeLa9903 cells to BPA, BPS, or 4,4'-BPF, either individually or as a mixture of three, at a range of concentrations (100 pM – 100 μ M) and measured nuclear hormone activation (Park *et al.*, 2020). Interestingly, the mixture of the three bisphenols activated the ER α , using a transfected ER α luciferase assay, at an order of magnitude lower concentration than the individual BPA alternatives. In a sub-experiment, the authors found that the mixture of BPA alternative chemicals also caused increased ER α degradation in MCF-7 cells at a lower concentration than the individual mixture components. Like the effects seen with the ER α , the bisphenols mixture decreased AR activity in AR-EcoScreen cells at a lower concentration relative to individual BPA alternatives. However, there was no difference between the mixture and its individual components in their ability to decrease AhR activity. Overall, these data suggest that exposure to binary mixtures of BPA alternative chemicals may cause effects that are greater than individual exposure. More research is needed to determine if these effects are additive, synergistic, or antagonistic to inform chemical risk assessments of BPA alternatives.

1.5 Using transcriptomics to study BPA alternative chemicals

Several studies have used HTTr to examine the transcriptional perturbations following *in vitro* exposure to BPA and BPA alternative chemicals. For example, one study conducted full transcriptome profiling using the Affymetrix microarray platform of estrogen-receptor positive MCF-7 cells exposed to BPA and 6 alternatives (0.1 nM – 100 μ M) for 48 hours (Mesnage *et al.*, 2017). This study highlighted that the BPA alternatives clustered alongside the positive control (17 β -estradiol), as well as enriching genes associated with cell cycle control, steroid hormones, and genes associated with breast cancer etiology (Mesnage *et al.*, 2017). Further, pathway analysis using Metacore revealed that BPS activated the most pathways relative to BPA. To compare with microarray data, the experiment was repeated using RNA-seq and exposure to BPA and BPAF that was largely in line with microarray data but identified 2 – 3 times more significantly altered genes (Mesnage *et al.*, 2017). Another study using H9 human embryonic stem cells exposed to a range of BPA, BPS, 4,4'-BPF, and 3,3',5,5'-tetrabromobisphenol A (TBBPA) concentrations (1 – 700 μ M) found that BPA, BPS, and 4,4'-BPF had overlapping median BMCs ($\sim$ 100 μ M), suggesting similar potencies, while that of TBBPA was much higher (Peshdary *et al.*, 2021). However, these data must be interpreted with caution as the absence of ERs in human embryonic stem cells underestimates the toxicity of BPA and like compounds, which act primarily through this nuclear hormone receptor.

Together, these data point toward similar toxicological potencies and overlapping mechanisms between BPA and its replacements *in vitro*.

1.5.1 Estrogen receptor alpha transcriptomic biomarker

A critical challenge in HTTr is making sense of the complex data sets in an efficient and objective way. A promising approach in this field is the use of transcriptomic biomarkers that are highly predictive of specific key events in adverse outcome pathways like endocrine disruption (Corton *et al.*, 2019). Toward this, a 46 gene biomarker associated with ER alpha (ER α) has been developed and validated in the ER α -positive MCF-7 cells (Ryan *et al.*, 2016). This biomarker predicts compounds that activate ER α through statistical comparisons of gene expression profiles of the 46 biomarker genes, with a pattern of 32 up-regulated and 14 down-regulated genes following ER activation. Such an approach can be used to identify the potential for the chemicals to interact with ER α , identify the lowest observed effect concentration (LOEC) and produce a composite BMC score for the biomarker to compare potencies for this specific molecular initiating event. This biomarker was used to validate the estrogenic activity of BPA and six alternatives in MCF-7 cells (Mesnage *et al.*, 2017).

1.6 Thesis Overview

1.6.1 Rationale

BPA is a plasticizer used in a variety of commercial products and industries. It is regularly detected in the general population worldwide. Due to its endocrine disrupting effects, BPA is regulated in many jurisdictions and banned in some products. Chemical alternatives to BPA are increasingly used, often as chemical mixtures, and are also detected in the general population in many countries. Unfortunately, most BPA alternative chemicals are data-poor, limiting regulatory risk assessment activities. Therefore, there is a need to screen BPA alternative chemicals to address knowledge gaps and identify potential hazards. **The goal of my thesis was to use HTTr to compare the toxicological potencies of BPA alternative chemicals and identify potential mechanisms of action *in vitro*.**

In Chapter 2, I screened BPA and 15 alternative chemicals in MCF-7 breast cancer cells for transcriptomic changes. Transcriptomic biomarker analysis was used to identify chemicals that activate the ER α and cellular stress response pathways. Global and ER α -specific transcriptomic points of departure (tPODs) were derived to rank chemicals for toxicological potency and compare their ability to activate the ER α , respectively. Gene expression data was compared to a propriety curated database to identify gene sets enriched by chemical exposure.

Chapter 3 tested whether mixtures of ER α active BPA alternative chemicals screened in Chapter 2 behaved additively *in vitro*. We used a similar experimental design

as Chapter 2 to screen seven mixtures of BPA alternatives and their individual components in MCF-7 breast cancer cells. Concentration response modelling was used to predict global and ER α -specific tPODs based on the potencies of individual BPA alternatives. Transcriptomic biomarkers and gene set enrichment analysis was applied to test for ER α activity and identify modes of action.

In Chapter 4, I sought to screen BPA alternative chemicals in system that shares more similarities to human tissue than cells of cancerous origin. To this end, we exposed human mammary epithelial cells (HMECs) from multiple donors to some of the most potent BPA alternative tested in Chapter 2 and 3 and did whole transcriptomic profiling to derive tPODs, perform transcriptomic biomarker and gene set enrichment analysis.

The two supplementary chapters further build on these data. Appendix A utilizes data collected from Chapter 2 to develop a novel transcriptomic biomarker to detect chemicals that activate the ER α *in vitro*. Appendix B also incorporates data from Chapter 2 to explore utility of tPODs from HTTr studies in support of regulatory assessment.

1.6.2 Chapter Objectives

The overall goal of my PhD thesis is to use *in vitro* models coupled to global transcriptomic profiling to compare the toxicological potencies and potential mechanisms of action of BPA alternatives individually and in mixtures to inform regulatory decision-making.

Chapter 2: *In vitro* transcriptomic analyses reveal pathway perturbations, estrogenic activities, and potencies of data-poor BPA alternative chemicals

Objective 1: Apply HTTr to explore mechanisms of action and derive points of departures for potency comparison of BPA and its alternative chemicals in breast cancer cells.

Objective 1a: Use tPODs to compare the toxicological potency of BPA alternative chemicals.

Objective 1b: Use Ingenuity Pathway Analysis to compare upstream regulators and canonical pathways activated by BPA alternative chemical exposure.

Importance: I show that most BPA alternative chemicals tested activate the ER α *in vitro*. BPAF was the most potent chemical tested overall, and activated ER α at the lowest concentration, followed by BPA. Most BPA alternatives had similar modes of action in breast cancer cells and enriched pathways associated with dysfunctional cellular proliferation and cancer-related pathways. This work demonstrates the utility of *in vitro* HTTr to screen data-poor BPA alternative chemicals for estrogen-like effects and compare their toxicological potency to a known EDC. Moreover, these findings also underscore the importance for additional research on BPAF.

Chapter 3: High-throughput transcriptomics analysis of equipotent and human relevant mixtures of BPA alternatives reveal additive effects *in vitro*

Objective 2: Apply HTTr to perform concentration-response modeling to mixtures of BPA alternative chemicals to determine if these chemicals exert additive effects *in vitro*.

Objective 2a: Use tPODs to compare potency of mixtures of BPA alternatives to their individual components.

Objective 2b: Use Ingenuity Pathway Analysis to compare upstream regulators and canonical pathways activated by exposure to mixtures of BPA alternative chemicals.

Importance: This data set supports that the ER α active BPA alternative chemicals behave additively in MCF-7 breast cancer cells. Mixtures of BPA alternatives enrich similar gene sets as their individual components, consistent with the model of additivity. Moreover, I produced a similar potency ranking as Chapter 2 and identified BPAF as the most potent chemical tested. These data demonstrate the replicability of tPODs across HTTr studies and supports the use of an additive model in risk assessments of mixtures of BPA alternative chemicals.

Chapter 4: Transcriptomic analysis of BPA alternative chemicals in human mammary epithelial cells

Objective 3: Apply HTTr to determine potency and mode of action of BPA and its alternatives in primary human mammary epithelial cells.

Objective 3a: Use tPODs to compare potency of BPA alternative chemicals previously screen in breast cancer cells.

Objective 3b: Use Ingenuity Pathway Analysis to compare upstream regulators and canonical pathways activated by BPA alternative chemical exposure.

Importance: These data are the first to derive tPODs from primary tissues from multiple human donors. I show that the BPA alternative chemicals tested have similar toxicological potencies in primary mammary cells and identifies BPAF as the most potent chemical tested, and that BPA as the least. Gene set enrichment revealed two largely opposing modes of action, with one group of chemicals inhibiting pathways associated with cancer-related outcomes, and the other activating the same pathways. This work highlights the differences in mode of action of BPA alternatives in mammary cells of primary and cancerous origin.

Supplementary Chapters

Appendix A: A 50-gene biomarker identifies estrogen receptor-modulating chemicals in a microarray compendium

Importance: Transcriptomic biomarkers are useful for identifying biological pathways that are perturbed by a condition like chemical exposure. This work validates a 50 gene biomarker developed from multiple HTTr data sets of MCF-7 cells exposed to endocrine disruptors. This biomarker can be used on HTTr data sets to detect chemicals that activate the ER α *in vitro*.

Appendix B: From vision toward best practices: Evaluating in vitro transcriptomic points of departure for application in risk assessment using a uniform workflow

Importance: Regulatory toxicology has traditionally relied on animal studies to inform decision making. This meta-analysis applies a uniform workflow to several HTTr data sets from multiple cell lines to compare toxicological potencies and derive human-relevant administered equivalent doses. It highlights that for many of the chemicals analyzed, *in vitro* potencies were equally or more protective than animal-derived potencies, supporting the use of *in vitro* HTTr data for regulatory toxicology assessment activities.

Chapter 2: *In vitro* transcriptomic analyses reveal pathway perturbations, estrogenic activities, and potencies of data-poor BPA alternative chemicals

Modified from: Matteo G, Leingartner K, Rowan-Carroll A, Meier M, Williams A, Beal MA, Gagné M, Farmahin R, Wickramasuriya S, Reardon AJF, Barton-Maclaren T, Christopher Corton J, Yauk CL, Atlas E. *In vitro* transcriptomic analyses reveal pathway perturbations, estrogenic activities, and potencies of data-poor BPA alternative chemicals. *Toxicological Sciences*, 2023 Feb 17;191(2): 266-275.

2.1 Abstract

Since initial regulatory action in 2010 in Canada, bisphenol A (BPA) has been progressively replaced by structurally related alternative chemicals. Unfortunately, many of these chemicals are data-poor, limiting toxicological risk assessment. We used high-throughput transcriptomics to evaluate potential hazards and compare potencies of BPA and 15 BPA alternative chemicals in cultured breast cancer cells. MCF-7 cells were exposed to BPA and 15 alternative chemicals (0.0005 – 100 μ M) for 48 hrs. TempO-Seq (BioSpyder Inc.) was used to examine global transcriptomic changes and estrogen

receptor alpha (ER α)-associated transcriptional changes. Benchmark concentration (BMC) analysis was conducted to identify two global transcriptomic points of departure (tPODs): (a) the lowest pathway median gene BMC and (b) the 25th lowest rank-ordered gene BMC. ER α activation was evaluated using a published transcriptomic biomarker and an ER α -specific tPOD was derived. Genes fitting BMC models were subjected to upstream regulator and canonical pathway analysis in Ingenuity Pathway Analysis. Biomarker analysis identified BPA and eight alternative chemicals as ER α active. Global and ER α tPODs produced highly similar potency rankings with bisphenol AF (BPAF) as the most potent chemical tested, followed by BPA and bisphenol C (BPC). Further, BPA and transcriptionally active alternative chemicals enriched similar gene sets associated with increased cell division and cancer-related processes. These data provide support for future read-across applications of transcriptomic profiling for risk assessment of data-poor chemicals and suggest that several BPA alternative chemicals may cause hazards at similar concentrations to BPA.

2.2 Introduction

Bisphenol A (BPA) is an endocrine disruptor that interacts with nuclear hormone receptors, as well as altering non-hormonal pathways, to exert its effects (Cimmino *et al.*, 2020). Exposure to BPA is associated with multi-organ toxicity and negative health outcomes including metabolic and endocrine dysfunction, reproductive and

developmental disorders, and hormone-related cancer (Ma *et al.*, 2019). Canada was the first country to take regulatory action to limit exposure to BPA in 2008, followed by the USA, the EU, and others (Rogers, 2021). As pressures to phase out BPA increase, manufacturers are relying more on BPA alternative chemicals.

BPA alternative chemicals are detected in the environment, consumer products, and in humans (Chen *et al.*, 2016). Many of these chemicals are structurally similar to BPA but differ in their adjoining atom and aryl substituents. A systematic review has highlighted the ability of bisphenol S (BPS), a sulfone-based bisphenol, and bisphenol F (4,4'-BPF), a BPA analogue missing the geminal methyl groups, to act as endocrine disruptors at similar concentrations to BPA (Rochester and Bolden, 2015). Some substitutes are less structurally related to BPA, like Pergafast201® and 4,4''-bis-(p-tolylsulfonyl ureido)-diphenylmethane (BTUM); these are nonphenolic sulphonyl urea-based alternative chemicals used primarily in thermal paper (Björnsdotter *et al.*, 2017). Several of these alternative chemicals are also associated with endocrine disruption, reproductive toxicity, and carcinogenicity (den Braver-Sewradj *et al.*, 2020). Thus, there is a clear need to identify potential hazards and determine relative potencies of BPA alternative chemicals.

High-throughput transcriptomics (HTTr) is increasingly used to identify hazards and establish bioactivity thresholds for regulatory toxicology (Thomas *et al.*, 2019). We recently applied HTTr to compare potencies of BPA alternative chemicals using benchmark concentration (BMC) analysis in H9 human embryonic stem cells exposed to

BPA, BPS, 4,4'-BPF, and 3,3',5,5'-tetrabromobisphenol A (TBBPA) at a range of concentrations (1 – 700 μ M) (Peshdary *et al.*, 2021). BPA, BPS, and 4,4'-BPF had overlapping median BMCs ($\sim$ 100 μ M), suggesting similar potencies, while the BMC of TBBPA was much higher. These data must be interpreted with caution as the absence of ERs in human embryonic stem cells underestimates the toxicity of BPA and similar compounds as the primary mode action of these chemicals is via these nuclear hormone receptors.

An important approach to screening BPA alternatives is to specifically compare potencies based on interactions with the estrogen receptor (ER). For example, Kitamura and colleagues tested the effects of several BPA alternatives using MCF-7 cells transfected with an estrogen response element (ERE) luciferase reporter (Kitamura *et al.*, 2005). Bisphenol AF (BPAF) and bisphenol B (BPB) stimulated estrogenic activity at lower concentrations than BPA, whereas BPS and 4,4'-BPF required higher concentrations. Mesnage *et al* 2017 found that BPAF, a hexafluoro bisphenol, was the most potent BPA alternative in stimulating an ERE-mediated luciferase reporter gene relative to BPB, bisphenol Z (BPZ), BPA, 4,4'-BPF, bisphenol AP (BPAP), and BPS (Mesnage *et al.*, 2017). The same study also conducted whole genome transcriptomic analysis with MCF-7 cells and observed that these BPA alternatives perturbed Gene Ontology (GO) gene sets associated with cell cycle, steroid hormone response, and breast cancer.

Nuclear receptor activation can also be assessed through the analysis of transcriptomic biomarkers. Toward this, a 46 gene transcriptomic biomarker associated with ER α was developed and validated in the ER α -positive MCF-7 cells (Ryan *et al.*, 2016). This biomarker predicts compounds that activate ER α through statistical comparisons of gene expression profiles of the 46 biomarker genes, with a pattern of 32 up-regulated and 14 down-regulated genes following ER activation. This approach can be used to identify the potential for the chemicals to interact with ER α and produce a composite BMC score for the biomarker to compare potencies for this specific molecular initiating event.

In this study, we explored the use of HTTr to establish mechanistic similarities, compare potencies, and increase understanding of data-poor BPA alternative chemical toxicity in support of prioritization and assessment activities by Health Canada. MCF-7 cells were exposed to BPA and 15 BPA alternatives (from 0.0005 – 100 μ M) or solvent (0.1% DMSO) for 48 hours. The exposure concentrations chosen span several orders of magnitude above and below the mean concentration of BPA in human urine (approximately 10 nM; Chen *et al.*, 2016) and are consistent with the US Environmental Protection Agency's ToxCast Program. HTTr using Templated Oligo-Sequencing (TempO-Seq; BioSpyder Inc.) was used to examine both general toxicological transcriptional effects as well as ER α -associated transcriptional changes induced by these chemicals. BMC analysis was conducted to identify both general and ER α -specific transcriptomic points of departure (tPODs). Finally, genes showing robust concentration-

response were subjected to upstream regulator and pathway analysis to explore the toxicodynamic similarity of these chemicals to each other and BPA.

2.3 Methods

2.3.1 Chemicals

See Table 2.1 for a list of chemicals and suppliers. The final concentration of DMSO in media was 0.1% for all chemicals tested.

2.3.2 Cell culture

MCF-7 cells (ATCC, Manassas, VA, USA) were routinely cultured in 100 mm dishes (Corning Falcon #353003, Corning, NY, USA) with Dulbecco's modified Eagle's medium (DMEM; Gibco, Thermo Fisher Scientific, Waltham, MA, USA) supplemented with 10% fetal bovine serum (FBS; Wisent Bioproducts, Saint-Jean Baptiste, QC, Canada), and 1% penicillin and streptomycin, at 37 °C in a humidified atmosphere with 5% CO₂. For the experiment, cells were seeded into clear 96-well plates (Corning Falcon #353075) at a density of 2.5×10^4 cells per well in phenol red-free DMEM supplemented with 5% charcoal-dextran stripped FBS (CD-FBS) to eliminate estrogenic components (Atlas *et al.*, 2003). The next day the medium was removed and replaced with DMEM with 5% CD-FBS containing the chemicals of interest. Cells were exposed, in quadruplicate,

to BPA and 15 alternative chemicals at 10 different concentrations (0.0005, 0.001, 0.01, 0.1, 0.5, 1, 5, 10, 50, 100 μ M), the positive control 17 β -estradiol (E2; 0.0001, 0.001, 0.01, 0.1, 1, 10 nM), dexamethasone (Dex – a non ER-interacting control; 0.0001, 0.001, 0.01, 0.1, 1 μ M) alongside solvent controls (0.1% DMSO; four solvent controls per plate) for 48 hours. The exposure timeframe was selected based on pilot studies to ensure that we observed a robust response in estrogen-responsive genes (data not shown). After exposure, cells were washed with phosphate-buffered saline (PBS) and lysed *in situ* using 2x TempO-Seq lysis buffer diluted with an equal amount of PBS for Templated Oligo-Sequencing (TempO-Seq®). We observed precipitate in some chemicals and removed these concentrations from further analysis: BPAF (50, 100 μ M), BPC (50, 100 μ M), BADGE (50, 100 μ M), D-8 (50, 100 μ M), DDS (10, 50, 100 μ M), BTUM (50, 100 μ M), TGSA (50, 100 μ M), BPS-MPE (50, 100 μ M).

2.3.3 Cell viability assay

To determine cytotoxic concentrations, MCF-7 cells were seeded into black 96-well plates (Corning Falcon #3603) and treated as described above. Cell viability was measured using the CellTiter-Blue Cell Viability Assay (Promega Corp, Madison, WI, USA), as per the manufacturer's instructions. This is a fluorometric assay that measures the metabolic capacity of cells based on resazurin reduction. Briefly, after 48 hour exposure, CellTiter-Blue reagent was added to each well, and the plates were incubated

at 37°C, 5% CO₂ for 2 hours. The fluorescence was read using excitation wavelength of 560nm and emission of 590nm using a SpectraMax M2 (Molecular Devices LLC, San Jose, CA, USA). Fluorescence readings were expressed as a % ratio to that of the DMSO control group. Cytotoxicity was defined as readings <50% of the control.

2.3.4 Cell proliferation assay

MCF-7 cells were seeded at a density of 2.0×10^4 cells/well in clear 24-well plates (Corning Costar #3524) and incubated at 37 °C and 5% CO₂ in phenol red free DMEM/F12 media containing 5% CD-FBS and 1% penicillin-streptomycin. Cells were treated with BPA and 15 alternative chemicals at 11 different concentrations (0.0001 - 100 uM) and positive control E2 (0.0001 - 10 nM). Chemical concentrations that yielded precipitates were again excluded from this analysis. Cells were counted on day seven using a TC20 automated cell counter (Bio-Rad, Hercules, CA, USA) as per manufacturer's instructions after being treated every two days with indicated concentration of the bisphenols and E2.

2.3.5 TempO-Seq library building and next-generation sequencing

Gene expression was measured using the TempO-Seq Human Whole Transcriptome v2.0 kit (BioSpyder Technologies Inc. Carlsbad, CA, USA) as per

manufacturer's instructions and previously described (Yeakley *et al.*, 2017). Cell lysates and positive technical controls (Human Universal Reference RNA - uhrRNA Agilent Cat# 740000, Santa Clara, CA, USA, and Human Brain Total RNA brRNA - ThermoFisher AM7962, Waltham, MA, USA), as well as no-cell negative controls (1X TempO-Seq lysis buffer alone) were hybridized to the Detector Oligo (DO) Pool in an annealing kit for the whole human genome supplied by BioSpyder for 10 minutes at 70 °C followed by a temperature gradient with a ramp rate of 0.5 °C/minute to 45 °C over 50 minutes with a 16-hour hold at 45 °C and then cooled to 25 °C. Nuclease digestion was employed to remove excess, unbound, or incorrectly bound DOs at 37 °C for 90 minutes. Amplification templates were generated by ligating DO pairs bound to adjacent target sequences for 1 hour at 37 °C, followed by enzyme denaturation for 15-minutes at 80 °C. Amplification templates (10 µL) were pipetted into a 96-well PCR Pre-Mix and Primer plate supplied by BioSpyder and amplified using a CFX96 Real-Time PCR Detection System (Bio-Rad) to incorporate a unique sample index/tag sequence and the sequencing adaptors for each sample. The following PCR settings were used: 37 °C for 10 minutes, 95 °C for 1 minute; 25 cycles of 95 °C for 10 seconds, 65 °C for 30 seconds, 68 °C for 30 seconds (with optical read for visual sample quality control (QC)); 68 °C for 2 minutes; and hold at 25 °C prior to library pooling and purification. For a list of attenuators used, see Supplementary information (Supplementary File 1.1).

NucleoSpin Gel and PCR Clean-up kits were used to pool and purify labeled amplicons. Next-generation sequencing (NGS) libraries were sequenced using NextSeq

500 High-Throughput Sequencing System (Illumina San Diego, CA, USA), using 50 cycles from a 75-cycle high-throughput flow cell to achieve a median read depth of 2 million reads per sample. Data were processed as described below and reads were aligned to the BioSpyder TempO-Seq Human Whole Transcriptome probe set (22,537 probes over 19,687 genes) using their purpose-built pipeline.

2.3.6 Data processing and quality control

Reads were demultiplexed from the BCL files and processed into FASTQ files using `bcl2fastq v2.20.0.422`. The FASTQ files were processed with the TempO-SeqR analysis pipeline in R (version 3.1) supplied by BioSpyder. This script uses STAR v2.7.8a (Dobin *et al.*, 2013) to perform alignment of raw reads to the reference sequence and the `qCount` function of the QuasR R package (v1.30.0; Gaidatzis *et al.*, 2015) to extract the feature counts from the aligned reads (BAM files) using features specified in a GTF file. The script generates a samples-by-probes count matrix.

Study-wide QC was performed on the count matrix using several methods to measure consistency and remove low-quality samples, using the methods in Harrill *et al.* (Harrill *et al.*, 2021) as a guideline. A cut-off of 0.1 for 1-Spearman's ρ was used to remove samples that were not correlated with others in the study. As in Harrill *et al.*, we also used a 10% cut-off of uniquely mapped reads as the number of target sequences (e.g., 100,000 reads to pass filter when the target is 1,000,000 for TempO-Seq experiments). We

removed any samples outside of Tukey's Outer Fence (3X interquartile range) for: 1) the number of probes capturing the top 80% of the signal, and 2) the number of detected probes (those with at least 5 mapped reads). Samples with a Gini coefficient (which measures inequality in distributions) greater than 0.95 were excluded. Samples removed by these criteria are listed in the Supplementary information (Supplementary File 1.2). We note that a technical error led to loss of the 0.001 and 0.01 μM concentrations of BPA-exposed cells (see Supplementary Table S2).

To create a matrix for biomarker analysis, individual pair-wise contrasts (pairwise comparisons) for each concentration and each chemical tested were created to the respective 0.1% DMSO control samples for each plate. Following the recommendations set out by the Omics Data Analysis Frameworks for Regulatory application (R-ODAF) guidelines (Verheijen *et al.*, 2022), genes were filtered for each contrast tested to include only those where 75% of at least one experimental group were above 0.5 Counts per Million (CPM), and spurious spikes were removed in which $[\text{max} - \text{median}]$ of counts were less than $[\text{sum of counts}]/[\text{number of replicates} + 1]$. We used DESeq2 1.30.0 (Love *et al.*, 2014) to estimate fold changes and normalize for library size within the TempO-Seq data. The ashR method was used to perform log2FoldChange shrinkage (Stephens, 2017). The code used to perform processing of high-throughput sequencing data is available at <https://github.com/R-ODAF>.

2.3.7 Gene expression biomarker analysis

To determine if the BPA alternatives activated ER or stress pathways, the expression profile of each chemical-concentration tested relative to solvent control (derived from the analysis of biosets) was compared to a number of characterized biomarkers as previously described (Kupersmidt *et al.*, 2010). Biosets are gene expression profiles of genes filtered by unadjusted p value < 0.05 and absolute fold change ≥ 1.2 . The biomarkers included those that predict modulation of ER (Ryan *et al.*, 2016), Nrf2 (Rooney *et al.*, 2020), heat shock factor 1 (Cervantes and Corton, 2021), metal-induced transcription factor 1 (Jackson *et al.*, 2020), TGx-DDI (Li *et al.*, 2019), TGx-HDACi (Cho *et al.*, 2021), and NF-kB (Korunes *et al.*, 2022). The correlations between each biomarker and the gene lists of BPA and alternatives were determined using the Running Fisher algorithm as described previously in the BaseSpace Correlation Engine (Ryan *et al.*, 2016). The Running Fisher algorithm provides an assessment of the statistical significance of the correlation of the overlapping genes between the biomarker and each gene list providing a summary p-value. A complete description of the Running Fisher test is provided in (Kupersmidt *et al.*, 2010). The results were exported and each p-value was converted to a $-\log(\text{p-value})$. Negative values were used to indicate negative correlation between the biomarker and the gene list. Thresholds for significance were set at $-\log(\text{p-value}) \geq 4$ for activation or ≤ -4 for suppression based on prior studies.

2.3.8 Points of departure analysis

We prefiltered probes using a Williams trend test ($p < 0.05$) and absolute fold change > 1.5 . Probes with the following criteria were removed: 1) having a BMC greater than the concentration used in the analysis after excluding cytotoxic/precipitating concentrations; 2) mapping to more than one gene; 3) having a model fit p -value < 0.1 determined by a likelihood ratio test; and 4) having a BMC upper (BMCU) to BMC lower (BMCL) ratio < 40 . BMDEExpress2.3 was used to derive BMCs for BPA alternative chemicals as previously described (Farmahin *et al.*, 2017). Concentration-response modeling was done by fitting genes best fit models, including polynomial 2°, linear, power (power term constrained to ≥ 1), and exponential models (degrees 2-5); best fit models were selected for each probe based on the lowest Akaike Information Criterion (AIC). The BMDEExpress parameters were: 250 maximum iterations, 0.95 confidence level, assume constant variance, BMR type standard deviation, BMR factor 1 standard deviation, restrict power ≥ 1 . The BMDEExpress model selection criteria were: compute and utilize in best model selection, best poly model test Nested Chi Square, p -value cut-off 0.05. Probes that met all the BMC filtering criteria were converted to their corresponding Entrez Identifiers. Data were further analyzed using R-4.0.5 statistical software. Gene accumulation plots were made by plotting the BMC of each gene on the horizontal axis and representing the gene accumulation number along the vertical axis. Two approaches were used for tPOD generation intended for: (1) general toxicity; and (2) ER α -specific response.

To establish general toxicity tPODs two methods were used:

- a) Lowest median pathway BMC: We produced the lowest median pathway BMC by aligning genes and their associated BMC/BMCL/BMCU values to the gene sets in Reactome Pathways database (<https://reactome.org/>), the Kyoto Encyclopedia of Genes and Genomes (KEGG; <https://www.genome.jp/kegg/>), or the Gene Ontology (GO) database (<http://geneontology.org>). Gene sets that contained at least three gene BMCs (genes that pass all criteria in the analysis) and were at least 5% populated (based on total annotated gene number) were selected. The tPOD chosen represents the gene set from either pathway database with the lowest median BMC. This approach was based on one recommended by the National Toxicology Program (NTP; National Toxicology Program, 2018).
- b) We set the 25th rank-ordered gene as a threshold to identify the concentration at which a robust change in the transcriptome has occurred (e.g., a concerted molecular response; Johnson et al., 2022). Chemicals that do not have at least 25 genes fitting BMCs were considered 'inactive' based on analysis of this tPOD.

To derive an ER α -specific tPOD, we calculated the median BMC of the 46 genes from the published ER α biomarker (Ryan *et al.*, 2016). Confidence Intervals for gene sets were estimated using a parametric bootstrap. For each gene in the gene set, a uniform

distribution was assumed with the lower limit represented by the gene BMCL and the upper limit represented by the gene BMCU. A bootstrapped sample consisted of randomly generated BMCs for each gene that had a BMC. The median value from each of the 2000 bootstrap samples was used to estimate the bootstrap distribution for the median. The 95% confidence interval was estimated using the 2.5th and 97.5th percentiles from the bootstrap distribution.

2.3.9 Pathway and upstream regulator enrichment analysis

Ingenuity Pathway Analysis (IPA; QIAGEN, Redwood City, CA, USA) was used to identify perturbed upstream regulators, canonical pathways, and disease & functions. For each chemical tested, we imported Excel files into IPA containing gene IDs (Ensembl and Gene Symbol) for the genes that fit a BMC (i.e., passed all filtering criteria), as well as their Williams trend test p-values from the BMC analysis, and the largest fold-changes of the gene relative to solvent controls (exported data from BMDExpress). This approach allowed us to identify enrichment of genes showing robust concentration-responses to the exposures. IPA Core Analysis with a gene expression threshold of absolute fold change ≥ 1.5 and FDR-adjusted p-value ≤ 0.05 was used with the direct and indirect relationship settings based on experimental and highly predicted data (focusing on human sources from breast cancer cell lines). Statistical significance of the overlap (FDR-adjusted p-value ≤ 0.05) between the data set and known targets of upstream regulators

in IPA were calculated using Fisher's Exact tests. The z-score was calculated using Fisher's Exact Test based on the expected relationship for directions between upstream regulators and target genes and those observed in the data set. A z-score of > 2 (activated) or < -2 (inhibited) was considered statistically significant.

2.4 Results

2.4.1 Cell viability assay

Cytotoxicity was evaluated using the CellTiter-Blue Cell Viability assay. There was no evidence of cytotoxicity using this assay, as there was no decline in fluorescence for BPA or any of the alternative chemicals at any of the concentrations tested following 48 hour exposure (Figure S1).

2.4.2 Cell proliferation assay

Cell proliferation was evaluated by counting cells following exposure to chemicals for seven days and comparing to respective solvent controls. BPA (50, 100 μ M) and several alternative chemicals decreased cell counts below 50% of solvent control, including: 4,4'-BPF (100 μ M), BPAP (10, 50, 100 μ M), BPS-MAE (50, 100 μ M), 2,4'-BPF (100 μ M) which suggests cell cytotoxicity (Figure S2).

2.4.3 General HTTr data summary

Transcriptomic data were processed using the R-ODAF pipeline that assesses quality of pre-processed data, implements study-wide alignment QCs (Harrill *et al.*, 2021), and produces fold changes and p-values for each gene with the DESeq2 R package (Love *et al.*, 2014). The median read depth was approximately 2 million per sample with a median range of 96.6% mapped reads. Of 960 samples, 59 were lost to quality assurance (QA)/QC analysis. All quality control reports are available in the Supplementary information (Supplementary File S2). The data were further filtered at the gene or probe level to produce a read count matrix for analysis in BMDExpress v2.3 to derive lists of concentration-responsive genes for enrichment analysis and tPOD derivation for toxicological potency ranking.

2.4.4 Stress response biomarker activation

We used previously published transcriptomic biomarkers to explore general stress responses and toxicities following exposure to BPA and its substitutes (Supplementary File 1.3). The biomarkers were specific to: (a) nuclear factor erythroid 2-related factor 2 (Nrf2) activation/suppression (Rooney *et al.*, 2020); (b) DNA damage induction (TGx-DDI; Li *et al.*, 2019); (c) metal responsive transcription factor 1 (MTF1) activation/suppression (Jackson *et al.*, 2020); (d) nuclear factor kappa beta (NFkB) activation/suppression (Korunes *et al.*, 2022) and a biomarker gene set (e) for activation of heat shock

transcription factor 1 (HSF1; Cervantes and Corton, 2021). We also investigated an epigenotoxicity biomarker for histone deacetylase inhibition (HDACi; Cho et al., 2021). Coactivation of these biomarkers within a condition generally occurs when there are broad, non-specific changes in transcription that result from overt cellular stress; i.e., the cytotoxic burst phenomenon that occurs at concentrations close to cell death (Escher *et al.*, 2020). Supplementary Table S1 provides a summary of all stress response biomarkers activated and inhibited.

BPAP activated most of the stress response biomarkers at the top two concentrations: Nrf2 (50, 100 μ M); TGx-DDI (100 μ M); MTF1 (50, 100 μ M); NF-kB (100 μ M); and HSF1 (50, 100 μ M). BPA activated both Nrf2 and MTF1 at 100 μ M. 2,4'-BPF activated Nrf2 (50, 100 μ M) and NF-kB (50 μ M). BPS activated NF-kB (100 μ M) and MTF (0.001 μ M). BPS-MAE activated Nrf2 and MTF1 at 100 μ M. P201 activated MTF1 (0.01, 0.1, 1, 10, 100 μ M) and NF-kB (100 μ M). A few other chemicals activated a single biomarker: 4,4'-BPF activated the Nrf2 biomarker at 100 μ M; 2,4'-BPS activated NF-kB at 100 μ M; and BADGE (0.1, 0.5, 1 μ M) and BTUM (5 μ M) activated MTF1. The HDACi biomarker was inhibited by BPA (5 - 50 μ M) and the following chemicals: 4,4'-BPF (5 – 100 μ M), BPAF (1 μ M), BPAP (5 – 50 μ M), BPC (5 – 10 μ M), and BPS (5 μ M).

2.4.5 Final sample inclusion and exclusion for ER α activity, potency comparison, and enrichment analysis

The stress response biomarker analysis was used to identify those concentrations of the chemicals that were causing overt cellular stress. Specifically, a chemical concentration was considered cytotoxic (overly stressed) if it activated two or more biomarkers or if it decreased cell counts below 50% of controls; these concentrations were subsequently removed from further analyses. In general, the highest exposure concentrations of BPA (50, 100 μ M) and several BPA alternative chemicals met these criteria, including: 2,4'-BPF (50, 100 μ M), 4,4'-BPF (100 μ M) BPAP (10, 50, 100 μ M), BPS-MAE (50, 100 μ M), and P201 (100 μ M). Supplementary Table S2 provides a summary of the final concentrations retained for: (a) ER α biomarker analysis; (b) BMC analysis; (c) tPOD derivation; (d) UR and pathway analysis. For some chemicals like BPA and BPAP, stress response biomarkers were activated at highest concentrations while ER α biomarker activation at these concentrations decreased. These data support that the cellular generalized stress response was associated with a loss of ER α signaling. Further, this stress response is in line with a cytotoxic 'burst' phenomenon associated with exposure to high concentrations of certain chemicals (Judson *et al.*, 2016).

2.4.6 Identification of ER α receptor agonists and potency ranking for ER α effects

The ability of each BPA alternative chemical tested to activate the ER α was predicted using a published 46 gene ER α biomarker (Ryan *et al.*, 2016). Nine chemicals

including BPA were predicted to be ER α agonists (Figure 2.1 and Supplementary File 1.3).

The median BMCs of genes that fit models using BMDEExpress from the ER α biomarker were also used to rank order the potency of these chemicals for receptor for ER α activity. The rank from most potent to least potent based on this was: BPAF, BPC, BPA, 4,4'-BPF, BPAP, BPS, 2,4'-BPF, P201, 2,4'-BPS (Figure 2.1).

2.4.7 Potency ranking using BMC analysis

The total number of genes fitting BMC models is shown in Figure 2.1. The top two most transcriptionally active by this analysis (i.e., chemicals with the most genes fitting BMC models) were BPS and 4,4'-BPF, with 1040 and 861 genes fitting models, respectively. In contrast, three BPA alternative chemicals (DDS, BTUM, BPS-MPE) had fewer than 25 genes fitting BMC models (presented in more detail below).

To rank order BPA alternative chemicals by toxicological potency, we calculated two general systemic tPODs to represent a 'tipping point' for induction of transcriptomic perturbations: (a) the 25th rank-ordered gene BMC, and (b) the lowest pathway gene set median BMC as per the US NTP's recommendation (National Toxicology Program, 2018b).

Gene accumulation plot shows the ranking of BPA and its alternatives potency according to their 25th ranked-ordered gene BMC (Figure 2.2). The order of potency from the most to least potent (lowest BMC to highest BMC) was: BPAF, BPA, BPC, 4,4'-BPF, BPAP, BPS, 2,4'-BPF, BADGE, P201, D-8, 2,4'-BPS, TGSA, BPS-MAE. Three BPA alternative chemicals (DDS, BTUM, BPS-MPE) did not have at least 25 genes with modelled BMC and thus were considered 'not transcriptionally active' at the concentrations tested (Figure 2.1). General system tPODs were used to compare potency to our prototype agent: BPA. The halogenated bisphenol BPAF was more potent than BPA (BMC = 0.00098 μ M and 0.0084 μ M for BPAF and BPA, respectively). BPA is more potent than BPC (BMC = 0.033 μ M) by one order of magnitude, followed by five alternative chemicals (4,4'-BPF, BPAP, BPS, 2,4'-BPF, BADGE) that had BMCs greater than one order magnitude higher (0.15 – 0.61 μ M) compared to BPC. Importantly, the 25th rank-ordered gene BMC of BPA is within one or order of magnitude of the mean human urinary BPA concentration (Chen *et al.*, 2016). The remaining five (i.e., least potent) transcriptionally active BPA alternative chemicals (P201, D-8, 2,4'-BPS, TGSA, BPS-MAE) had BMCs greater than the previous five compounds by one order of magnitude (2.9 – 7.4 μ M).

We then ranked BPA and its alternatives by potency based on the lowest median BMC derived from KEGG and Reactome pathways (Figure 2.1). The order of potency from most to least potent was: TGSA (GO), BPAF (GO), 2,4'-BPF (Reactome), 2,4'-BPS (KEGG), BPA (GO), BPC (GO), 4,4'-BPF (Reactome), BPAP (GO), BPS (GO), P201

(Reactome), D-8 (GO), BPS-MAE (GO). A full list of gene sets for each chemical tested is included in Supplementary materials (Supplementary File 1.4).

2.4.8 Pathway and Upstream Regulator Enrichment analysis

We uploaded the lists of genes with BMCs, along with the maximum fold change and Williams trend test p-values, into IPA to determine if there was enrichment of genes fitting models to specific canonical pathways and/or diseases, or associated with regulation by specific upstream molecules. This approach allowed us to compare chemicals across all concentrations tested and revealed that the positive control (E2), BPA, and 10 BPA alternative chemicals (BPAF, BPC, 4,4'-BPF, BPAP, BPS, BPS-MAE, 2,4'-BPF, 2,4'-BPS, BPC, P201, and D-8) perturbed gene sets associated with predicted upstream regulators and canonical pathways (Figure 2.3). A full list of upstream regulators, canonical pathways, and disease and functions is in the Supplementary information (Supplementary File 1.5). The gene set enrichment analyses were remarkably similar across the ER α active chemicals and E2, which indicates a high degree of concordance in the transcriptional alterations that they induce.

Upstream regulator analysis (Figure 2.3) provides additional support to the weight of evidence that many of the BPA alternative chemicals are ER α activators (see Figure 2.3) operating through similar mechanisms. As expected, the estrogen receptor 1 (ER α) was predicted to be activated by E2 and all the ER α agonists identified in the biomarker

analysis. Further, E2 and many of the BPA alternative chemicals were also predicted to activate the retinoic acid receptor alpha (RAR α). From this analysis, D-8 and BPS-MAE were predicted to activate ESR1 (i.e., ER α) and RAR α , albeit weakly, although these chemicals were not predicted to be ER α active using the biomarker. E3 ubiquitin ligase constitutive photomorphogenic 1 (COP1) and Paf1/RNA polymerase II complex component (CTR9) were predicted to be activated only by BPA, BPAF, 4,4'-BPF, BPAP, and BPS and ring finger protein (RNF181) activation was predicted for BPAF, 4,4'-BPF, BPAP, and BPS. The two upstream regulators most inhibited by E2 and many of the BPA alternative chemicals were: lysine demethylase 5B (KDM5B) and the tumor protein 53 (TP53). Overall, BPA, BPAF, 4,4'-BPF, BPAP, and BPS perturbed the most similar upstream regulators.

In general, the most affected canonical pathways were associated with sustained proliferative signaling and downregulation of growth suppressors (Figure 2.3). For example, E2 and most ER α active chemicals tested were predicted to activate the kinetochore metaphase signaling pathway and inhibit the cell cycle and DNA damage checkpoint control regulation pathway. Some perturbed pathways, including aryl hydrocarbon receptor and salvage pathways of pyrimidines ribonucleotides, likely represent a non-specific transcriptomic response to xenobiotic exposure. Full interpretation of altered canonical pathways is beyond this scope of this study but serves to group BPA alternative chemicals based on possible mode of action. In sum, most ER α

active BPA alternative chemicals, including BPA, BPAF, 4,4'-BPF, BPAP, BPS, and 2,4'-BPF perturbed similar canonical pathways.

2.5 Discussion

A major goal of this study was to provide information on the mode of action and potency of data-poor BPA alternatives relative to more well-characterized replacements. We used a previously published transcriptomic ER α biomarker (Ryan *et al.*, 2016) to identify ER α active compounds and derive a tPOD for potency comparison. The biomarker analysis identified nine agonists among the 16 chemicals tested, in line with previous data. Indeed, several BPA alternative chemicals have been identified as ER α active using other methods, including BPC, 4,4'-BPF, BPS, BPAF, and BPAP (see Pelch *et al.*, 2019; Keminer *et al.*, 2020). BPAF had the lowest ER α median BMC, followed by BPC. BPA was the third most estrogenic compound tested, and had comparable estrogenicity to 4,4'-BPF, BPAP, BPS, and 2,4'-BPF. Importantly, our data align with Mesnage *et al.* who also applied this biomarker in MCF-7 cells to identify the following BPA replacements with ER α activity: BPAF, 4,4'-BPF, BPS, and BPAP (Mesnage *et al.*, 2017). Moreover, BPAF was identified as the most potent chemical tested among 20 BPA alternatives using ToxCast data for gene sets associated with nuclear hormone receptors, including the ER α (Bone *et al.*, 2016). Our data add to the literature base to support that BPAF is more estrogenic than BPA and other alternative chemicals *in vitro* indicating that

more information is required to scrutinize BPAF as an alternative to BPA. This information could serve to support read-across chemical safety evaluations.

Several data-poor BPA replacements demonstrated weak or no estrogenicity based on our biomarker analysis. Specifically, 2,4'-BPF was less estrogenic than its isomer (4,4'-BPF) by one order of magnitude. To our knowledge, we are the first group to report on the estrogenicity of 2,4'-BPF in MCF-7 cells. Previous work suggests that P201 is non-estrogenic (US Environmental Protection Agency (USEPA), 2014; Keminer *et al.*, 2020; Goldinger *et al.*, 2015); our results suggest that this alternative chemical is weakly estrogenic and ER α agonism occurs at high concentrations ($\geq 50 \mu\text{M}$). 2,4'-BPS was previously predicted to be ER α inactive (Pelch *et al.*, 2019; Keminer *et al.*, 2020); herein, we predicted ER α activity at high concentrations ($\geq 50 \mu\text{M}$). Other data-poor BPA alternatives, including TGSA, D-8, BPS-MAE, and BPS-MPE were predicted to be ER α inactive, in line with previous data (Pelch *et al.*, 2019; Keminer *et al.*, 2020; Kuruto-Niwa *et al.*, 2005; Goldinger *et al.*, 2015). We were unable to find other published reports on the estrogenicity of DDS or BTUM. Our data indicate that further investigations using established methods (e.g., competitive binding assays) are warranted to confirm the estrogenic activity of 2,4'-BPF and P201.

We explored two approaches to identify a 'general toxicological response tPOD' that reflects a concentration at which the transcriptome is perturbed based on (a) the 25th ranked gene BMC or (b) the lowest median pathway BMC. We previously used the 5th

percentile as a tPOD for deriving bioactivity exposure ratios to be conservative (Rowan-Carroll *et al.*, 2021); we note that percentile gene BMCs are influenced by the top concentration and total number of responsive genes. The 5th percentile can also be derived from very few genes and we sought to identify a method that ensures some minimal level of transcriptional change. Thus, the 25th rank-ordered gene was used for the first time as a fixed point that represents a robust transcriptomic changes relative to matched solvent controls (Reardon *et al.*, 2021); the 25th gene is a more stable point for direct comparisons across chemicals that is less influenced by inclusion/exclusion of top concentrations relative to the 5th percentile BMC. This was an arbitrary but conservative selection based on our previous analysis that such an approach provides reproducible potency rankings for per- and poly-fluorolkyated chemicals and representing approximately 0.1% of the genes in the genome. The lowest median pathway was included because it is also less influenced by inclusion/exclusion of top concentrations and has previously been proposed as an acceptable approach by a panel of experts (National Toxicology Program, 2018).

The general toxicological response tPODs yielded potency rankings that largely aligned with the ER α biomarker approach. In particular, the 25th gene BMC yielded very similar potency rankings as the ER α BMC, but was generally more conservative by one order of magnitude. This suggests that there is biological activity impacting genes not directly regulated by ER α at lower exposure concentrations than those regulated by the receptor. BPAF had the lowest 25th gene BMC, followed by BPA, BPC; all other chemicals

had 25th gene BMCs within two orders of magnitude of BPA. The lowest pathway median BMC was also very similar to the 25th gene BMC in most cases. Examination of the accumulation plots indicates the 25th gene BMC serves as a suitable representation of the concentration at which transcriptional activity ramps up. Interestingly, TGSA, 2,4'-BPF, and 2,4'-BPS had very low pathway median BMCs, while they were among the least potent chemicals tested according to the other two tPODs. This may represent a limitation to the NTP approach, which is not encountered using the 25th gene BMC. Further investigation is warranted into the transcriptomic effects of data-poor BPA alternative chemicals. Overall, these data align with previous data from our group that found considerable overlap between pathway- and global-based tPODs from microarray data of *in vivo* toxicological dose-response studies (Farmahin *et al.*, 2017).

We used a suite of previously published transcriptomic biomarkers to explore stress responses induced by the chemicals in the MCF-7 cells (Jackson *et al.*, 2020; Korunes *et al.*, 2022). Most cellular stress response biomarkers, as well as the DNA damage biomarker, were activated by BPA and alternative chemicals at the highest (≥ 50 μ M) concentrations tested. We also used a novel HSF1 biomarker to assess cellular stress and observed a similar trend. Interestingly, BPA and several alternatives inhibited the HDACi biomarker at non-cytotoxic concentrations. Thus, these BPA alternatives may decrease histone acetylation, perhaps indirectly, adding to a growing body of evidence that supports that *in vitro* and *in vivo* BPA exposure affects this pathway (Qin *et al.*, 2021).

Overall, those alternative chemicals with the most structural homology to BPA had the most similar patterns of biomarker perturbation.

Genes fitting BMCs revealed toxicodynamic similarities across the chemical sets supporting read-across for data-poor BPA alternative chemicals. All chemicals identified as ER α active by biomarker analysis were also predicted to activate ER α as an upstream regulator, along with RAR α . RAR α is necessary for ERE-mediated transcription and proliferation in breast cancer cells (Ross-Innes *et al.*, 2010) and has significantly overlapping target genes with ER α (Hua *et al.*, 2009). Most ER α active BPA alternative chemicals were predicted to inhibit KDM5B. Previous work supports a role for KDM5B in cultured breast cancer cell proliferation (Zhang *et al.*, 2019), and its predicted inhibition in our dataset may represent a compensatory mechanism in response to sustained proliferative signaling by BPA alternative chemicals. Furthermore, a subset of the ER α active chemicals activated upstream regulators that facilitate ER α signaling including CTR9 (Zeng *et al.*, 2016), RNF31 (Zhu *et al.*, 2014), RNF181 (Zhu *et al.*, 2020), and COP1 (Tang *et al.*, 2022). These data support our identification of ER α active BPA alternatives and suggest that they cause robust activation of the ER α proliferative network in MCF-7 cells.

Our pathway analysis suggests that several BPA alternatives induce transcriptional changes in gene sets associated with non-specific proliferative signaling and dysregulated cell cycle control. MAPK1 and CREB1 were predicted as activated upstream

regulators by a subset of chemicals tested. Another group identified CREB1 as an upstream regulator in hippocampi of BPA-exposed mice using RNA-seq (Henriksen *et al.*, 2020). Further, the most perturbed canonical pathways in this study by BPA and alternative chemicals were the kinetochore metaphase signaling pathway, mitotic roles of polo-like kinase, and G2/M DNA damage checkpoint. Importantly, a previous transcriptomic profiling study using RNA-seq in MCF-7 cells found that BPA and 5 alternative chemicals induced GO pathways associated with cell cycle and breast cancer (Mesnage *et al.*, 2017). Collectively, these enriched gene set data suggest that exposure to BPA and several alternative chemicals affect the ability of MCF-7 cells to regulate cell division.

In conclusion, our HTTr data analysis pipeline enabled us to identify potential ER α agonists within our 15 BPA alternative chemicals, several of which are supported by additional evidence in the literature. Application of BMC modeling facilitated toxicological potency ranking based both on ER-specific hazard as well as more generalized transcriptomic responses. A novel approach to gene set enrichment based on genes fitting BMC models revealed a high degree of similarity across the gene expression profiles induced by these chemicals, supporting their grouping for this and future analyses. This study also supports previous data indicating that BPAF is a potent ER α agonist that induces transcriptional changes at low concentrations, indicating that further research of halogenated BPA alternative chemicals is warranted. Given that these data are derived from an *in vitro* cell model that overexpresses the ER α , they are limited in

their ability to generalize to human toxicity. Thus, further investigation in primary cell models would be beneficial. Together, our study provides key support for future read-across applications and suggests that several of the proposed BPA alternative chemicals may cause hazards at similar concentrations to BPA.

2.6 Tables and Figures

Table 2.1 List of chemicals and suppliers.

Chemical Name	Abbreviation	CAS	Purity	Supplier
2,2-Bis(4-hydroxyphenyl)propane	BPA	80-05-7	> 99%	Sigma-Aldrich Inc (St. Louis, MO, USA)
Bis(4-hydroxyphenyl)methane	4,4'-BPF	620-92-8	> 98%	
1,1-Bis(4-hydroxyphenyl)-1-phenylethane	BPAP	1571-75-1	99%	
4,4'-Sulfonyldiphenol	BPS	80-09-1	98%	
2,2-Bis[4-(glycidyloxy)phenyl]propane	BADGE	1675-54-3	98%	
4,4'-Dichlorodiphenyl sulfone	DDS	80-07-9	98%	
4,4'-(Hexafluoroisopropylidene)diphenol	BPAF	1478-61-1	97%	
Dimethyl sulfoxide	DMSO	67-68-5	> 99%	
17 β -estradiol	E2	50-28-2	> 98%	
Dexamethasone	Dex	50-02-2	98%	Toronto Research Chemicals (Toronto, ON, Canada)
4-[[4-(2-Propen-1-yloxy)phenyl]sulfonyl]phenol	BPS-MAE	97042-18-7	97%	
4,4''-Bis-(p-tolylsulfonyl ureido)-diphenylmethane	BTUM	151882-81-4	98%	
4-((4-Isopropoxyphenyl)sulfonyl)phenol	D-8	95235-30-6	98%	
4-Methyl-N-[[[3-[[4-methylphenyl)sulfonyl]oxy]phenyl]amino]carbonyl]benzenesulfonamide	Pergafast201 (P201)	232938-43-1	98%	

2,4'-Dihydroxydiphenyl sulfone	2,4'-BPS	5397-34-2	98%	
4-(4-hydroxy-3-prop-2-enylphenyl)sulfonyl-2-prop-2-enylphenol	TGSA	41481-66-7	98%	
2,2-Bis(4-hydroxy-3-methylphenyl)propane	BPC	79-97-0	98%	Accustandard (New Haven, CT, USA)
2,4'-Dihydroxydiphenylmethane	2,4'-BPF	2467-03-0	> 98%	TCI America (Portland, OR, USA)

Chemical	ER α tPOD			general toxicity tPODs		
	ER α BM Genes with BMCs	ER α Median BMC	ER α Prediction	Genes with BMCs	25 th Gene BMC	Lowest Pathway BMC
BPAF	21	0.037	Agonist	609	0.00098	0.00074
BPA	12	0.73	Agonist	311	0.0084	0.0037
BPC	23	0.43	Agonist	620	0.033	0.047
4,4'-BPF	24	0.73	Agonist	861	0.15	0.13
BPAP	13	2.4	Agonist	573	0.16	0.13
BPS	25	2.7	Agonist	1040	0.50	0.36
2,4'-BPF	14	4.8	Agonist	456	0.60	0.0012
BADGE			Inactive	34	0.61	
P201	12	23	Agonist	416	2.9	2.6
D-8			Inactive	322	4.1	3.2
2,4'-BPS	11	56	Agonist	469	4.8	0.0012
TGSA			Inactive	39	7.0	0.00041
BPS-MAE			Inactive	44	7.4	5.1
DDS			Inactive	22		
BTUM			Inactive	15		
BPS-MPE			Inactive	13		

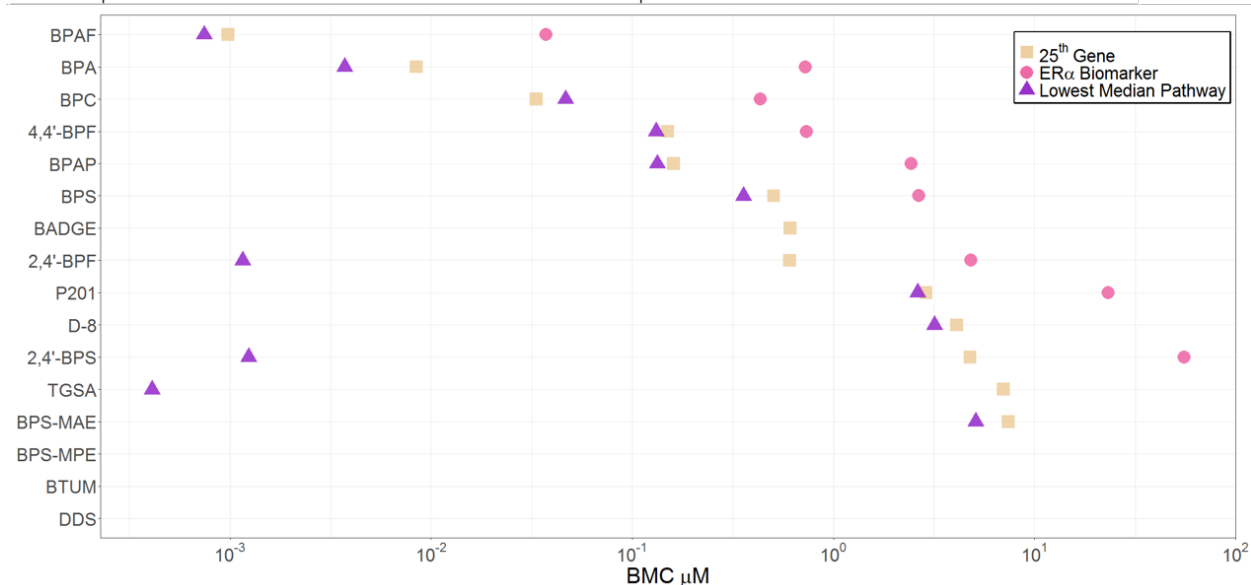


Figure 2.1 Comparison of ER α bioactivity and tPODs derived from the whole transcriptome analysis in MCF-7 cells ($n = 3 - 4$ per concentration) exposed to BPA and 15 alternative chemicals at a range of concentrations (0.0005 – 100 μ M) for 48 hours. ER α prediction is based on at least one concentration yielding an agonist or antagonist call. For tPOD derivation, data were prefiltered using the Williams trend test ($p \leq 0.05$) and a fold-change of ≥ 1.5 or ≤ 1.5 . Data were also post-filtered with the following settings in BMDExpress v2.3: Best BMD/BMDL < 20, Best BMDU/BMD < 20, Best BMDU/BMDL < 40, and Best fitPvalue > 0.1. tPODs representing the 25th rank ordered gene BMC (shown in μ M), the median gene BMC for the lowest pathway (at least 3 genes and 5% of pathway) as well as the median gene BMC for the ER α biomarker gene set are shown in the table and in the top panel. The chemicals are shown in decreasing order of potency based on tPODs from the 25th gene BMC.

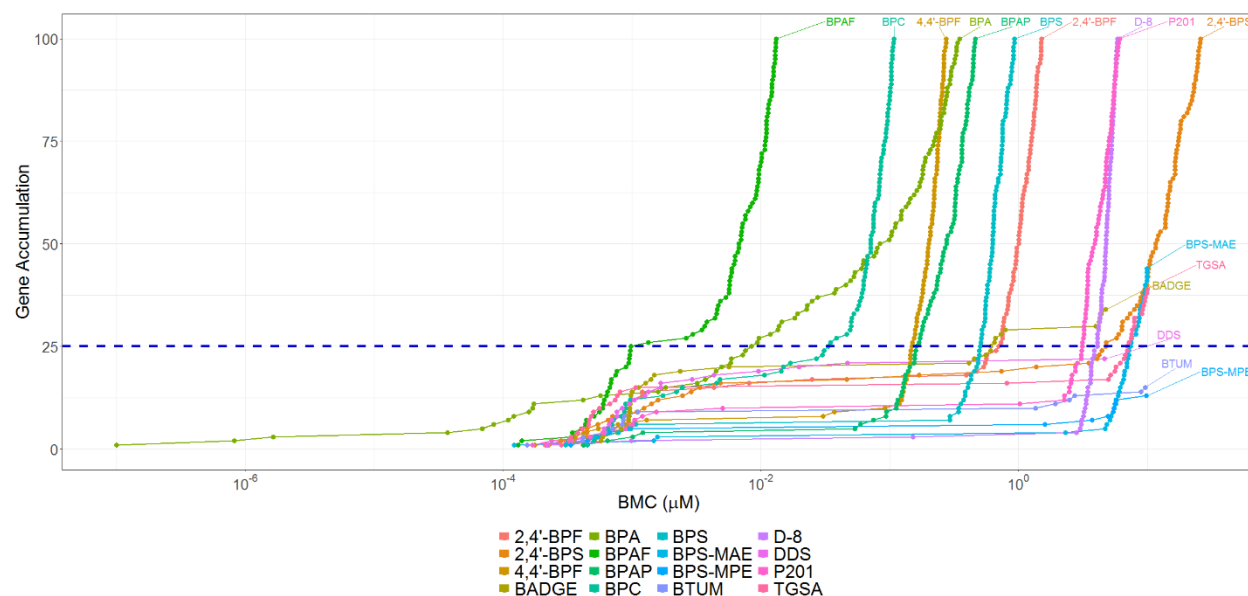


Figure 2.2 Gene accumulation plot of MCF-7 cells ($n = 3 - 4$ per concentration) exposed to BPA and 15 alternative chemicals ($0.0005 - 100 \mu\text{M}$) for 48 hours. The graph was modified to highlight the 25th ranked gene (blue dashed line). Note that the text labels are descriptive and rank order potency is based on the order of chemicals crossing the 25th gene line. Data were prefiltered using the Williams trend test ($p < 0.05$) and an absolute fold-change filter of ≥ 1.5 , and post filtered with the following settings in BMDEExpress v2.3: Best BMD/BMDL < 20 , Best BMDU/BMD < 20 , Best BMDU/BMDL < 40 , and Best fitPvalue > 0.1 .

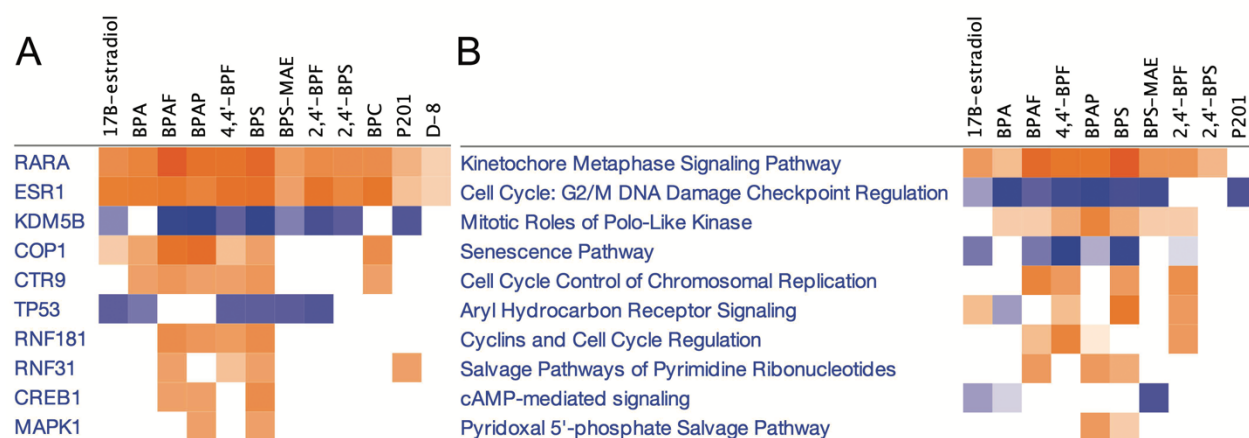


Figure 2.3 Ingenuity Pathway Analysis of genes fitting BMC models in the MCF-7 cells ($n = 3 - 4$ per concentration) exposed to BPA and 15 alternative chemicals for 48 hours, as well as positive control 17 β -estradiol. A list of genes fitting models, the Williams trend test p-value and the maximum fold change were imported into IPA for this comparison. Z-score and adjusted p-value filters were set to ≥ 2.0 and ≤ 0.05 , respectively. A: Top 10 Upstream Regulators. B: Top 10 Canonical Pathways. Orange is predicted activation and blue is predicted inactivation.

Chapter 3: High-throughput transcriptomics analysis of equipotent and human relevant mixtures of BPA alternatives reveal additive effects *in vitro*

Authors: Matteo G, Cho E, Rigden E, Eickmeyer DC, Bradford LM, Meier MJ, Williams A, Corton JC, Yauk CL, Atlas E

Modified From: Matteo G, Cho E, Rigden M, Eickmeyer DC, Bradford LM, Meier MJ, Williams A, Corton JC, Yauk CL, Atlas E. High-throughput transcriptomics analysis of equipotent and human relevant mixtures of BPA alternatives reveal additive effects *in vitro*. Arch Toxicol. 2025 Sep;99(9):3707-3719.

3.1 Abstract

While many jurisdictions have phased out use of bisphenol A (BPA), there is increasing exposure to mixtures of BPA alternatives. Like BPA, some alternatives perturb nuclear hormone receptors and are endocrine disruptors. We used high-throughput transcriptomics (HTTr) to evaluate the potency and modes of action of seven mixtures of BPA alternatives and their 12 individual components in breast cancer cells. Our aim was to explore whether alternatives present in mixtures act additively. MCF-7 cells were exposed to chemicals (0.001 – 50 μ M) for 48 h and gene expression analysis was used

to measure global and estrogen receptor alpha (ER α)-specific transcriptomic changes. Transcriptomic points of departure (tPODs) were derived using benchmark concentration (BMC) modelling. We first identified concentrations at which global transcriptional activity was robustly altered. Then, we applied a ER α transcriptomic biomarker to identify ER α agonists and predict ER α activation tPODs. We employed mixtures modelling to predict potency of BPA alternatives and test for additive effects *in vitro*. Ingenuity Pathway Analysis (IPA; Qiagen) was used to identify upstream regulators and canonical pathways from genes fitting BMCs. BPAF was the most potent individual chemical tested overall, followed by BPA and BPC. All seven mixtures had additive effects across all tPODs modelled. The ER α transcriptomic biomarker classified all mixtures as ER α activators along with several BPA alternatives. All mixtures and most individual components perturbed similar upstream regulators and pathways, suggesting common modes of action. These data support the value of HTTr in identifying additive effects and toxicological potency of mixtures *in vitro*.

3.2 Introduction

Bisphenol A (BPA) is an endocrine disruptor that affects nuclear hormone receptors including the estrogen and androgen receptors. BPA is primarily used in the manufacture of polycarbonate plastics, epoxy resins, and thermal paper; today, most people have measurable levels of BPA and its metabolites (Colorado-Yohar *et al.*, 2021). Given empirical evidence demonstrating various adverse health effects, BPA has been

increasingly replaced with alternative chemicals (Manzoor *et al.*, 2022). Epidemiological data support that humans are exposed to multiple BPA alternatives (Moreno-Gómez-Toledano *et al.*, 2023; S. Kim *et al.*, 2021; Gys *et al.*, 2020). Exposure primarily occurs orally as BPA and alternatives present as polymers degrade into monomers that can be ingested (Manzoor *et al.*, 2022). Another common route of exposure is by contact with thermal paper (e.g., receipt paper), as BPA and alternatives like Pergafast 201 (P201) are used as developing agents (Björnsdotter *et al.*, 2017). Despite their widespread use, many BPA alternative chemicals are data-poor, and there is mounting evidence that some may pose human health hazards.

Many BPA alternatives share common modes of action through ER α . For example, bisphenol AF (BPAF), bisphenol C (BPC), 4,4'-bisphenol F (4,4'-BPF), bisphenol B (BPB), bisphenol Z (BPZ), bisphenol AP (BPAP), and 4,4'-bisphenol S (BPS) are ER α agonists *in vitro* (Pelch *et al.*, 2019; Keminer *et al.*, 2020). Some, like BPAF, activate ER α at lower concentrations than BPA (Matteo *et al.*, 2023). Importantly, evidence suggests that BPA alternatives have additive effects when present as mixtures *in vitro*. For example, a binary mixture containing BPA and BPAF behaved additively in ER α luciferase assay in T47D-KBluc cells (Bermudez *et al.*, 2010). In HeLa cells expressing ER α , a ternary mixture of BPA, 4,4'-BPF, and BPS activated ER α at a concentration (0.43 μ M) lower than predicted according to a logistic model (Park *et al.*, 2020). Moreover, in a follow-up study using the

same *in vitro* system, the authors found that binary and ternary mixtures of BPA, BPAF, BPB, BPC, BPE, and BPZ activated ER α in an additive manner (Lee *et al.*, 2023). More research investigating the behaviour of BPA alternatives and their interaction with ER α in mixtures is needed to better understand these effects.

High-throughput transcriptomics (HTTr) is increasingly being used for chemical potency assessment, mode of action analysis, and biological read-across (Meier *et al.*, 2024). We previously used HTTr to show that several BPA alternatives have comparable potencies to BPA in MCF-7 breast cancer cells and activate ER α at similar concentrations (Matteo *et al.*, 2023). Beal and colleagues used a similar approach to screen 11 BPA alternatives also in MCF-7 cells (Beal *et al.*, 2024). In MCF-10 breast epithelial cells, exposure to equimolar mixtures of BPA alternatives activated KEGG pathways associated with metabolism, ER protein processing, spliceosome function, and ubiquitin-mediated proteolysis (Mesnage *et al.*, 2022). However, these mixtures were not assessed for additivity, synergy, or antagonism. These studies demonstrate the utility of HTTr in informing the toxicity of BPA alternatives for regulatory decision making (Reardon *et al.*, 2023). Applying these methods to clarify how these chemicals interact in mixtures to alter endogenous gene expression would be highly valuable to advance knowledge.

Here, we used HTTr to determine if simple and complex mixtures of BPA alternatives behave additively *in vitro*. MCF-7 breast cancer cells were exposed to seven different mixtures of BPA alternative chemicals across a range of eight concentrations (0.001 – 50 μ M), along with their 12 individual components, solvent control (0.1% DMSO),

and a positive control (17 β -estradiol, 1 and 10 nM) for 48 h in serum-free conditions. Templated Oligonucleotide Sequencing (TempO-Seq; BioSpyder Inc.) was used to quantify changes to the transcriptome and ER α -related gene sets. Transcriptomic biomarkers were used to assess if chemicals were ER α agonists and identify cellular stress conditions. Cytotoxicity of mixtures was assessed using a cell viability assay. BMC modeling was applied to gene expression data to derive transcriptomic points of departure (tPODs) that assess global, pathway, and ER α -specific transcriptomic changes. To test for additive effects, we applied a mathematical model previously used by our lab (Addicks *et al.*, 2023) to predict the potency of BPA alternatives based on their individual components. Finally, genes fitting BMCs underwent pathway over-representation analysis to examine the toxicological mechanisms associated with the individual chemicals and their mixtures.

3.3 Materials and methods

3.3.1 Chemicals and mixtures

Table 3.1 lists the chemicals used and their respective suppliers. All chemicals and mixtures (Figure 3.1, Table 3.1) were dissolved in dimethyl sulfoxide (DMSO); the final DMSO concentration in exposure media was 0.1%. BPA and alternative chemicals were combined to form 6 different equimolar mixtures (Mix 1 – Mix 6) and one complex (non-equimolar) mixture (Mix 7). Each mixture concentration (i.e., total chemical concentration)

is the sum of the individual chemicals (e.g., 50 μ M Mix 1 contains 25 μ M BPA and 25 μ M BPAF). The complex mixture (Mix 7) was made using the proportions of alternative chemicals detected in human urine (Gys *et al.*, 2020). The final concentrations used in Mix 7 are the sum of the individual concentrations of each alternative chemical. Table 3.2 describes each mixture's components and relative proportions of chemicals.

3.3.2 Cell culture

MCF-7 cells (ATCC, Manassas, Virginia) were cultured in 10 cm dishes (Corning Falcon No. 353003, Corning, New York) with Dulbecco's modified Eagle's medium (DMEM; Gibco, Thermo Fisher Scientific, Waltham, Massachusetts) containing 10% fetal bovine serum (FBS; Wisent Bioproducts, Saint-Jean Baptiste, QC, Canada), and 1% penicillin and streptomycin (P/S; Wisent), at 37°C in a humidified atmosphere (5% CO₂). For the experiment, cells were washed with phosphate-buffered saline (PBS) and phenol red-free DMEM supplemented with 5% charcoal-dextran stripped FBS (CD-FBS) was used to eliminate estrogens present in commercial medium (Atlas *et al.*, 2003) for 48 h. Cells were then seeded into clear 96-well plates (Corning Falcon No. 353075) at a density of 2.5×10^4 cells per well in phenol red-free DMEM supplemented with 5% CD-FBS. After 48 h, the medium was removed and replaced with phenol red-free DMEM containing 2% bovine serum albumin (BSA; supplier), 1% P/S, and chemicals of interest. BSA was used in place of CD-FBS in the exposure medium to further reduce the concentrations of estrogens and other hormones. Cells were exposed for 48 h, in quadruplicate, to BPA,

the 12 alternative chemicals, the six equimolar mixtures of alternative chemicals at eight different concentrations (0.001, 0.01, 0.1, 0.5, 1, 5, 10, 50 μ M), and one complex mixture. Concentrations for the complex mixture were 0.66, 1.3, 2.6, 5.2, 10.5, 21 μ M. The experiment included cells exposed to 17 β -estradiol (E2; 0.1, 1 nM) as a positive control and at least four plate-matched solvent controls (0.1% DMSO). The exposure timeframe was selected based on our previous study (Matteo *et al.*, 2023). After exposure, cells were washed with PBS and lysed *in situ* using 2x TempO-Seq lysis buffer diluted with an equal amount of PBS for TempO-Seq library building.

3.3.3 Cell viability assay

To determine if chemical exposure decreased cell viability, MCF-7 cells were seeded into black 96-well plates (Corning Falcon No. 3603) and treated as described above. We tested all mixtures and a subset of chemicals (BPB and BPZ) that were not included in our previous study (Matteo *et al.*, 2023). Cell viability was measured using the CellTiter-Blue Cell Viability Assay (Promega Corp, Madison, Wisconsin), as per the manufacturer's instructions. This is a fluorometric assay that measures the metabolic capacity of cells based on resazurin reduction. Briefly, after 48-h exposure, CellTiter-Blue reagent was added to each well, and the plates were incubated (37°C, 5% CO₂) for 2 h. The fluorescence was read using an excitation wavelength of 560 nm and emission of 590 nm using a SpectraMax M2 (Molecular Devices LLC, San Jose, California).

Fluorescence readings were expressed as a percent ratio to that of the DMSO control group. Cytotoxicity was defined as readings <50% of the control.

3.3.4 TempO-Seq library building and next-generation sequencing

Gene expression was measured using the TempO-Seq Human Whole Transcriptome v2.1 kit (BioSpyder Technologies Inc, Carlsbad, California) as per the manufacturer's instructions and as previously described (Matteo et al., 2023). Cell lysates and positive technical controls (Human Universal Reference RNA—uhrRNA Agilent Cat No. 740000, Santa Clara, California, and Human Brain Total RNA brRNA—ThermoFisher AM7962, Waltham, Massachusetts), as well as no-cell negative controls (1x TempO-Seq lysis buffer) were hybridized to the Detector Oligo (DO) Pool using an annealing kit for the whole human genome supplied by BioSpyder. The hybridization mixture was incubated for 10 min at 70°C followed by a temperature gradient with a ramp rate of 0.5°C /min to 45°C over 50 min with a 16-h hold at 45°C and then cooled to 25°C. Nuclease digestion was employed to remove excess, unbound, or incorrectly bound DOs at 37°C for 90 min. Amplification templates were generated by ligating DO pairs bound to adjacent target sequences for 1 h at 37°C, followed by enzyme denaturation for 15 min at 80°C. Amplification templates (10 µL) were pipetted into a 96-well PCR Pre-Mix and Primer plate supplied by BioSpyder and amplified using a CFX96 Real-Time PCR Detection System (Bio-Rad) to incorporate a unique sample index/ tag sequence and the sequencing adaptors for each sample. The following PCR settings were used: 37°C for

10 min, 95°C for 1 min; 25 cycles of 95°C for 10s, 65°C for 30s, 68°C for 30s (with optical read for visual sample quality control [QC]); 68°C for 2 min; and hold at 25°C prior to library pooling and purification.

NucleoSpin Gel and PCR Clean-up kits were used to pool and purify labeled amplicons. Next-generation sequencing libraries were sequenced using a NextSeq 2000 High-Throughput Sequencing System (Illumina, San Diego, California), using 50 cycles from a 75-cycle high-throughput flow cell to achieve a median read depth of 2 million reads per sample. Data were processed as described below and reads were aligned to the BioSpyder TempO-Seq Human Whole Transcriptome probe set (22 537 probes over 19 687 genes).

3.3.5 Data processing and quality control

Briefly, reads were demultiplexed from the BCL files and processed into FASTQ files using DRAGEN analysis v1.3.0 on the Illumina Basespace Sequence Hub (Illumina, Inc., San Diego, CA, USA). Processing and quality control of the FASTQ files, differential expression analysis, and exploratory statistical analyses were performed using the R-ODAF_Health_Canada pipeline (https://github.com/R-ODAF/R-ODAF_Health_Canada; accessed December 2024), which manages workflows using snakemake v8.0 (Mölder *et al.*, 2021) and relies on scripts in R (v.4.3.2). The preprocessing steps use fastp v0.23.2 for trimming (Chen *et al.*, 2018), STAR v2.7.8a (Dobin *et al.*, 2013) to perform alignment of raw reads to the reference sequence and the qCount function of the QuasR R package

(v1.30.0; Gaidatzis et al., 2015) to extract the feature counts from the aligned reads (BAM files) using features specified in a GTF file. A samples-by-probes count matrix was produced. Study-wide quality control was performed on the count matrix using several methods to measure consistency and remove low-quality samples, following the methods in Harrill et al. (2021) as a guideline. A cut-off of 0.1 for 1-Spearman's ρ was used to remove samples that were not correlated with others in this study. Like Harrill et al. (2021), we also used a 10% cut-off of uniquely mapped reads as the number of target sequences (e.g., 100 000 reads to pass filter when the target is 1 000 000 for TempO-Seq experiments). We removed any samples outside of Tukey's Outer Fence (3x interquartile range) for: (1) the number of probes capturing the top 80% of the signal, and (2) the number of detected probes (those with at least 5 mapped reads). Samples with a Gini coefficient (a measure of inequality in distributions) > 0.95 were excluded. Samples removed by these criteria are listed in Supplementary File 2.1.

To create a matrix for biomarker analysis, individual pairwise contrasts for each concentration and each chemical tested were created to the respective 0.1% DMSO control samples for each plate. Following the recommendations set out by the Omics Data Analysis Frameworks for Regulatory application (R-ODAF) guidelines (Verheijen et al., 2022), genes were filtered for each contrast tested to include only those where 75% of at least one experimental group were above 0.5 counts per million (CPM), and spurious spikes were removed in which $[\text{max} - \text{median}]$ of counts were less than $[\text{sum of counts}]/[\text{number of replicates} + 1]$. We used DESeq2 1.30.0 (Love et al., 2014) to estimate

fold changes and normalize for library size within the TempO-Seq data. The ashR method was used to perform $\log_2$ fold change shrinkage (Stephens, 2017). Data is available through the NCBI Gene Expression Omnibus with GEO accession GSE276869.

3.3.6 Gene expression biomarker analysis

To determine if the BPA alternatives activate ER α or other pathways, the expression profile of each chemical-concentration tested relative to solvent control (derived from the analysis of gene sets with unadjusted p value < 0.05 and absolute linear fold change ≥ 1.2) was compared to several characterized biomarkers as previously described (Kupersmidt *et al.*, 2010). The biomarkers included those that predict modulation of ER α (Corton, Matteo, *et al.*, 2024), nuclear factor erythroid 2-related factor 2 (NRF2; Rooney *et al.*, 2020), heat shock factor 1 (HSF1; Cervantes and Corton, 2021), metal-induced transcription factor 1 (MTF1; Jackson *et al.*, 2020), aryl hydrocarbon receptor (AhR; Oshida *et al.*, 2015), DNA damage (TGx-DDI; Li *et al.*, 2019), histone deacetylase inhibition (TGx-HDACi; Cho *et al.*, 2021), nuclear factor kappa B (NF- κ B; Korunes *et al.*, 2022), hypoxia inducible factor 1 (HIF1; Corton *et al.*, 2025), and cellular proliferation (CP; Corton, Ledbetter, *et al.*, 2024). As well, we included an unpublished biomarker that predicts unfolded protein response mediated by the X-box binding protein 1 (XBP1; unpublished data). The correlations between each biomarker and the gene lists of BPA and alternatives were determined using the Running Fisher algorithm in the BaseSpace Correlation Engine as described previously (Ryan *et al.*, 2016). The Running

Fisher algorithm provides an assessment of the statistical significance of the correlation of the overlapping genes between the biomarker and each gene list. A complete description of the Running Fisher test has been published (Kupersmidt *et al.*, 2010). The results were exported, and each p-value was converted to a $-\log(\text{p-value})$. Negative values were used to indicate negative correlation between the biomarker and the gene list. Thresholds for significance were set at $-\log(\text{p-value}) \geq 4$ for activation or ≤ -4 for suppression based on prior studies using this threshold (Matteo *et al.*, 2023).

3.3.7 Transcriptomic point of departure analysis

BMC modeling was done as described previously (Matteo *et al.*, 2023) in BMDEExpress3. A Williams trend test ($p < 0.05$) and absolute fold change ≥ 1.5 was applied to filter probes. Best fit models from polynomial 2°, linear, power (power term constrained to ≥ 1), and exponential models (degrees 3 and 5) were selected for each probe based on the lowest Akaike Information Criterion. A benchmark response of 1 standard deviation was selected. Probes with the following criteria were removed: 1) having a BMC greater than the highest concentration used in the analysis after excluding cytotoxic or non-soluble concentrations; 2) mapping to more than one gene; 3) having a model fit p-value < 0.1 determined by a likelihood ratio test; 4) having a BMC to BMC lower (BMCL) ratio > 20 , 5) having a BMC upper (BMCU) to BMCL ratio > 40 . Probes that met all the BMC filtering criteria were converted to their corresponding Entrez Identifiers. Data were further analyzed using R statistical software (version 4.3.2). Gene

accumulation plots were generated by rank ordering genes fitting BMCs (lowest to highest) and plotting the BMC of each gene on the x-axis and rank along the y-axis. Two types of tPODs were produced that we called: (1) general toxicity; and (2) ER α -specific response.

Three methods were used to derive general toxicity tPODs:

3. 25th rank-ordered gene BMC: As previously described (Matteo *et al.*, 2023), genes fitting BMCs were rank ordered from lowest to highest BMC and the BMC for the 25th gene was selected as an early transcriptomic tPOD. Confidence intervals were derived using the BMCU and BMCL as upper and lower limits, respectively.
4. Lowest pathway BMC: We derived the lowest pathway tPOD by aligning genes and their associated BMCL/BMC/BMCU values to the gene sets in Reactome Pathways database (<https://reactome.org/>) or the Kyoto Encyclopedia of Genes and Genomes (KEGG; <https://www.genome.jp/kegg/>). Gene sets that contained at least three genes fitting BMCs (genes that pass all criteria in the analysis), were at least 5% populated (based on total annotated gene number), and contained at least 40 genes in the whole gene set were selected. The 5th percentile BMC of the selected gene sets were selected as tPODs.
5. Lowest Consistent Response Dose (LCRD): We used an approach by Crizer and colleagues to identify the lowest concentration at which a consistent change in biological activity begins to occur (Crizer *et al.*, 2021).

We derived two ER α -specific tPODs by calculating (a) the median BMC and (b) the 5th percentile BMC of 50 genes that comprise a published ER α biomarker (Corton, Matteo, *et al.*, 2024). We first confirmed that a compound or mixture activated the ER α biomarker for at least one exposure concentration as described below (see Gene expression biomarker analysis). We further filtered by requiring a minimum of three genes with BMCs in the biomarker gene set to derive an ER α BMC. For all tPODs, confidence intervals for gene sets were estimated using a parametric bootstrap where the residuals were assumed to be normally distributed. For each gene, 100 bootstrap samples were generated followed by BMC analysis. Experiments were simulated using the bootstrapped results to obtain a BMC distribution. Experiment inclusion probabilities for each gene was estimated using the proportion of bootstrap samples that resulted in a BMC. For each simulated experiment a uniform random number between 0 and 1 was generated for each gene. If the random number was less than the gene's inclusion probability, the gene was retained and a BMC from the list of estimated BMCs for that gene was randomly selected. The median value from each of the 10000 simulated experiments was used to estimate the tPOD median. The 95% confidence interval was estimated using the 2.5th and 97.5th percentiles from the simulated distribution.

3.3.8 Mixture modelling

Since most of the BPA alternative chemicals tested activate ER α (Matteo *et al.*, 2023), we assumed that most would have additive effects when tested *in vitro*. We used the same method as previously described (Addicks *et al.*, 2023; Altenburger *et al.*, 2000) to predict the BMCs of mixtures of chemicals (BMC_{pred}) from the BMCs of the individual components (BMC_i) and their molar fractions (p_i) in equation 1.

$$\text{Equation 1: } BMC_{pred} = \left(\sum_{i=1}^n \frac{p_i}{BMC_i} \right)^{-1}$$

3.3.9 Ingenuity Pathway Analysis

Ingenuity Pathway Analysis (IPA; QIAGEN, Redwood City, CA, USA) was used to identify changes to upstream regulators, canonical pathways, and disease & functions. For each chemical tested, we imported Excel files into IPA containing gene IDs (Gene Symbol and Entrez ID) for the genes that fit a BMC (i.e., passed all filtering criteria), as well as their Williams trend test p-values from the BMC analysis, and the largest fold-changes of the gene relative to solvent controls (exported filtered data from BMDExpress3). This approach allowed us to analyze for the enrichment of genes showing robust concentration-responses to the exposures. IPA Core Analysis with a gene expression threshold of absolute fold change ≥ 1.5 and false discovery rate (FDR) adjusted p-value ≤ 0.05 was used with the direct and indirect relationship settings based on experimental and highly predicted data. Statistical significance of the overlap (FDR-

adjusted p-value ≤ 0.05) between the data set and known targets of upstream regulators in IPA were calculated using Fisher's Exact tests. The z-score was calculated using Fisher's Exact Test based on the expected relationship for directions between upstream regulators and target genes and those observed in the data set. A z-score of > 2 (activated) or < -2 (inhibited) was considered statistically significant. As a secondary analysis, we binned genes fitting BMCs by rank order (least to greatest BMC) set to 1 – 100, 101 – 200, and 201 – 300 to study specific transcriptomic changes initiated at increasing concentrations.

3.4 Results

3.4.1 Cell viability

The CellTiter-Blue Cell Viability assay was used to assess cytotoxicity after 48 hours of chemical exposure. There was no evidence of cytotoxicity following exposure to any mixture or individual BPA alternative chemical (Supplementary Figure 3.1).

3.4.2 General transcriptomic data summary

The median read depth was approximately 2 million per sample with a median range of 89.8% mapped reads. Of 672 samples, 70 experimental samples were lost to QC, along with 28 reference RNA and lysis buffer control samples (Supplementary File 2.1). The total number of samples per condition is listed in Supplementary Table 3.1.

The total number of genes fitting BMCs for each chemical tested is shown in Table 3.3. All mixtures of BPA alternatives and their individual components were transcriptionally active. Most chemicals and mixtures tested had 300 – 600 genes fitting BMCs. TMBPF (1191), M7 (714), and M1 (703) had the most genes fitting BMCs; P201 (205), 2,4'-BPF (200), and 2,4'-BPS (168) had the least. The gene accumulation plot shows the first 100 genes fitting BMCs for each mixture and chemical tested (Figure 3.2).

3.4.3 Stress response biomarker activation

We used published transcriptomic biomarkers to explore general cellular stress responses following exposure to mixtures of BPA alternative chemicals and their individual components. The biomarkers were specific to: (a) NRF2 (Rooney *et al.*, 2020); (b) HSF1 (Cervantes and Corton, 2021); (c) MTF1 (Jackson *et al.*, 2020); (d) AhR (Oshida *et al.*, 2015); (e) TGx-DDI (Li *et al.*, 2019); (f) TGx-HDACi (Cho *et al.*, 2021); (g) NF-kB (Korunes *et al.*, 2022); (h) HIF1 (Corton *et al.*, 2025); (i) CP (Corton, Ledbetter, *et al.*, 2024); (j) XBP1 (unpublished data). As in our previous work (Matteo *et al.*, 2023) we considered activation of at least two of these biomarkers within a condition as an indication of overt cellular stress (Escher *et al.*, 2020) and thus filtered out these concentrations from subsequent BMC analyses. Supplementary File 2.2 provides a summary of all stress response biomarkers activated and inhibited.

Many BPA alternative chemicals and all mixtures activated the CP biomarker at the highest exposure concentrations tested: 2,4'-BPF (50 μ M), 4,4'-BPF (10, 50 μ M), BPA (5, 10, 50 μ M), BPAF (5, 10, 50 μ M), BPB (5, 10, 50 μ M), BPC (5, 10, 50 μ M), BPS (50 μ M), BPZ (10, 50 μ M), M1 (5, 10, 50 μ M), M2 (5, 10, 50 μ M), M3 (10, 50 μ M), M4 (1, 5, 10, 50 μ M), M5 (5, 10, 50 μ M), M6 (10, 50 μ M), M7 (10.5, 21 μ M). Some chemicals inhibited the CP biomarker at exposure concentrations ≤ 1 μ M: 2,4'-BPS (0.5, 1, 5 μ M), BPAP (0.001, 0.5 μ M), BPB (0.01, 0.1 μ M), BPS (0.001, 0.01, 0.1, 0.5), BPZ (0.01, 0.1 μ M), P201 (0.01, 0.1, 0.5, 1 μ M). Few of the other transcriptomic biomarkers used in the study were perturbed by chemical exposure. Two chemicals inhibited the HIF1 biomarker BPB (50 μ M) and BPC (10 μ M), while the XBP1 biomarker was inhibited by M6 (10 μ M). Based on this analysis, no treatment conditions were considered overtly cytotoxic.

3.4.4 ER α biomarker activation

We used a 50 gene transcriptomic ER α biomarker (Corton, Matteo, *et al.*, 2024) to test the ability of mixtures of BPA alternative chemicals and their individual components to activate/inhibit ER α . Many BPA alternatives tested, including all mixtures, activated the ER α biomarker at the highest exposure concentrations tested (Supplementary Figure 3.2). As well, some chemicals inhibited the ER α biomarker at exposure concentrations ≤ 1 μ M (Supplementary Figure 3.2). Three chemicals did not activate the ER α biomarker at any concentrations tested: 2,4'-BPS, P201, and TMBPF. See Supplementary File 2.2 for a detailed breakdown of ER α biomarker activity. Overall, the results show that all mixtures

and most individual components activated the ER α biomarker at exposure concentrations $\geq 5 \mu\text{M}$ and some chemicals inhibited the biomarker at lower concentrations.

3.4.5 Transcriptomic Point of Departure (tPOD) derivation

We used two approaches to generate tPODs based on either (a) general toxicological effects or (b) ER α activation (mode of action-specific).

(a) Potency ranking for general toxicity

Given that transcriptomic activity other than that produced by ER α may cause adverse effects, we derived general toxicity tPODs to identify concentrations producing a ‘concerted molecular change’ (Johnson *et al.*, 2022). The gene accumulation plot shows the rank order of the potency of the mixtures of BPA alternatives and individual chemicals based on the 25th gene BMC, from most (left on graph) to least potent (right side of graph). BPAF was the most potent chemical tested, followed by M1, and M2 (Figure 3.2).

We used three approaches to identify exposure concentrations that cause a robust change in transcriptomic activity: (1) the 25th gene BMC, (2) the lowest pathway gene set 5th percentile BMC, and (3) the LCRD.

The 25th gene BMC is used to filter out genes that are within the noise range of the system and chemicals with no biological activity, while identifying the low end of the concentration range where the transcriptome is altered (see Figure 3.2). All the chemicals and mixtures tested had a 25th gene BMC, with most being between 0.1 – 1 μM (Table

3.3, Figure 3.3). BPAF was the most potent chemical tested based on the 25th gene (0.27 μ M), followed by M1 (0.35 μ M), and M2 (0.43 μ M). BPAP (5.0 μ M), 2,4'-BPS (5.8 μ M), and 2,4'-BPF (9.9 μ M) were the least potent chemicals tested in this experiment.

The 5th percentile BMC for the lowest pathway BMC was derived because this represents an approach akin to that originally recommended by the US National Toxicology Program (National Toxicology Program, 2018). The tPODs derived using this approach showed highly similar values and trends to the 25th gene BMC (Supplementary Table 3.2, Supplementary Figure 3.4). BPAF had the lowest 5th percentile lowest pathway BMC (0.19 μ M), followed by M2 (0.30 μ M) and M1 (0.30 μ M). BPS (6.6 μ M), 2,4'-BPF (19 μ M), and 2,4'-BPS (30 μ M) had the highest 5th percentile lowest pathway BMCs.

The LCRD identifies an inflection point in transcriptomic activity on the accumulation curve. For most chemicals and mixtures tested, the LCRD was one order of magnitude lower than the 25th gene and the pathway BMC (Supplementary Table 3.2, Supplementary Figure 3.4). BPAF (0.11 μ M), M1 (0.17 μ M), and M2 (0.18 μ M) had the lowest LCRDs, while BPAP (0.96 μ M), P201 (1.21 μ M), and 2,4'-BPF (1.64 μ M) had the highest.

(b) Potency ranking based on the ER α biomarker

To explore the potencies of the BPA alternatives and mixtures in activating the ER α , we identified the genes fitting BMCs that are part of the 50 gene transcriptomic ER α biomarker. Most ER α biomarker genes were activated at higher concentrations than the

25th gene BMC (Supplementary Figure 3.3). The BMCs for genes in the ER α biomarker were used to produce a 5th percentile BMC to rank ER α -specific potency of individual BPA alternatives and mixtures (Table 3.3, Figure 3.3). Most chemicals tested had ER α 5th percentile BMCs between 0.1 – 1.0 μ M. M1 had the lowest ER α 5th percentile BMC (0.082 μ M), followed by BPAF (0.15 μ M), and M3 (0.19 μ M). M7 (1.9 μ M), BPS (7.0 μ M), and 2,4'-BPF (37 μ M) had the highest ER α 5th percentile BMCs. We note that 2,4'-BPS, P201, and TMBPF did not activate the ER α biomarker and thus were not used to produce a 5th percentile ER α BMC. Two chemicals, BPAP and P201, did not meet the criteria for minimum number of genes fitting BMCs in the biomarker.

We also derived a median BMC using the 50 gene ER α biomarker (Supplementary Table 3.2, Supplementary Figure 3.4) as a more robust central metric for potency ranking. All mixtures of BPA alternative chemicals and most of their individual components had ER α biomarker median BMCs within one order of magnitude of each other. M1 had the lowest ER α biomarker median BMC (0.50 μ M), followed by BPAF (0.58 μ M), and M3 (0.64 μ M). BPZ (4.6 μ M), BPS (22 μ M), and 2,4'-BPF (37 μ M) had the highest ER α biomarker median BMCs. Although the BMCs for these medians were generally higher, the potency rank order was nearly identical to the 5th percentile ER α .

3.4.6 Mixture modelling

A main objective of this research was to determine if the BPA alternative chemicals interact or cause additive effects when present in mixtures. To address this, the BMCs of

individual BPA alternative chemicals were used to predict their BMCs when present in mixtures using an additive model (Equation 1). We compared the predicted mixture tPOD to the empirically derived tPOD using the 25th gene BMC, the lowest 5th percentile pathway BMC, and ER α 5th percentile BMC (Table 3.3, Figure 3.3), as well as ER α biomarker median BMC, and the LCRD (Supplementary Table 3.2, Supplementary Figure 3.4).

We found that most mixture BMC predictions were within the BMCL/BMCU of their experimentally derived BMCs (Table 3.3, Figure 3.3, Supplementary Table 3.2, Supplementary Figure 3.4). For M2 – M7, the predicted tPODs were virtually identical to their empirically derived tPODs. However, there were some notable differences for M1. The ER α 5th percentile BMC was one order of magnitude lower than predicted, but still within the empirical BMCL/BMCU (Table 3.3, Figure 3.3). Overall, these data support that mixtures of BPA alternative chemicals tested behave additively at concentrations near their BMCs.

3.4.7 Ingenuity Pathway Analysis

To determine the primary mechanisms altered by the chemicals and mixtures, we used the genes fitting BMCs, their maximum fold changes, and their p-values, in an IPA canonical pathway and upstream regulator analysis as described previously (Matteo *et al.*, 2023). The mixtures and many of their individual chemical components perturbed similar upstream regulators and pathways, supporting common modes of action. A full list

of canonical pathways and upstream regulators affected by chemical exposure is available in Supplementary File 2.3.

All the mixtures of BPA alternatives tested perturbed similar upstream regulators and canonical pathways along with several of their individual components: BPA, BPAF, BPC, 4,4'-BPF, BPB, and BPZ (Figure 3.4A and 4A). Most of the pathways predicted to be activated were associated with increased cellular proliferation and estrogen-related signaling (Figure 3.4A). Importantly, many of these same chemicals activated the CP biomarker (Supplementary File 2.2). All mixtures and most individual chemicals were predicted to inhibit MAGI1, NUPR1, and TP53 (Figure 3.4B). Further, these same chemicals were predicted to activate TFEB, CKAP2L, RABL6, FOXM1, MITF, COP1, and ERBB2. In contrast, BPS, BPAP, 2,4'-BPF, 2,4'-BPS, P201, and TMBPF mostly caused opposite effects relative to the other chemicals tested, but these typically did not reach statistical significance (Figure 3.4B). All mixtures and most individual BPA alternative chemicals that were predicted to activate ESR1 were identified as ER α activators using the transcriptomic biomarker (Supplementary File 2.2). Some chemicals that were identified as inhibitors of the ER α biomarker, such as 2,4'-BPS and P201, were also predicted to inhibit ESR1 using this IPA-based approach.

We rank ordered and binned the genes fitting BMCs into three bins to explore the mechanisms associated with transcriptional changes occurring in different BMC ranges:

- Bin 1: genes BMCs ranked 1 – 100 (Lowest BMCs, Supplementary Figure 3.5)

- Bin 2: genes BMCs ranked 101 – 200 (Medium BMCs, Supplementary Figure 3.6)
- Bin 3: genes BMCs ranked 201 – 300 (High BMCs, Supplementary Figure 3.7)

For Bin 1, most of the mixtures, along with BPA, BPAF, and BPC, were enriched for pathways and upstream regulators associated with cellular proliferation and the estrogen receptor (Supplementary Figure 3.5). Bin 2 gene sets were over-represented by the mixtures, with fewer genes fitting BMCs leading to enriched pathways for BPA, BPAF, and BPC (Supplementary Figure 3.6). Interestingly, P201 inversely modulated the genes enriched in this bin relative to most of the other chemicals in Bin 2. In Bin 3, most mixtures, as well as 4,4'-BPF, BPB, and BPZ were enriched in similar pathways and upstream regulators associated with cellular proliferation and the estrogen receptor (Supplementary Figure 3.7). Finally, the genes with BMCs for the chemicals 2,4'-BPS (168 genes fitting BMCs; Bins 1 and 2 only), 2,4'-BPF (200 genes fitting BMCs; Bins 1 and 2 only), and TMBPF (1191 genes fitting BMCs) were not enriched in the same pathways and upstream regulators as the other chemicals and mixtures.

3.5 Discussion

Given the growing number of BPA replacements and human exposure to these compounds in mixtures, we tested whether mixtures of BPA alternatives have additive effects in MCF-7 cells using HTTr. We previously determined the potency of BPA and 15 alternative chemicals in MCF-7 cells (Matteo *et al.*, 2023) and used the most potent ER α

active chemicals from that study to design six equimolar mixtures. We produced a seventh complex mixture using the relative proportions of BPA alternatives detected in human urine samples (Gys *et al.*, 2020). We did BMC modelling on gene expression data to derive ER α -specific and general toxicity tPODs. We applied a predictive model previously used by our lab to test for additive effects *in vitro* (Addicks *et al.*, 2023). All the mixtures in our study activated ER α based on a transcriptomic biomarker analysis. In addition, using tPODs derived from the ER α biomarker and general toxicity methods it was observed that all predicted mixture tPODs were nearly identical to their experimentally derived tPODs. Genes fitting BMCs analyzed with IPA revealed that mixtures perturbed similar upstream regulators and canonical pathways as their individual components. Collectively, our data support additive effects of BPA alternative chemicals *in vitro* through both their ER α activities and alternative mechanisms of toxicity that they perturb.

All mixtures and most of the BPA alternatives tested were identified as ER α active. Importantly, equimolar mixtures, as well as the complex mixtures tested, activated ER α at similar concentrations as BPA. This aligns with previous reports of mixtures of BPA alternatives (that include BPA) activating ER α *in vitro* (Bermudez *et al.*, 2010; Park *et al.*, 2020; Lee *et al.*, 2023). Interestingly, three chemicals that we previously determined to be ER α agonists (2,4'-BPS, BPAP, and P201) (Matteo *et al.*, 2023), inhibited the 50 gene ER α biomarker in this study. Further, several BPA alternatives tested inhibited the 50 gene ER α biomarker at concentrations below 1 μ M despite being ER α active above this

concentration. Together, these data suggest that BPA alternatives may have non-monotonic effects, wherein low exposure concentrations ($< 1 \mu\text{M}$) inhibit ER α , which aligns with the low dose effects of BPA and other nuclear hormone endocrine disruptors (Vandenberg *et al.*, 2012; Shioda *et al.*, 2013). We also used several transcriptomic biomarkers to detect concentrations at which cells would be overtly stressed by chemical exposure, alongside an established benchtop assay for cell viability. No conditions were cytotoxic in this study, in line with previous data from our lab (Matteo *et al.*, 2023). Our data demonstrate that mixtures of BPA alternatives activate ER α at similar concentrations as BPA and that some alternatives may inhibit ER α at low exposure concentrations.

Equimolar and human relevant mixtures of BPA alternative chemicals tested had additive effects in MCF-7 breast cancer cells. Across all tPODs derived in this study, the predicted and empirical BMCs were within an order of magnitude, supporting the use of concentration addition models for assessing mixtures of BPA alternatives *in vitro*. Others have previously used transcription-based assays to predict additive ER α transactivation by mixtures of BPA alternatives (Lee *et al.*, 2023; Bermudez *et al.*, 2010). Moreover, another group exposed MCF-7 cells to an equimolar mixture of BPA alternatives ($0.35 \mu\text{M}$) and detected almost no changes in transcriptomic activity using HTTr after 24 h exposure, despite causing significant biological activity in MCF-10 breast epithelial cells that lack ER α expression (Mesnage *et al.*, 2022)). We also observed very few differentially expressed genes in MCF-7 cells using a mixture of similar composition (Mix 7) at concentrations below $1 \mu\text{M}$ (data not shown). In both studies, the mixtures were mostly

composed of ER α active BPA alternatives, which is the primary mode of action of BPA and many alternative chemicals in MCF-7 cells. It is possible that exposure to mixtures of BPA alternatives at concentrations below 1 μ M may not be sufficient to cause robust transcriptional activity in MCF-7 cells. These data support the use of HTTr data to test for chemical additivity *in vitro* and that mixtures of BPA alternatives have additive effects in ER α positive breast cancer cells. Further research of the effects of exposure to BPA alternative chemicals in physiologically relevant systems like normal breast epithelial cells is warranted.

We used several tPODs to rank BPA alternative chemicals based on toxicological potency *in vitro*. The 25th gene BMC tPOD was used to highlight early changes to the transcriptome caused by chemical exposure and rank chemicals for overall toxicological potency. We also employed a distribution-based approach to derive a conservative gene set tPOD to capture relevant biological activity as previously described (Reardon *et al.*, 2023). Further, an LCRD was used to determine the lowest concentration associated with significant transcription. All tPODs produced similar potency rankings and most were within one order of magnitude of each other. BPAF, a fluorinated alternative, was the most potent chemical tested overall, followed by M1, and M2. In line with additive effects, mixtures containing the most potent individual chemicals had the lowest tPODs. BPA was the second most potent individual chemical tested, followed by BPC, similar to our previous HTTr potency ranking exercise (Matteo *et al.*, 2023). Interestingly, TMBPF, a non-estrogenic alternative to BPA (Soto *et al.*, 2017), caused the most transcriptional

activity of all chemicals tested, despite being less potent than many chemicals tested. Overall, both general and ER α tPODs aligned to produce similar potency rankings, with mixtures of BPA alternatives among the most potent chemicals tested. These data support the use tPODs to determine the potency of chemical mixtures *in vitro*.

All genes fitting BMCs were analyzed using IPA to identify potential modes of action of mixtures of BPA alternatives in MCF-7 cells. In line with our previous study (Matteo *et al.*, 2023), many individual BPA alternatives perturbed similar upstream regulators including ESR1, RARA, and COP1. Moreover, mixtures and most of their individual components affected similar upstream regulators and pathways associated with loss of cell cycle control, supporting additive effects in breast cancer cells. Specifically, all mixtures tested were predicted to inhibit tumor suppressor genes like P53 (Lim *et al.*, 2009) and MAGI1 (Wörthmüller *et al.*, 2023). Interestingly, NUPR1, a gene associated with tamoxifen resistance in breast cancer cells (Wang *et al.*, 2021), was predicted to be inhibited in IPA. Since NUPR1 interacts with ESR1 (Wang *et al.*, 2021), it is possible that exposure to mixtures of BPA alternatives inhibited this pathway. Further, all mixtures were predicted to activate genes associated with breast cancer tumor progression like TFEB (T. Wang *et al.*, 2023), RABL6 (Li *et al.*, 2013), FOXM1 (Ziegler *et al.*, 2019), MITF (Y. Zhang *et al.*, 2024), and CKAP2L (dos Santos *et al.*, 2022). When we binned genes according to their BMC rank order, we found that most of the affected upstream regulators and pathways were within the first 200 genes fitting BMCs. This suggests that for most chemicals tested, exposure to higher concentrations of mixtures of BPA

alternatives caused non-specific transcriptomic activity. Finally, despite having the most genes fitting BMCs, TMBPF perturbed the fewest upstream regulators and pathways, suggesting that it caused non-specific biological activity in MCF-7 cells at all concentrations tested. Collectively, these data support that mixtures of BPA alternatives have additive effects and dysregulate cellular proliferation *in vitro*.

In summary, our HTTr data add to the weight of evidence that mixtures of BPA alternative chemicals have additive effects *in vitro*. Transcriptomic biomarker analysis revealed that all mixtures and most BPA alternatives tested activated ER α at similar concentrations. General toxicity and ER α -related tPODs produced comparable potency rankings, with BPAF as the most potent chemical overall and ER α activator tested, in line with previous data from our lab and others. All mixtures and most individual chemicals had common modes of action via dysfunctional cellular proliferation and ER α -related pathways. Further research on the effects of individual BPA alternatives and mixtures thereof in physiologically relevant systems, such as normal breast epithelial cells, is warranted. Overall, these data support the use of HTTr to test mixtures as part of hazard characterization for human health risk assessment.

3.6 Tables and Figures

Table 3.1 List of chemicals and suppliers.

Chemical	Acronym	CAS Number	Supplier
Bisphenol A	BPA	80-05-7	Sigma-Aldrich Inc (St. Louis, MO, USA)
Bisphenol AF	BPAF	1478-61-1	
4,4'-Bisphenol F	4,4'-BPF	620-92-8	
Bisphenol S	BPS	80-09-1	
Bisphenol AP	BPAP	1571-75-1	
Estradiol	E2	50-28-2	
Dimethyl sulfoxide	DMSO	67-68-5	
Bisphenol B	BPB	77-40-7	
Bisphenol Z	BPZ	843-55-0	
Tetramethyl bisphenol F	TMBPF	5384-21-4	
Pergafast 201	P201	232938-43-1	Toronto Research Chemicals (Toronto, ON, Canada)
2,4'-Bisphenol S	2,4'-BPS	5397-34-2	
Bisphenol C	BPC	79-97-0	Accustandard (New Haven, CT, USA)
2,4'-Bisphenol F	2,4'-BPF	2467-03-0	TCI America (Portland, OR, USA)

Table 3.2 List of mixtures and relative proportions.

Mixtures	Components	Proportions
M1	BPA:BPAF	Equimolar
M2	BPA:BPAF:BPC	Equimolar
M3	BPA:BPAF:BPC:4,4'-BPF	Equimolar
M4	BPA:BPAF:BPC:4,4'-BPF:BPAP	Equimolar
M5	BPA:BPAF:BPC:4,4'-BPF:BPAP:BPS	Equimolar
M6	BPA:BPAF:BPC:4,4'-BPF:BPAP:BPS:2,4'-BPF:P201:2,4'-BPS	Equimolar
M7	BPA:BPAF:4,4'-BPF:BPAP:BPS:BPB:BPZ	318:1:1789:6:320:3:13

Table 3.3 Comparison of transcriptomic points of departure (tPODs).

Chemical	# Genes Fitting BMCs	25th Gene BMC	25th Gene BMCL/BM CU	25th Gene Predicted BMC	# Genes Fitting ERa biomarker	ERa 5th Percentile BMC	ERa 5th Percentile BMCL/BM CU	ERa 5th Percentile Predicted BMC
BPA	609	0.49	0.10 - 1.56		21	0.53	0.17 - 1.43	
BPAF	413	0.27	0.05 - 0.58		20	0.15	0.04 - 0.41	
BPC	562	0.61	0.19 - 4.54		23	0.51	0.12 - 1.65	
BPB	559	0.61	0.50 - 1.02		21	0.54	0.25 - 1.86	
4,4'-BPF	511	0.93	0.19 - 3.17		20	1.14	0.45 - 2.56	
BPZ	447	0.97	0.51 - 2.70		14	1.20	0.29 - 3.38	
TMBPF	1191	3.88	1.05 - 7.81		4			
P201	205	5.34	3.29 - 13.74		1			
BPS	356	5.67	3.56 - 13.55		12	6.96	4.13 - 10.0	
BPAP	265	8.03	4.99 - 11.6		2			
2,4'-BPS	168	9.28	5.75 - 9.60		0			
2,4'-BPF	200	9.89	4.57 - 41.21		3	36.8	22.2 - 106	
M1	703	0.35	0.060 - 1.69	0.35	18	0.08	0.013 - 0.36	0.23
M2	498	0.43	0.080 - 1.13	0.40	20	0.24	0.070 - 0.64	0.28
M3	527	0.48	0.10 - 1.02	0.47	19	0.19	0.054 - 0.52	0.34
M4	539	0.69	0.33 - 1.43	0.58	23	0.83	0.38 - 1.94	0.43
M5	634	0.75	0.37 - 4.09	0.68	24	0.79	0.39 - 1.32	0.51
M6	585	1.24	0.32 - 5.00	0.98	17	0.75	0.22 - 2.13	0.76

M7	714	1.4 0	1.17 - 9.27	1.40	17	1.91	1.24 - 2.84	1.66
----	-----	----------	----------------	------	----	------	----------------	------

tPODs were derived by prefiltering data using the Williams trend test ($p < 0.05$) and an absolute fold-change filter of ≥ 1.5 and postfiltered with the following settings in BMDExpress v3: Best BMD/BMDL < 20 , Best BMDU/BMDL < 40 , and Best fitPvalue ≥ 0.1 ; tPODs (shown in μM) representing the 25th rank ordered gene benchmark concentration (BMC) and the 5th percentile gene BMC for the estrogen receptor alpha (ER α) biomarker gene set are shown; most chemicals are shown in decreasing order of potency based on tPODs from the 25th gene BMC, followed by the mixtures

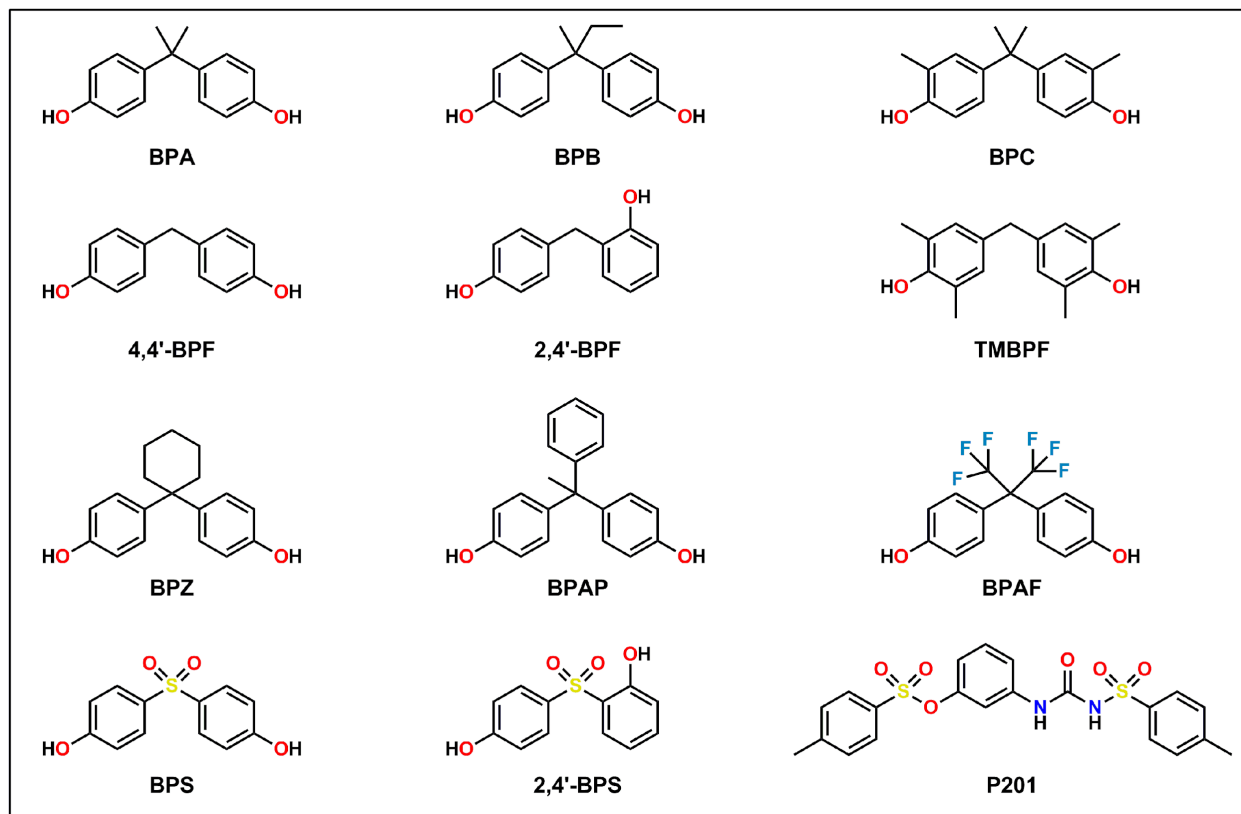


Figure 3.1 Chemical structure of bisphenol A (BPA) and alternative chemicals.

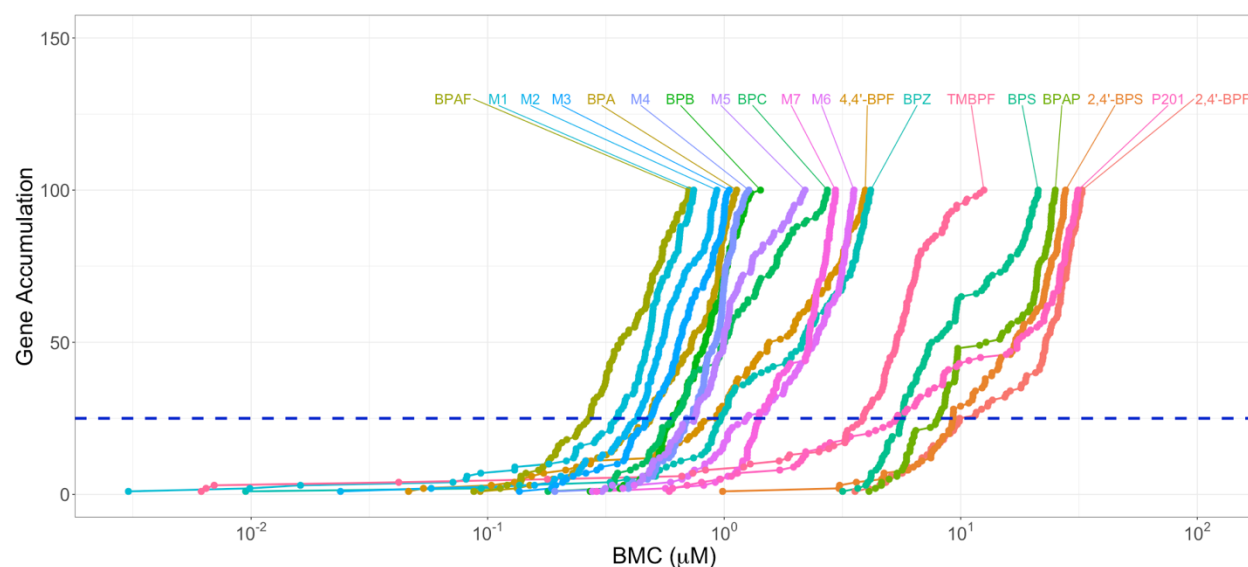


Figure 3.2 Gene accumulation plot of MCF-7 cells ($n = 2 - 4$ per concentration) exposed to seven mixtures of bisphenol A (BPA) alternatives and 12 individual chemicals at a range of concentrations ($0.001 - 50 \mu\text{M}$) for 48 h. The blue dashed line indicates the 25th ranked gene. Data were prefiltered using the Williams trend test ($p < 0.05$) and an absolute fold-change filter of ≥ 1.5 , and post filtered with the following settings in BMDExpress v3: Best BMD/BMDL < 20 , Best BMDU/BMD < 20 , Best BMDU/BMDL < 40 , and Best fitPvalue ≥ 0.1 .

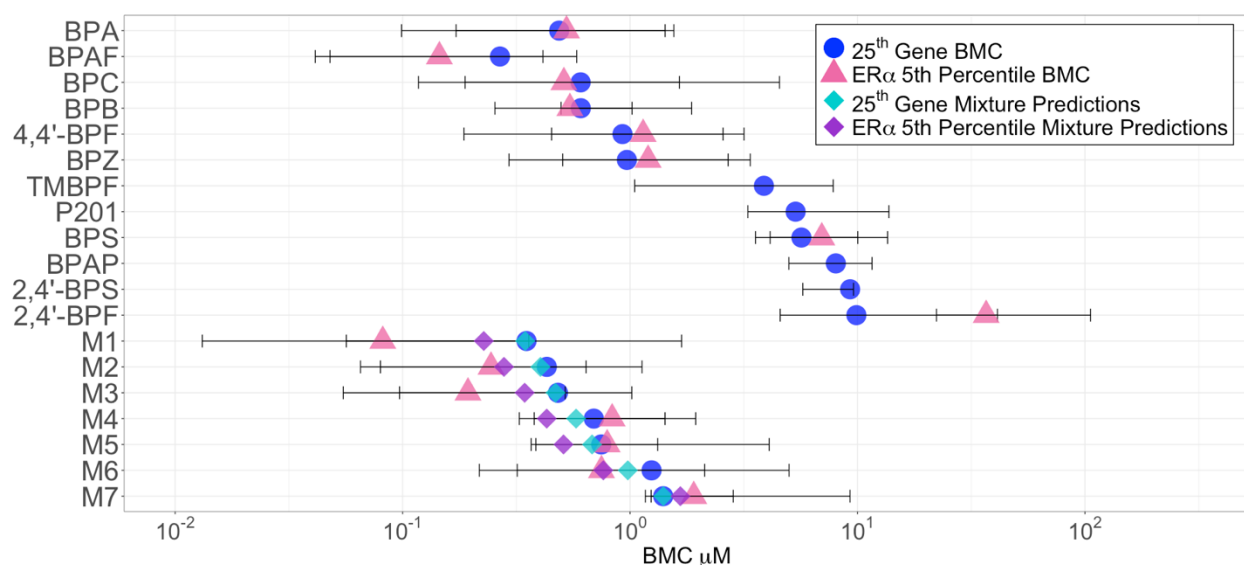


Figure 3.3 Comparison of transcriptomic points of departure (tPODs) derived from whole transcriptome analysis of MCF-7 cells ($n = 2 - 4$ per concentration) that were exposed to seven mixtures of bisphenol A (BPA) alternatives and 12 individual chemicals at a range of concentrations ($0.001 - 50 \mu\text{M}$) for 48 h. tPODs were derived by prefiltering data using the Williams trend test ($p < 0.05$) and an absolute fold-change filter of ≥ 1.5 . Data were also postfiltered with the following settings in BMDEExpress v3: Best BMD/BMDL < 20 , Best BMDU/BMD < 20 , Best BMDU/BMDL < 40 , and Best fitPvalue ≥ 0.1 . tPODs (shown in μM) representing the 25th rank ordered gene benchmark concentration (BMC) and the 5th percentile gene BMC for the estrogen receptor alpha ($\text{ER}\alpha$) biomarker gene set are shown. The $\text{ER}\alpha$ tPOD is based on at least one concentration yielding an agonist or antagonist call and at least 3 genes. The reference chemical BPA is shown at the top. BMCL and BMCU are used for the lower and upper bounds, respectively. Most chemicals are shown in decreasing order of potency based on tPODs from the 25th gene BMC, followed by the mixtures. Predictions based on additivity of individual components are shown as diamonds for the mixtures.

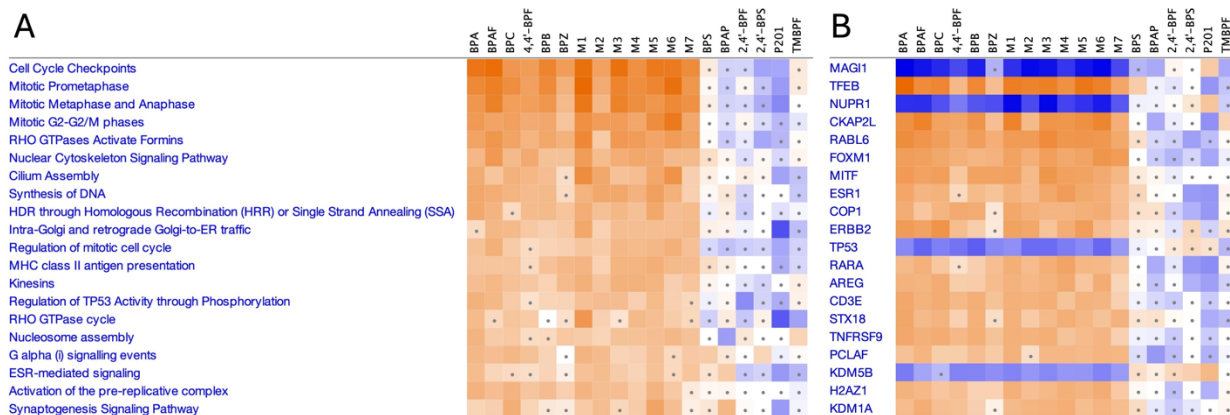


Figure 3.4 Ingenuity Pathway Analysis of all genes fitting benchmark concentration models in MCF-7 cells ($n = 2 - 4$ per concentration) that were exposed to seven mixtures of bisphenol A (BPA) alternatives and 12 individual chemicals at a range of concentrations ($0.001 - 50 \mu\text{M}$) for 48 h. A list of genes fitting models, the Williams trend test p value, and the maximum fold-change were imported into IPA for this comparison. Filters were set to Z-score ≥ 2.0 and p-value < 0.05 . Orange denotes predicted activation and blue predicted inactivation (dots signify Z-score < 2). (A) Top 20 canonical pathways. (B) Top 20 upstream regulators.

Chapter 4: Transcriptomic analysis of BPA alternative chemicals in human mammary epithelial cells

Authors: Matteo G, Rowan-Carrol A, Bradford LM, Meier MJ, Williams A, J. Corton JC, Yauk CL, Atlas E

Target journal: *Environment International*

4.1 Abstract

Some everyday consumer products contain endocrine disruptors like bisphenol A (BPA) and its functional alternatives. To date, most *in vitro* chemical screening to evaluate these compounds has been accomplished using immortalized cell lines, which differ significantly from human tissues. Our goal was to test BPA and select alternatives previously screened in breast cancer cells for toxicological potency and mechanism of action in human mammary epithelial cells (HMECs). HMECs from three human donors were exposed to BPA and four alternative chemicals (in concentration response format from 0.001 – 50 μ M) for 48 h and global transcriptomic changes were quantified. Transcriptomic biomarker analysis was employed to explore chemically induced estrogen receptor alpha (ER α) activation and alterations in stress response pathways. Benchmark concentration (BMC) analysis was applied to gene expression data to derive

transcriptomic points of departure (tPODs) to compare chemicals for potency. Gene enrichment analysis of upstream regulators and canonical pathways was applied among the genes fitting BMCs. All chemicals had tPODs within an order of magnitude of each other. BPAF was the most potent, followed by TMBPF, BPC, BPS, and BPA. None of the chemicals activated the ER α active biomarker. Some stress response biomarkers were activated at high exposure concentrations. Genes fitting BMCs clustered chemicals into two groups, with one group (BPAF and TMBPF) primarily inhibiting gene sets and the other (BPC, BPS, and BPA) with mostly activating expression patterns. These data suggest that the BPA alternatives tested have similar toxicological potencies in HMECs and oppositely enrich cancer related gene sets.

4.2 Introduction

Bisphenol A (BPA) is used to make polycarbonate plastics and epoxy resins, along with being used as a developing agent in some thermal papers. BPA is an endocrine disrupting chemical (EDC) because of its interactions with several nuclear hormone receptors (Manzoor *et al.*, 2022). Consequently, it is regulated in many jurisdictions worldwide and banned in some products like baby bottles in Canada. Chemical alternatives to BPA are increasingly used in commercial products such as canned foods, thermal paper and more (Ni *et al.*, 2023). Thus, exposure to these alternatives is increasing in the general population (Acevedo *et al.*, 2025). Some BPA alternatives have been subjected to toxicological screening, mostly in immortalized cell lines of cancerous

origin (e.g., Meier et al., 2024; Beal et al., 2024; Matteo et al., 2023, 2025). While these models share some key characteristics with human tissues, they differ significantly from normal tissues in many ways.

The human mammary gland is hormonally sensitive and a known target of endocrine disruptors. Mammary epithelial cells are composed of subpopulations including luminal epithelial and myoepithelial cells, with the former thought to be the origin of most breast tumors (Yersal and Barutca, 2014). These cells express nuclear hormone receptors like estrogen, progesterone, and prolactin (Hannan *et al.*, 2023). The mammary epithelium is exposed to EDCs, either directly through circulation or indirectly via the mesenchyme/stroma. BPA has been detected in breast adipose tissue of breast cancer patients (Reeves *et al.*, 2018) and levels of BPA metabolites have been associated with concentrations of BPA in breast adipose tissue (Keshavarz-Maleki *et al.*, 2021). Further, evidence from rodent studies supports that exposure to BPA and alternatives like bisphenol S (BPS) during pregnancy and adulthood negatively breast development and lactation, respectively (for a review, see Vandenberg, 2021). These data add to the weight of evidence that the mammary gland is exposed to chemicals like BPA that can negatively impact its physiology. However, the potential mechanisms by which BPA and its alternatives interact with mammary epithelial cells are not well elucidated.

Immortalized breast cells lines like breast cancer cells that overexpress the ER α have been used to screen EDCs like BPA and its alternatives for ER α activity for chemical risk assessment. For example, our lab (Matteo *et al.*, 2023) and others (Harrill *et al.*, 2024;

Beal *et al.*, 2024) have used high-throughput transcriptomics (HTTr) to compare the toxicological potency and modes of action of BPA to several alternative chemicals in MCF-7 breast cancer cells. These data support that several BPA alternatives are estrogenic, i.e. can activate ER α , and have a similar potency as BPA in breast cancer cells. However, only a subpopulation (< 10 %) of non-cancerous breast epithelial cells are ER α positive in human tissue (Petersen *et al.*, 1987), and these cells require selective culturing to maintain ER α expression (Fridriksdottir *et al.*, 2015). Therefore, BPA and its alternatives may differ in toxicological potency and mechanisms in primary mammary cells that do not overexpress the ER α .

Human mammary epithelial cells (HMECs) can be isolated and cultured from non-cancerous donor breast tissue. Exposure of HMECs to nanomolar concentrations of BPA is associated with increased cellular proliferation (Qin *et al.*, 2012) and perturbation of genes associated with cancer related pathways (Weng *et al.*, 2010; Qin *et al.*, 2012). Exposure to BPA increases the number of estrogen receptor alpha (ER α)-positive cells and leads to nuclear translocation of the ER α in HMECs cultured in 3D for several weeks (Weng *et al.*, 2010). Like breast tissue, only a subpopulation of HMECs cultured in 3D express the ER α (Rosenbluth *et al.*, 2020), suggesting that BPA exposure promotes survival of ER α -positive epithelial subpopulations after prolonged culturing. Interestingly, when compared to BPA and 4,4'-bisphenol F (4,4'-BPF), BPS caused the most significant changes to HMEC morphology and branching in 3D (Winkler *et al.*, 2022). These data

support that exposure to BPA and its alternatives affect morphology and gene expression in HMECs.

To address the need for toxicological data from BPA alternatives in more human-relevant models, we assessed the transcriptional activity and toxicological potency in HMECs from three human donors. We were interested in exploring inter-donor variability and motivated to explore potential mechanisms in these cells given that ER α is only expressed in a minority of primary breast cells. HMECs from three donors were exposed to BPA and four alternative chemicals across eight concentrations (0.001 – 50 μ M) for 48 h, 17 β -estradiol (E2; 1 and 10 nM), and solvent control (0.1% dimethylsulfoxide; DMSO). The BPA alternative chemicals selected were the most toxicologically potent in our previous HTTr studies in breast cancer cells (Matteo *et al.*, 2023, 2025). Our results reveal that while the tested BPA alternatives exhibited similar toxicological potencies, they showed divergent effects on stress response pathways and cancer-related gene networks, highlighting distinct mechanisms of action that may be masked in traditional breast cancer cell models.

4.3 Materials and methods

4.3.1 Chemicals

Table 4.1 lists chemicals used and their respective suppliers. All chemicals (Figure 4.1, Table 4.1) were dissolved in DMSO with a final concentration in exposure media of 0.1%.

4.3.2 Cell culture

HMECs (CC-2551, Lonza Bioscience, Basel, Switzerland) were sourced from three biological female donors (see Supplementary Table 4.1). Cells were cultured in T75 flasks (Corning Falcon No. 156499, Corning New York) with supplemented Mammary Epithelial Cell Growth Medium containing phenol red (MEGM; Lonza Bioscience) at 37°C in a humidified atmosphere (5% CO₂). For the experiment, cells (passage 7) were seeded at 2.0×10^4 cells per well in supplemented MEGM without phenol red (Lonza Bioscience). After 48 h, the media were removed and replaced with supplement-free MEGM without phenol red containing chemicals of interest at eight concentrations (0.001, 0.01, 0.1, 0.5, 1, 5, 10, 50 µM) and E2 (0.1 and 1 nM). The experiment included at least six plate-matched solvent controls (0.1% DMSO). Four technical replicates from each donor were exposed for 48 h based on our previous experiments with MCF-7 cells (Matteo *et al.*, 2023). After exposure, cells were washed with phosphate-buffered saline (PBS) and lysed *in situ* using 2x TempO-Seq lysis buffer diluted with an equal amount of PBS for TempO-Seq library building.

4.3.3 Cell viability assay

To determine if chemical exposure decreased cell viability, HMECs were seeded into black 96-well plates (Corning Falcon No. 3603) and treated as described above (three technical replicates per donor). Cell viability was measured using the CellTiter-Blue Cell Viability Assay (Promega Corp, Madison, Wisconsin), as per the manufacturer's instructions. This is a fluorometric assay that measures the metabolic capacity of cells based on resazurin reduction. Briefly, after 48-h exposure, CellTiter-Blue reagent was added to each well, and the plates were incubated (37°C, 5% CO₂) for 2 h. The fluorescence was read using an excitation wavelength of 560 nm and emission of 590 nm using a SpectraMax M2 (Molecular Devices LLC, San Jose, California). Fluorescence readings were expressed as a percent ratio to that of the DMSO control group. Cytotoxicity was defined as readings < 50% of the control.

4.3.4 TempO-Seq library building and next-generation sequencing

Gene expression was measured using the TempO-Seq Human Whole Transcriptome v2.0 and 2.1 kits (BioSpyder Technologies Inc, Carlsbad, California) as per the manufacturer's instructions and as previously described (Matteo *et al.*, 2023, 2025). Cell lysates and positive technical controls (Human Universal Reference RNA—uhrRNA Agilent Cat No. 740000, Santa Clara, California, and Human Brain Total RNA brRNA—ThermoFisher AM7962, Waltham, Massachusetts), as well as no-cell negative

controls (1x TempO-Seq lysis buffer) were hybridized to the Detector Oligo (DO) Pool using an annealing kit for the whole human genome supplied by BioSpyder. The hybridization mixture was incubated for 10 min at 70°C followed by a temperature gradient with a ramp rate of 0.5°C /min to 45°C over 50 min with a 16-h hold at 45°C and then cooled to 25°C. Nuclease digestion was employed to remove excess, unbound, or incorrectly bound DOs at 37°C for 90 min. Amplification templates were generated by ligating DO pairs bound to adjacent target sequences for 1 h at 37°C, followed by enzyme denaturation for 15 min at 80°C. Amplification templates (10 µL) were pipetted into a 96-well PCR Pre-Mix and Primer plate supplied by BioSpyder and amplified using a CFX96 Real-Time PCR Detection System (Bio-Rad) to incorporate a unique sample index/ tag sequence and the sequencing adaptors for each sample. The following PCR settings were used: 37°C for 10 min, 95°C for 1 min; 25 cycles of 95°C for 10s, 65°C for 30s, 68°C for 30s (with optical read for visual sample quality control [QC]); 68°C for 2 min; and hold at 25°C prior to library pooling and purification.

NucleoSpin Gel and PCR Clean-up kits were used to pool and purify labeled amplicons. Next-generation sequencing libraries were sequenced using a NextSeq 500 High-Throughput Sequencing System (Illumina, San Diego, California), using 50 cycles from a 75-cycle high-throughput flow cell to achieve a median read depth of 2 million reads per sample. Data were processed as described below and reads were aligned to the BioSpyder TempO-Seq Human Whole Transcriptome probe set (22 537 probes over 19 687 genes).

4.3.5 Data processing

Reads were demultiplexed from the BCL files and processed into FASTQ files using DRAGEN analysis v1.3.0 on the Illumina Basespace Sequence Hub (Illumina, Inc., San Diego, CA, USA). Processing and quality control of the FASTQ files, differential expression analysis, and exploratory statistical analyses were performed using the R-ODAF_Health_Canada pipeline (https://github.com/R-ODAF/R-ODAF_Health_Canada; accessed December 2024), which manages workflows using snakemake v8.0 (Mölder *et al.*, 2021) and relies on scripts in R (v.4.3.2). The preprocessing steps use fastp v0.23.2 for trimming (Chen *et al.*, 2018), STAR v2.7.8a (Dobin *et al.*, 2013) to perform alignment of raw reads to the reference sequence, and the qCount function of the QuasR R package (v1.30.0; Gaidatzis *et al.*, 2015) to extract the feature counts from the aligned reads (BAM files) using features specified in a GTF file. A samples-by-probes count matrix was produced.

4.3.6 Quality control

Study-wide quality control was performed on the count matrix using several methods to measure consistency and remove low-quality samples, following the methods in Harrill *et al.* (2021) as a guideline. A cut-off of 0.1 for 1-Spearman's ρ was used to remove samples that were not correlated with others in this study. Like Harrill *et al.* (2021),

we also used a 10% cut-off of uniquely mapped reads as the number of target sequences (e.g., 100 000 reads to pass filter when the target is 1 000 000 for TempO-Seq experiments). We removed any samples outside of Tukey's Outer Fence (3x interquartile range) for: (1) the number of probes capturing the top 80% of the signal, and (2) the number of detected probes (those with at least 5 mapped reads). Samples with a Gini coefficient (a measure of inequality in distributions) > 0.96 were excluded. Samples removed by these criteria are listed in Supplementary File 3.1.

4.3.7 Transcriptomic biomarkers

To create a matrix for biomarker analysis, individual pairwise contrasts for each concentration and each chemical tested were created to the respective 0.1% DMSO control samples for each plate. Following the recommendations set out by the Omics Data Analysis Frameworks for Regulatory application (R-ODAF) guidelines (Verheijen *et al.*, 2022), genes were filtered for each contrast tested to include only those where 75% of at least one experimental group were above 0.5 counts per million (CPM), and spurious spikes were removed in which $[\text{max} - \text{median}]$ of counts were less than $[\text{sum of counts}]/[\text{number of replicates} + 1]$. We used DESeq2 1.30.0 (Love *et al.*, 2014) to estimate fold changes and normalize for library size within the TempO-Seq data. The ashR method was used to perform $\log_2$ fold change shrinkage (Stephens, 2017).

To determine if the BPA alternatives activate biological pathways, the expression profile of each chemical-concentration tested relative to solvent control (derived from the

analysis of gene sets with unadjusted p value < 0.05 and absolute linear fold change ≥ 1.2) was compared to several characterized biomarkers as previously described (Kupersmidt *et al.*, 2010). The biomarkers predict modulation of ER α (Corton, Matteo, *et al.*, 2024), nuclear factor erythroid 2-related factor 2 (NRF2; Rooney *et al.*, 2020), heat shock factor 1 (HSF1; Cervantes and Corton, 2021), metal-induced transcription factor 1 (MTF1; Jackson *et al.*, 2020), DNA damage (TGx-DDI; Li *et al.*, 2019), histone deacetylase inhibition (TGx-HDACi; Cho *et al.*, 2021), nuclear factor kappa B (NF- κ B; Korunes *et al.*, 2022), hypoxia inducible factor 1 alpha (HIF1 α ; Corton *et al.*, 2025), and cellular proliferation (CP; Corton, Ledbetter, *et al.*, 2024). We also analyzed an unpublished biomarker that predicts unfolded protein response mediated by the X-box binding protein 1 (XBP1; Corton, unpublished data). The correlations between each biomarker and the gene lists for each concentration / chemical / donor combination were determined using the Running Fisher algorithm in the BaseSpace Correlation Engine as described previously (Ryan *et al.*, 2016). The Running Fisher algorithm provides an assessment of the statistical significance of the correlation of the overlapping genes between the biomarker and the gene lists. A complete description of the Running Fisher test has been published (Kupersmidt *et al.*, 2010). The results were exported, and each p-value was converted to a $-\log(\text{p-value})$. Thresholds for significance were set at $-\log(\text{p-value}) \geq 4$ for activation or ≤ -4 for suppression based on prior studies using this threshold (Matteo *et al.*, 2023).

4.3.8 Transcriptomic point of departure analysis

BMC modeling was done as described previously (Matteo *et al.*, 2023) in BMDEpress3 (<https://github.com/auerbachs/BMDEpress-3/releases>). A Williams trend test ($p < 0.05$) and absolute fold change ≥ 1.5 was applied to filter probes. Best fit models from polynomial 2°, linear, power (power term constrained to ≥ 1), and exponential models (degrees 3 and 5) were selected for each probe based on the lowest Akaike Information Criterion. A benchmark response of 1 standard deviation from controls was selected. Probes with the following criteria were removed: 1) having a BMC greater than the highest concentration used in the analysis after excluding cytotoxic or non-soluble concentrations; 2) mapping to more than one gene; 3) having a model fit p-value < 0.1 determined by a likelihood ratio test; and 4) having a BMC upper (BMCU) to BMC lower (BMCL) ratio > 40 . Probes that met all the BMC filtering criteria were converted to their corresponding Entrez Identifiers. Data were further analyzed using R statistical software (version 4.3.2). Gene accumulation plots were generated by rank ordering genes fitting BMCs (lowest to highest) and plotting the BMC of each gene on the x-axis and rank along the y-axis.

We derived two global tPODs to assess general toxicity:

- a) 25th rank-ordered gene BMC: Genes fitting BMCs were rank ordered from lowest to highest BMC and the BMC for the 25th gene was selected as a tPOD to highlight early perturbations to the transcriptome. This approach has been described in detail (Reardon *et al.*, 2023). Confidence intervals were estimated using a

parametric bootstrap where the residuals were assumed to be normally distributed as previously described (Matteo *et al.*, 2025).

4. Lowest pathway 5th percentile BMC: We derived the lowest pathway tPOD by aligning genes and their associated BMCL/BMC/BMCU values to the gene sets in Reactome Pathways database (<https://reactome.org/>) or the Kyoto Encyclopedia of Genes and Genomes (KEGG; <https://www.genome.jp/kegg/>). This approach is based on recommendations by an expert panel at the US National Toxicology Program (National Toxicology Program, 2018). Gene sets that contained at least three genes fitting BMCs (genes that pass all criteria in the analysis), were at least 5% populated (based on total annotated gene number), and contained at least 40 genes in the whole gene set were selected. The 5th percentile BMC of the selected gene sets were selected as tPODs. We applied the 5th percentile BMC as a more conservative approach compared to the bootstrap median. This is based on other studies investigating data-distribution approaches including the 5th percentile BMC (Rowan-Carroll *et al.*, 2021; Bourdon *et al.*, 2013; Reardon *et al.*, 2023) to increase the sensitivity in the calculation of the BMC and to be comparable to the approach used previously for ToxCast data (Paul Friedman *et al.*, 2020).

4.3.9 Average donor transcriptomic point of departure

We used a parametric bootstrap approach to calculate an average 25th gene BMC for each chemical across all donors. For each gene, the standard deviation of the log-transformed BMC was approximated from the profile likelihood confidence interval. BMC

values were assumed to follow a log-normal distribution. For each donor, 10,000 random samples were drawn from the bootstrap distributions. To account for additional uncertainty, we multiplied the approximated standard deviation by $\sqrt{df/rchisq(df)}$ for each simulation run and for each gene, where df is the degrees of freedom and $rchisq(df)$ is a value randomly drawn from a chi-square distribution with df degrees of freedom, following the approach described by Moerbeek et al. (2004). Bias correction was applied by adjusting the 25th gene BMC based on the median of the bootstrap distribution. Confidence intervals (95%) were derived from the 2.5th and 97.5th percentiles of the bias-corrected bootstrap samples. The average 25th gene BMC was calculated as the median of the exponentiated mean log-transformed values across bootstrap samples, with corresponding 95% confidence intervals derived from the 2.5th and 97.5th percentiles of the sampled distribution.

4.3.10 Ingenuity Pathway Analysis

Ingenuity Pathway Analysis (IPA; QIAGEN, Redwood City, CA, USA) was used to identify changes to upstream regulators, canonical pathways, and disease & functions. For each chemical tested, we imported Excel files into IPA containing gene IDs (Gene Symbol and Entrez ID) for the genes that fit a BMC (i.e., passed all filtering criteria), as well as their Williams trend test p-values from the BMC analysis, and the largest fold-changes of the gene relative to solvent controls (exported filtered data from BMDExpress3). This approach allowed us to analyze for the enrichment of genes

showing robust concentration-responses to the exposures. IPA Core Analysis with a gene expression threshold of absolute fold change ≥ 1.5 and false discovery rate (FDR) adjusted p-value ≤ 0.05 was used with the direct and indirect relationship settings based on experimental and highly predicted data. Statistical significance of the overlap (FDR-adjusted p-value ≤ 0.05) between the data set and known targets of upstream regulators in IPA were calculated using Fisher's Exact tests. The z-score was calculated using Fisher's Exact Test based on the expected relationship for directions between upstream regulators and target genes and those observed in the data set. A z-score of > 2 (activated) or < -2 (inhibited) was considered statistically significant.

4.4 Results

4.4.1 Only the highest concentration of BPAF reduced cell viability in HMECs

We used the CellTiter-Blue Cell Viability assay to assess cytotoxicity after 48 h of chemical exposure. The highest exposure concentration of BPAF (50 μM) caused a significant decrease in cell viability for all three donors (Supplementary Figure 4.1) and was subsequently removed from the analysis. No other condition affected cell viability.

4.4.2 Most BPA alternatives induced broad transcriptional changes in donors

The median read depth was approximately 2.6 million reads per sample with a median range of 77.2% mapped reads. Of 636 experimental samples, 21 experimental samples were lost to QC (Supplementary File 3.1). The total number of samples per condition is listed in Supplementary Table 4.2. The 36 reference RNA and lysis buffer control samples produced the expected results, showing high reproducibility across plates and low signal, respectively (data not shown).

We applied BMC modeling to identify the genes that produced robust concentration responses for each donor. Table 2 summarizes the total number of genes fitting BMCs for each chemical. BPA and most of its alternatives were transcriptionally active across all three donors (Supplementary Figure 4.2). Chemicals tested generally had 200 - 600 genes fitting BMCs. BPAF donor 2 (1911), TMBPF donor 2 (1339), and BPAF donor 3 (667) had the most genes fitting BMCs while BP3 donor 3 (286), BPS donor 2 (270), and BPC donor 2 (239) had the least. Donor 1 did not have any genes fitting BMCs for BPAF.

We used a gene accumulation plot to compare the first 100 genes fitting BMCs for each chemical and donor and ranked from most (left on the graph) to least potent (right) based on the 25th gene BMC (horizontal line; Figure 4.2). BPAF was the most potent chemical tested overall, followed by TMBPF, BPC, BPS, and BPA. There were no significant differences in toxicological potency between the three donors included in this study.

4.4.3 Transcriptomic biomarker analysis reveals cellular stress response and no ER α activation

We used a 50 gene transcriptomic ER α biomarker (Corton, Matteo, *et al.*, 2024) to test the ability of BPA alternative chemicals to activate/inhibit the ER α . Supplementary File 3.2 provides a summary of all stress response biomarkers activated and inhibited. No chemical tested activated or inhibited the 50 gene ER α biomarker.

Transcriptomic biomarker analysis was also used to explore general cellular stress responses following exposure to BPA alternative chemicals. The biomarkers were specific to: (a) NRF2 (Rooney *et al.*, 2020); (b) HSF1 (Cervantes and Corton, 2021); (c) MTF1 (Jackson *et al.*, 2020); (d) TGx-DDI (Li *et al.*, 2019); (e) TGx-HDACi (Cho *et al.*, 2021); (f) NF-kB (Korunes *et al.*, 2022); (g) HIF1 α (Corton *et al.*, 2025) (h) CP (Corton, Ledbetter, *et al.*, 2024); (i) XBP1 (unpublished data).

Several of the chemicals tested activated the HIF1 α biomarker at the highest exposure concentration tested: BPA (donor 1, 50 μ M), BPAF (donor 3, 10 μ M), BPC (donor 1 – 3, 50 μ M), TMBPF (donor 3, 50 μ M). Many of these conditions also activated the MTF1 biomarker: BPAF (donor 3, 10 μ M), BPC (donor 1 – 3, 50 μ M), TMBPF (donor 2, 0.1 μ M ; donor 3, 50 μ M). Two conditions also activated the NF-kB biomarker (BPC donor 2 – 3, 50 μ M) and one activated both the NRF2 and TGX-DDI biomarkers (BPC donor 3, 50 μ M). Together, these results suggest that some BPA alternatives can activate

cellular stress response pathways at high concentrations, though activation patterns varied by chemical and donor.

4.4.4 BPA alternatives had similar toxicological potencies in HMECs

To assess general toxicity and evaluate potency differences in donor/chemical combinations, we produced two tPODs: (a) the 25th rank-ordered gene BMC; and (b) the 5th percentile of the lowest pathway BMC. We first derived donor-specific general toxicity tPODs to identify the concentrations that produce an initial ‘burst’ in transcriptomic activity (referred to as the concerted molecular response —where the curve becomes steep; Johnson et al., 2022).

The 25th gene BMC filters out genes in the noise range to identify an initial concerted molecular response (horizontal line; Figure 4.2). Confidence intervals were derived using parametric bootstrap modelling. Most chemicals had a 25th gene BMC between 2 – 8 μM (Figure 4.3, Table 4.2). BPAF had the lowest donor average 25th gene BMC (0.73 μM), followed by TMBPF (3.67 μM), BPC (5.10 μM), BPS (7.32 μM), and BPA (7.71 μM). For each chemical tested, individual donor 25th gene BMCs were highly similar (Figure 4.3, Table 4.2). BPAF donor 3 had the lowest 25th gene BMC (0.61 μM), followed by BPAF donor 2 (0.88 μM), and TMBPF donor 2 (2.00 μM). BPA donor 3 (7.38 μM), BPS donor 2 (8.20 μM), and BPA donor 2 (8.61 μM) had the highest 25th gene BMCs. These data highlight that BPAF was the most potent chemical tested overall, and that all BPA

alternatives had 25th gene BMCs within an order of magnitude of each other and were highly similar between donors.

The lowest 5th percentile pathway BMC is an approach derived from recommendations published by the US National Toxicology Program (National Toxicology Program, 2018) that is more widely used than the 25th gene. Most chemicals produced a lowest pathway 5th percentile BMC within one order of magnitude of each other (1 – 9 μM ; Figure 4.3, Table 4.2). Compared to the 25th gene BMC, the 5th percentile pathway BMC varied more between donors for each chemical, but most were within an order of magnitude of each other (Figure 4.3, Table 4.2). BPAF donor 3 (0.67 μM), TMBPF donor 2 (0.80 μM), and BPS donor 2 (1.10 μM) had the lowest 5th percentile pathway BMCs. BPS donor 1 (7.53 μM), BPA donor 3 (8.54 μM), and BPS donor 3 (25.9 μM) had the highest. In sum, most BPA alternatives tested had pathway BMCs within an order of magnitude of each other and did not differ significantly between donors.

4.4.5 BPA alternatives oppositely enriched gene sets associated with cancer related outcomes

To identify biological mechanisms altered by chemical exposure, we used the genes fitting BMCs, their maximum fold changes, and their p values, in an IPA canonical pathway and upstream regulator analysis as described previously (Matteo *et al.*, 2023, 2025). A full list of canonical pathways and upstream regulators perturbed by chemical exposure is available in Supplementary File 3.3.

Several canonical pathways related to cancer, inflammation, and transcription were enriched in response to BPA alternatives, with distinct activation or inhibition patterns across chemicals and donors. The top 10 canonical pathways are shown in Figure 4.4A. These pathways are associated with cancer (molecular mechanisms of cancer), protein processing (protein ubiquitination pathway; neddylation), transcription (chromatin organization; generic transcription pathway; processing of capped intron-containing pre-mRNA), inflammation (class I MHC mediated antigen processing and presentation; neutrophil degranulation; S100 family signaling pathway), and cell cycle and survival (cell cycle checkpoints). BPAF donor 2 and 3 were predicted to significantly inhibit most pathways. TMBPF donor 2 had similar predictions to BPAF, while TMBPF donor 1 and 3 were largely opposite to BPAF but mostly non-significant. BPC was predicted to activate most pathways for all three donors. BPS, BPA, and E2 had similar expression profiles as BPC, but most did not reach significance. No biological pathways associated with estrogen receptor signalling were significantly enriched by chemical exposure. Together, these data point towards opposing gene expression profiles between the BPA alternative chemicals tested, with one group (BPAF and TMBPF) activating a broad range of biological pathways, and the other (BPC, BPS, and BPA) inhibiting the same gene sets.

Upstream regulator analysis revealed opposing regulatory patterns between BPA alternatives, particularly between BPAF and BPC. The top 10 upstream regulators are shown in Figure 4.4B. Like the canonical pathways, chemical exposure altered genes associated with upstream regulators in cancer related pathways (HIF1 α), protein

processing (RNF187), transcription (NUPR1, ATF4, NF- κ B, RELA), inflammation (TNF, IFNG), and cell cycle and survival (PTPRR). In general, BPAF and BPC had mostly opposing gene expression patterns across all donors. For example, BPAF was predicted to inhibit TNF, ATFF, INFG, NF- κ B, and RELA; whereas, BPC was predicted to activate these upstream regulators. Interestingly, the top concentration for BPC also activated the HIF1 α and NF- κ B transcriptomic biomarkers for most donors. TMBPF donor 1 and 2, as well BPA and BPS donor 1, had similar profiles for upstream regulator enrichment for BPAF, while BPS and BPA donor 2 and 3 were more qualitatively similar to BPC, though few predictions reached significance. Genes fitting BMCs following E2 exposures in HMECs were not significantly enriched for any upstream regulator gene sets. No chemical was predicted to affect upstream regulators associated with the ER. Overall, these data highlight that BPAF and TMBPF largely inhibited enriched gene sets across most donors, while BPC, BPS, and BPA mostly activated those same gene sets.

4.5 Discussion

Humans are exposed to endocrine disruptors like BPA and its alternatives through contact with consumer products and their environment. Most *in vitro* chemical screening accomplished to date has utilized immortalized cell lines, which differ significantly from primary tissues. To our knowledge, this is the first HTTr study to compare toxicological potency and mechanism of action of BPA alternative chemicals in human primary mammary epithelial cells derived from different human donors. We previously identified

BPAF as the most estrogenic and potent chemical tested overall, followed by BPA (Matteo *et al.*, 2023, 2025). We selected some of the most potent BPA alternatives from our previous HTTr studies in breast cancer cells (Matteo *et al.*, 2023, 2025) and exposed HMECs to these chemicals using a similar experimental design. We did BMC modelling to derive general toxicity tPODs and identified BPAF as the most potent chemical tested overall and BPA as the least potent. Transcriptomic biomarkers analysis suggested that none of the chemicals are ER α active chemicals. However, BPAF, BPC, and TMBPF activated stress response pathways at high exposure concentrations. Gene set enrichment analysis revealed two groups of BPA alternatives with largely opposing activity associated with cancer related pathways. For all donors tested, most chemicals had similar toxicological potencies and gene set enrichment. We note that due to availability, HMECs were sourced from the biological donors of similar ethnic backgrounds that did not differ significantly in height or weight. Therefore, the generalizability of these data is limited and would benefit from validation using more diverse samples. Together, these data suggest that most BPA alternatives tested have similar toxicological potencies in HMECs and that some have opposing, non-ER α mediated gene expression patterns relative to each other *in vitro*.

BPA and the four alternative chemicals tested had similar toxicological potencies in HMECs. For all chemicals tested, all 25th gene BMCs and most pathway BMCs derived were within an order of magnitude of each other and did not differ significantly between biological donors. BPAF, a fluorinated derivative detected in humans (Niu *et al.*, 2021; H.

Wang *et al.*, 2023) and in products like processed foods (Lucarini *et al.*, 2023; Liao and Kannan, 2013) and dairy milk (Wan *et al.*, 2023), was the most potent chemical tested overall in HMECs. This aligns with previous HTTr data that rank BPAF as the most potent alternative in MCF-7 breast cancer cells (Mesnage *et al.*, 2017; Matteo *et al.*, 2023, 2025). In general, most chemicals tested in HMECs had general toxicity tPODs an order of magnitude higher than our previous HTTr study in MCF-7 cells (Matteo *et al.*, 2025). However, aside from BPAF, chemical potency rankings differed significantly between breast epithelial cells of primary and cancerous origin. For example, BPA, previously reported as one of the most toxicologically potent in its class in breast cancer cells (Beal *et al.*, 2024; Matteo *et al.*, 2023, 2025), was the least potent chemical tested overall in HMECs. Interestingly, some BPA alternatives, like BPAF and BPC, had more than twice the number of genes fitting BMCs in HMECs than MCF-7 cells (Matteo *et al.*, 2023, 2025). Thus, these chemicals are highly transcriptionally active in primary mammary cells. Overall, this HTTr data set adds to the weight of evidence that BPAF is among the most toxicologically potent BPA alternatives *in vitro* and highlight differences in potencies between breast cells of primary and cancerous origin.

Transcriptomic biomarker analysis identified cellular stress response pathways activated by exposure to BPA alternative chemicals. Specifically, high exposure concentrations of the most potent BPA alternatives tested (BPAF, TMBPF, and BPC; 50 μ M) activated the HIF1 α and MTF1 biomarker in at least one donor, while BPC activated the NF-kB biomarker in two donors. Given that HIF1 α (Generali *et al.*, 2006), MTF1 (Shi *et*

et al., 2010), and NF- κ B (Biswas *et al.*, 2004) are overexpressed in some breast tumors, these data suggest that high exposure concentrations of BPA alternatives may cause a cancerous phenotype in HMECs. Importantly, none of the BPA alternative chemicals tested activated the 46 gene (Ryan *et al.*, 2016) or 50 gene ER α biomarker (Corton, Matteo, *et al.*, 2024) in HMECs. This is likely because only a subpopulation of HMECs express the ER α (Rosenbluth *et al.*, 2020; Petersen *et al.*, 1987), which can be further affected by culturing conditions (Fridriksdottir *et al.*, 2015). Most of these BPA alternative chemicals (BPA, BPAF, BPC, and BPS) are ER α active in breast cancer cells that overexpress the ER α (Mesnage *et al.*, 2017; Matteo *et al.*, 2023, 2025). In sum, transcriptomic biomarker analysis suggests no BPA alternatives are ER α active in HMECs and that some may activate cellular stress response pathways associated with breast cancer at high exposure concentrations. However, the transcriptomic biomarkers applied in this study were developed using gene expression data from immortalized cells from various tissues and some of cancerous origin. Therefore, these results should be interpreted with caution and warrant further validation.

Gene set enrichment revealed two groups of BPA alternative chemicals with largely opposing profiles in HMECs. The data suggest that BPAF and TMBPF mostly inhibit similar upstream regulators and biological pathways associated with cancer and cellular processes like transcription, protein processing, and inflammation. Interestingly, BPAF significantly inhibited almost all enriched gene sets in two donors but failed to cause a robust transcriptomic response in one donor. One possible explanation is inter-

individual variability in transcriptional sensitivity to BPAF; the transcriptomic effects in this donor may fall below the threshold of detection at the concentrations tested (i.e., a weak responder). By contrast, BPA, BPS, and in particular, BPC, were generally predicted to activate these same gene sets. In line with transcriptomic biomarker analysis, BPC was predicted to activate HIF1 α and NF- κ B, whose expression is associated with breast cancer (Guo *et al.*, 2024; Zhi *et al.*, 2024), and interact to support tumor survival in some cancers (Zhao *et al.*, 2021). Further, other enriched pathways like transcription (Pellacani *et al.*, 2019), protein ubiquitination (Deng *et al.*, 2020), neddylation (S. Zhang *et al.*, 2024), and inflammation (Danforth, 2021), are dysregulated in breast cancer. These data suggest that exposure to some BPA alternatives like BPC may promote cancer-related outcomes in HMECs. This finding broadly aligns with previous reports that BPA exposure perturbs pathways associated with cancer in HMECs (Weng *et al.*, 2010; Qin *et al.*, 2012). Compared to previous HTTr experiments in breast cancer cells (Matteo *et al.*, 2023, 2025), where most BPA alternatives enriched gene sets associated with the ER and cellular proliferation, a broad range of non-ER related cellular pathways were affected by BPA alternatives in HMECs. These findings highlight two groups of BPA alternatives with mostly opposing gene set enrichment of cancer related pathways in HMECs. More research is warranted to determine possible modes of action of BPA alternative chemicals in HMECs.

In summary, we tested BPA and four alternative chemicals in HMECs sourced from three human donors. We identified BPAF as the most potent chemical tested overall in

HMECs in line with previous data from breast cancer cells. BPA was the least potent chemical tested. No chemicals were ER α active in HMECs and some BPA alternatives activated cellular stress response biomarkers at high exposure concentrations. Gene set enrichment revealed two groups of BPA alternatives with mostly gene expression profiles for cancer related outcomes. Taken together, these data suggest that the BPA alternatives tested have similar toxicological potencies and may have opposing mechanisms of action relative to each other in HMECs.

4.6 Tables and Figures

Table 4.1: List of chemicals and suppliers

Chemical	Acronym	CAS Number	Supplier
Bisphenol A	BPA	80-05-7	Sigma-Aldrich Inc (St. Louis, MO, USA)
Bisphenol AF	BPAF	1478-61-1	
Bisphenol S	BPS	80-09-1	
Estradiol	E2	50-28-2	
Dimethyl sulfoxide	DMSO	67-68-5	
Tetramethyl bisphenol F	TMBPF	5384-21-4	
Bisphenol C	BPC	79-97-0	Accustandard (New Haven, CT, USA)

Table 4.2: Comparison of transcriptomic points of departure in human mammary epithelial cells from three biological donors exposed to BPA and four alternatives.

Chemical	Donor	# Genes fitting BMCs	25th Gene BMC (μM)	25th Gene BMCL – BMCU (μM)	Lowest Pathway 5th Percentile BMC (μM)
BPAF	average		0.73	0.57 - 0.90	
BPAF	1	0	N/A	N/A	N/A
BPAF	2	1911	0.88	0.17 - 3.57	1.99
BPAF	3	1280	0.61	0.14 - 0.62	0.65
TMBPF	average		3.67	3.20 - 4.26	
TMBPF	1	651	4.47	3.08 - 8.08	3.35
TMBPF	2	1339	2.01	0.47 - 8.32	0.80
TMBPF	3	671	5.50	3.55 - 12.2	3.45
BPC	average		5.10	4.25 - 5.93	
BPC	1	1743	4.49	3.03 - 8.41	4.97
BPC	2	1508	5.28	1.07 - 8.13	5.73
BPC	3	1243	5.82	3.59 - 33.1	6.92
BPS	average		7.32	6.46 - 8.21	
BPS	1	450	6.66	4.02 - 20.8	7.53
BPS	2	270	8.20	4.23 - 43.1	1.10
BPS	3	306	7.23	4.20 - 26.8	25.9
BPA	average		7.71	6.82 - 8.70	
BPA	1	505	7.23	4.26 - 21.1	7.05
BPA	2	552	8.61	5.24 - 9.96	4.56
BPA	3	371	7.38	4.30 - 24.2	8.54

Human mammary epithelial cells ($n = 2 - 4$ per concentration) from three biological donors exposed to bisphenol A (BPA) and four alternatives at a range of concentrations ($0.001 - 50 \mu\text{M}$) for 48 h. tPODs were derived by prefiltering data using the Williams trend test ($p < 0.05$) and an absolute fold-change filter of ≥ 1.5 . Data were also postfiltered with the following settings in BMDEExpress v3: Best BMD/BMDL < 20 , Best BMDU/BMDL < 40 , and Best fitPvalue ≥ 0.1 . tPODs (shown in μM) representing the 25th rank ordered gene benchmark concentration (BMC) and the 5th percentile gene BMC for the lowest pathway are shown. Most chemicals are shown in decreasing order of potency based on tPODs from the 25th gene BMC. Note that BPAF Donor 1 did not have any genes fitting BMCs.

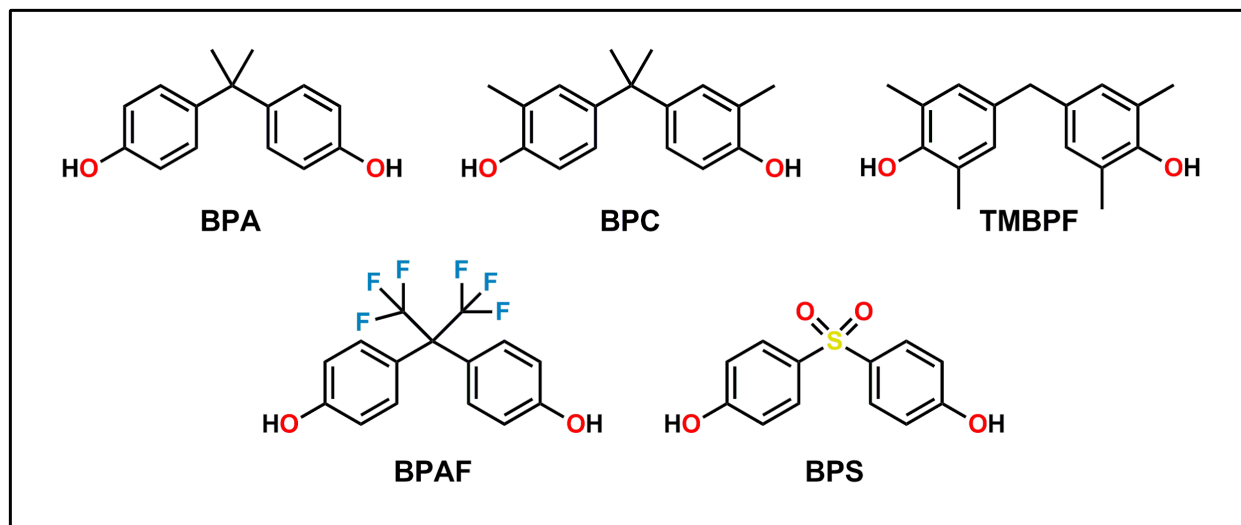


Figure 4.1: Chemical structures of bisphenol A (BPA) alternative chemicals.

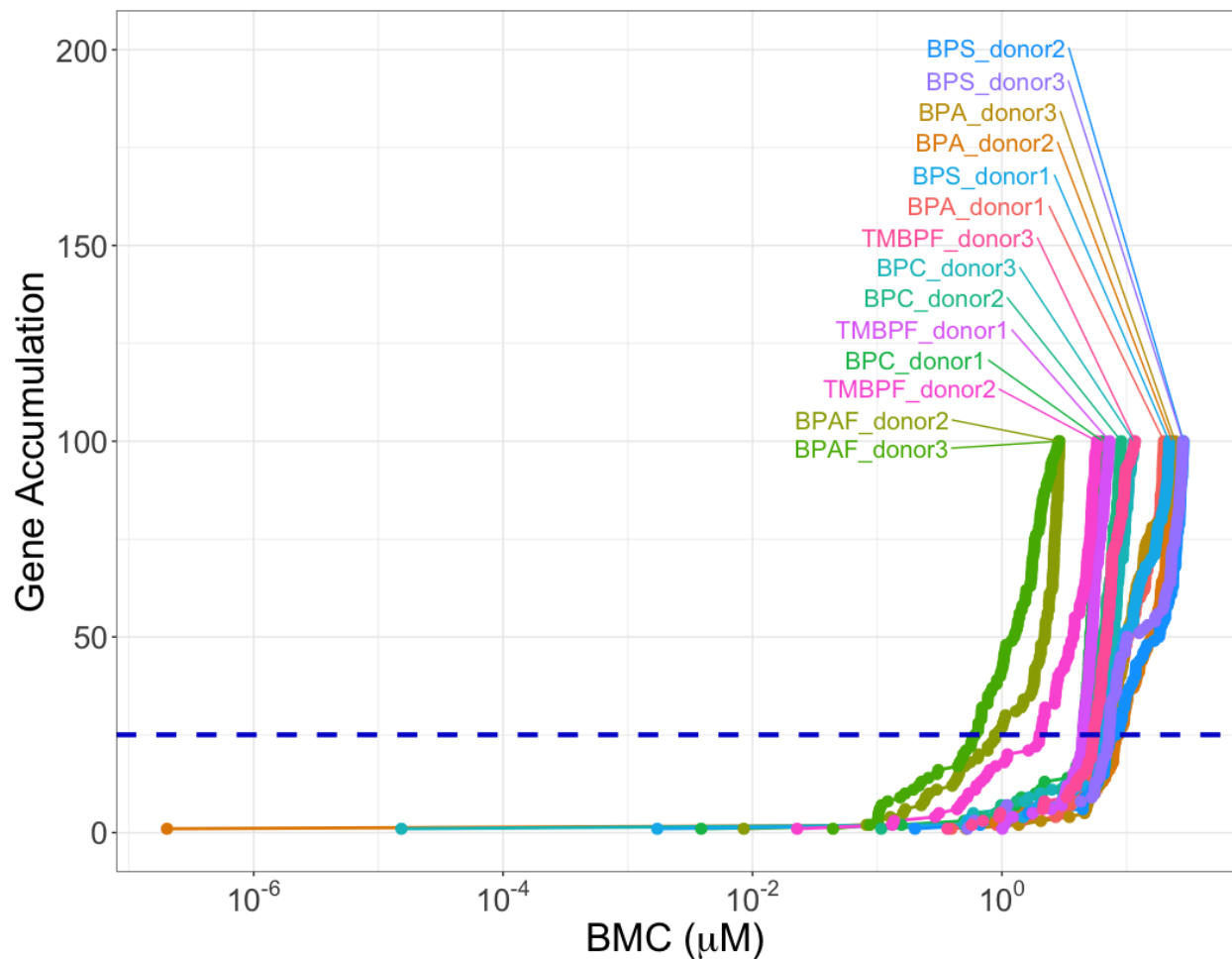


Figure 4.2: Gene accumulation plot of human mammary epithelial cells ($n = 2 - 4$ per concentration) for the first 100 genes fitting benchmark concentrations (BMCs) from three biological donors exposed to bisphenol A (BPA) and four alternatives at a range of concentrations ($0.001 - 50 \mu\text{M}$) for 48 h. The blue dashed line indicates the 25th ranked gene. Data were prefiltered using the Williams trend test ($p < 0.05$) and an absolute fold-change filter of ≥ 1.5 , and post filtered with the following settings in BMDExpress v3: Best BMD/BMDL < 20 , Best BMDU/BMDL < 40 , and Best fitPvalue ≥ 0.1 .

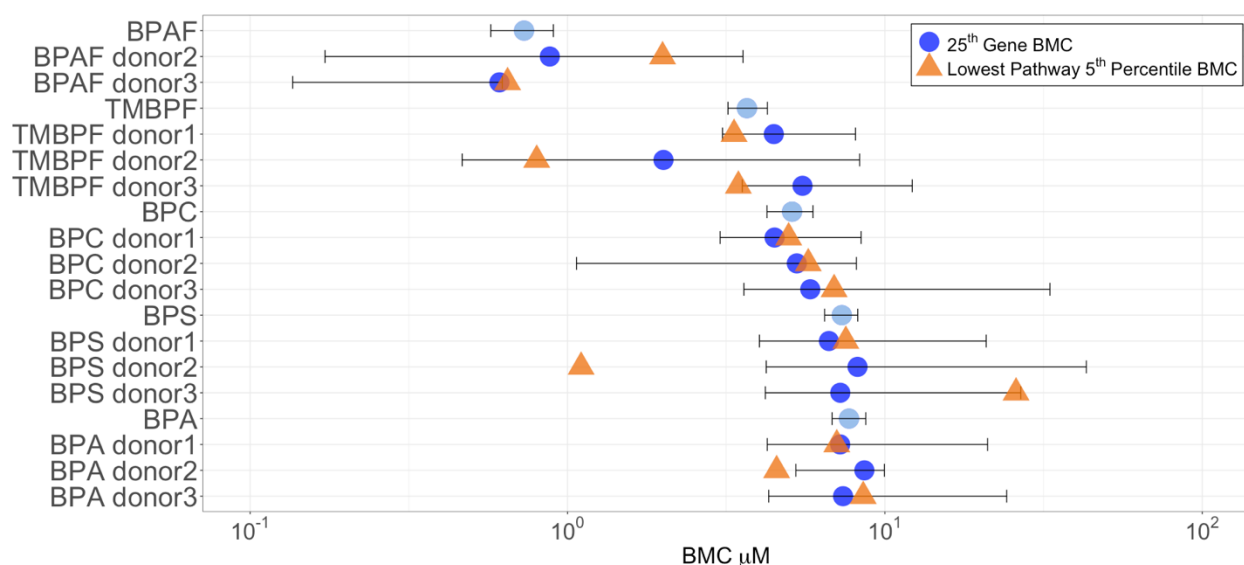
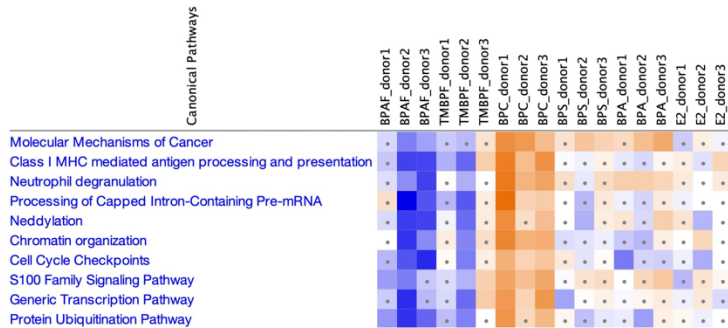


Figure 4.3: Comparison of transcriptomic points of departure (tPODs) derived from whole transcriptome analysis of human mammary epithelial cells ($n = 2 - 4$ per concentration) from three biological donors exposed to bisphenol A (BPA) and four alternatives at a range of concentrations ($0.001 - 50 \mu\text{M}$) for 48 h. tPODs were derived by prefiltering data using the Williams trend test ($p < 0.05$) and an absolute fold change filter of ≥ 1.5 . Data were also postfiltered with the following settings in BMDExpress v3: Best BMD/BMDL < 20 , Best BMDU/BMDL < 40 , and Best fitPvalue ≥ 0.1 . tPODs (shown in μM) representing the 25th rank ordered gene benchmark concentration (BMC; blue circles) for individual donors, the average 25th gene BMC calculated using parametric bootstrap approach (light blue circles), and the 5th percentile gene BMC for the lowest pathway (orange triangles) are shown. BMCL and BMCU are used for the lower and upper bounds of the 25th gene BMC, respectively. Most chemicals are shown in decreasing order of potency based on tPODs from the 25th gene BMC. Note that BPAF Donor 1 did not have any genes fitting BMCs.

A

Activation z-score
-6.835 7.462

B

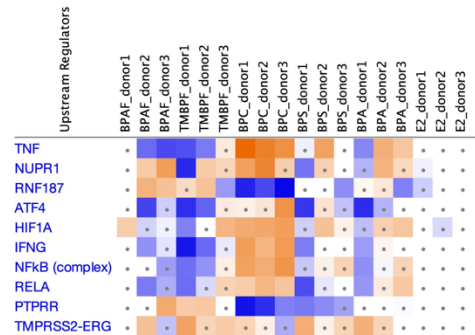
Activation z-score
-5.145 6.051

Figure 4.4: Ingenuity Pathway Analysis of all genes fitting benchmark concentration models in human mammary epithelial cells ($n = 2 - 4$ per concentration) from three biological donors exposed to bisphenol A (BPA) and four alternatives at a range of concentrations ($0.001 - 50 \mu\text{M}$) and 17β -estradiol (E2; 1 and 10 nM) for 48 h. A list of genes fitting models, the Williams trend test p value, and the maximum fold-change were imported into IPA for this comparison. Filters were set to Z-score ≥ 2.0 and p-value < 0.05 . Orange denotes predicted activation and blue predicted inactivation (dots signify Z-score < 2). (A) Top 10 canonical pathways. (B) Top 10 upstream regulators.

Chapter 5: Final Discussion and Conclusion

5.1 Summary of Outcomes

The overall goal of my thesis was to compare the toxicological potency and identify potential hazards of BPA alternative chemicals *in vitro*. Chapter 2 established an HTTr framework to screen BPA to alternatives chemicals for toxicological potency and mechanisms of action in breast cancer cells. For Chapter 3, a concentration addition model was integrated to the HTTr workflow to test mixtures of BPA alternatives for additive effects *in vitro*. In Chapter 4, BPA and select alternative chemicals were tested in a human-relevant primary mammary cell model. Together, my thesis outlines an approach to screen data-poor BPA alternative chemicals and advances the use of HTTr for regulatory toxicology. The findings and impacts for each chapter are described below.

Chapter 2

Objectives:

Objective 1: Apply HTTr to explore mechanisms of action and derive points of departures for potency comparison of BPA and its alternative chemicals in breast cancer cells.

Objective 1a: Use tPODs to compare the toxicological potency of BPA alternative chemicals.

Objective 1b: Use Ingenuity Pathway Analysis to compare upstream regulators and canonical pathways activated by BPA alternative chemical exposure.

Outcomes and Conclusions:

I present a framework that outlines an *in vitro* HTTr approach to screen BPA alternative chemicals to inform regulatory toxicology. BPA and 15 alternative chemicals were tested in MCF-7 cells and underwent global transcriptomic profiling. Most BPA alternatives caused robust transcriptional activity in breast cancer cells. Transcriptomic biomarker analysis identified BPA and eight alternatives as ER α active; two of which had no previous ER α activity data. Both global and ER α -specific tPODs ranked BPAF as the most potent chemical tested overall, followed by BPA, and BPC. Most BPA alternatives activated the ER α at similar concentrations as BPA. Gene set enrichment highlighted that BPA and most alternatives had similar mechanisms of action via the ER α and

dysregulated cellular proliferation and caused the expression of genes associated with breast cancer in MCF-7 cells. This HTTr data set supports that most BPA alternatives activate the ER α , have comparable toxicological potencies as BPA, and have similar modes of action in MCF-7 cells. This study provides a baseline data set and framework for how novel and data-poor BPA replacements can be tested moving forward.

Chapter 3

Objectives:

Objective 2: Apply HTTr to perform concentration-response modeling to mixtures of BPA alternative chemicals to determine if these chemicals exert additive effects *in vitro*.

Objective 2a: Use tPODs to compare potency of mixtures of BPA alternatives to their individual components.

Objective 2b: Use Ingenuity Pathway Analysis to compare upstream regulators and canonical pathways activated by exposure to mixtures of BPA alternative chemicals.

Outcomes and Conclusions:

I tested whether mixtures of BPA alternatives chemicals behave additively in breast cancer cells. MCF-7 cells were exposed to mixtures of BPA alternative chemicals and their individual components following a similar approach as outlined in Chapter 2. Concentration addition modelling identified that BPA alternatives tested behaved additively in MCF-7 cells. Potency ranking of individual BPA alternatives were consistent with Chapter 2, with BPAF as the most potent individual chemical tested overall, followed by BPA, and BPC. In line with additive effects, all mixtures tested activated the ER α and enriched similar upstream regulators and pathways as their individual components. In

sum, these data demonstrate the utility of HTTr to screen mixtures of BPA alternatives chemicals for additive effects *in vitro*. Further, I add to the weight of evidence that BPAF is toxicologically potent BPA alternative and highlight the replicability of tPOD potency rankings between HTTr data sets.

Chapter 4

Objectives:

Objective 3: Apply HTTr to determine potency and mechanism of action of BPA and its alternatives in primary human mammary epithelial cells.

Objective 3a: Use tPODs to compare potency of BPA alternative chemicals previously screen in breast cancer cells.

Objective 3b: Use Ingenuity Pathway Analysis to compare upstream regulators and canonical pathways activated by BPA alternative chemical exposure.

Outcomes and Conclusions:

BPA and select alternatives previously screened in MCF-7 cells were tested in HMECs from multiple donors. Global transcriptomic profiling demonstrated that BPAF and TMBPF caused more gene expression in HMECs than MCF-7 cells. No chemicals activated the ER α in HMECs and biomarkers associated with breast cancer were activated by the highest exposure concentrations of BPAF and BPC. All chemicals had similar toxicological potencies in HMECs from all three donors, while tPODs were generally an order of magnitude higher than those previously derived in MCF-7 cells. Gene set enrichment analysis revealed that BPA alternatives had opposing mechanisms of action, with one group (BPAF and TMBPF) inhibiting and the other (BPA, BPS, and

BPC) activating gene sets associated with dysregulated cellular proliferation and breast cancer-related phenotypes. Overall, this study highlights differences in toxicological potencies and mechanisms of action between mammary epithelial cells of primary and cancerous origin, and that exposure to some BPA alternatives like BPC may cause breast cancer-related outcomes in HMECs.

5.2 Contributions to Scientific Knowledge

Together, my thesis addresses knowledge gaps for several data-poor BPA alternative chemicals to which humans are exposed. The studies included herein provide a framework to use gene expression data to assess data-poor chemicals for estrogenic effects, rank by toxicological potency, and identify potential mechanisms of action *in vitro*. Overall, the data presented are critical in supporting the implementation of NAMs to inform chemical risk assessments and meet the growing demands to screen new and data-poor chemicals.

Chapter 2 established an *in vitro* HTTr approach to screen individual BPA alternative chemicals. This study demonstrated the utility of applying transcriptomic biomarkers to identify ER α active chemicals from among data-poor BPA alternatives. Transcriptomic profiling revealed that ER α active BPA alternatives caused robust and similar gene expression profiles in MCF-7 cells. In addition, this study is the first application of an ER α -specific tPOD in an HTTr data set. Both global and ER α -specific measures of potency were highly concordant, endorsing the use of multiple approaches to tPOD derivation for chemical potency ranking. Overall, these data support grouping of ER α active BPA alternatives for chemical read across and were used to inform a regulatory assessment of BPA alternative chemicals by Health Canada as part of the CMP.

In Chapter 3, the chemical screening framework was applied to mixtures of BPA alternative chemicals. Concentration addition modeling accurately predicted additive effects of equimolar and complex mixtures of BPA alternatives in MCF-7 cells. This model was previously used in another HTTr study to accurately predict additive effects of per- and polyfluoroalkyl substances (PFAS) in liver cells (Addicks *et al.*, 2023), demonstrating the transferability of methods between different HTTr studies using similar chemical screening frameworks. Further, mixtures of BPA alternatives activated ER α biomarkers and enriched similar gene sets as their individual components; thus, additive effects were likely mediated by the ER α in MCF-7 cells. Potency rankings were highly similar with those derived in Chapter 2, highlighting the replicability of tPODs across HTTr studies. In sum, this study adds to the weight of evidence that ER α active BPA alternative behave additively when present as mixtures in MCF-7 cells and supports the use of HTTr to screen chemical mixtures for additive effects *in vitro*.

For Chapter 4, BPA and select alternative chemicals were screened in a human-relevant model. This is the first HTTr data set to compare global transcriptomic responses of primary breast cells from different human donors to BPA alternative chemicals. While the number and diversity of donors included in this study was limited due to availability, inter-donor differences to BPA alternative chemical exposure highlight the need to incorporate multiple donors in chemical potency ranking studies. This study also showed that BPA alternatives act through non-ER α related mechanisms in HMECs. Moreover, this study demonstrated that BPA alternatives have different toxicological potencies in

HMECs compared to MCF-7 cells. Collectively, these data support that BPA alternatives act through distinct mechanisms in HMECs compared to breast cancer cells and highlight the need to incorporate primary cells in HTTr chemical screening studies.

Appendix A utilized data from Chapter 2 to develop a 50 gene transcriptomic biomarker for the ER α . This is a powerful tool that enables researchers to chemicals for ER α activity in HTTr data sets derived from MCF-7 cells. Importantly, this biomarker is not limited to BPA alternative chemicals and can be used to screen a broad range of chemicals for ER α activity.

Appendix B also incorporates data from Chapter 2 to derive administered equivalent doses for human exposure estimates of BPA alternative chemicals. This study demonstrates that the HTTr approaches described would be protective of human health based on being equally or more conservative conventional animal tests. These data are critical for building confidence in NAMs for chemical risk assessment and reducing reliance on animal testing.

5.3 Future Directions

My thesis lays the foundation for several future studies:

- 1) Identify BPA alternatives with high exposure levels in the Canadian population.**

My thesis involved the screening of 19 BPA alternatives and seven mixtures, as well as the identification of 10 ER α active alternatives in breast cancer cells. However, there are over 100 BPA alternative chemicals listed in Canada's Chemical Management Plan to date (Government of Canada, 2020). There is a lack of publicly available data on levels of BPA alternatives in the general Canadian population (e.g., urine, blood). These data are critical for informing study design by prioritization screening of BPA alternatives with high exposure levels and selecting physiologically relevant exposure concentrations. In addition, this will enable researchers to screen human-relevant mixtures of BPA alternatives that model those detected in the general population. Therefore, more biomonitoring data are warranted to prioritize BPA alternatives with high human exposures for *in vitro* chemical screening.

2) Screen more data-poor BPA alternatives individually and in mixtures

The HTTr framework established facilitates efficient screening of BPA alternatives individual and in mixtures. This approach has recently been applied to screen more data-poor BPA alternative chemicals in breast cancer cells by our lab (Geronimo Matteo, unpublished data) and others (Beal *et al.*, 2024). However, many BPA alternative chemicals remain data-poor, limiting regulatory toxicological assessment activities. Therefore, there is an urgent need to screen more BPA alternatives individually and in mixtures for toxicological potency and hazard

identification to inform regulatory toxicology. Further, application of *in vitro* to *in vivo* extrapolation (IVIVE) models should be applied to all HTTr endpoints (as described in Appendix B) to improve relevance to humans. Importantly, these data can be incorporated into a database in parallel with the DSL to identify and compare potencies of estrogen-like chemicals. These efforts will enable regulators to prioritize substances on the DSL that are potential EDCs to identify potential adverse effects and protect human health.

3) Screen BPA alternative chemicals individual and as mixtures for potential hazards in different cell lines

Evidence supports that BPA and alternatives chemicals can act through non-ER α pathways and that these effects may be masked in breast cancer cells that overexpress the ER α . For example, some non-ER α active alternatives like TMBPF are adipogenic and may act via the peroxisome proliferator-activated receptor gamma in murine adipocytes (Singh *et al.*, 2024). Further research is also needed to understand how mixtures of BPA alternative chemicals behave in different *in vitro* models. For example, some evidence suggests that mixtures of BPA alternatives are also obesogenic *in vitro* (Reina-Pérez *et al.*, 2022). Therefore, there is a need to incorporate diverse cell lines and primary tissues into chemical screening strategies for BPA alternatives to identify potential hazards in different tissues.

4) Use donors of diverse backgrounds to screen more BPA alternative chemicals

In vitro chemical screening has traditionally relied on immortalized cell lines, often of cancerous origin. However, cells harvested from donor tissues are more representative of primary tissues than immortalized cell lines, and better model human-relevant outcomes for regulatory toxicology applications. Moreover, using tissues of donors from diverse backgrounds captures individual differences to chemical exposure and increases generalizability of data to the general population. My thesis demonstrated that BPA and some alternatives have different potencies and mechanisms of action in primary breast cells compared to those of cancerous origin. Further, chemical exposure caused distinct gene expression profiles in HMECs of some donors relative to others. These data highlight the need to screen BPA alternatives in primary tissues such as HMECs and incorporate multiple donors of diverse backgrounds. Therefore, future *in vitro* HTTr chemical screening of BPA alternatives should integrate primary cell models.

5) Investigate the mechanisms of action of BPAF

BPAF is a fluorinated BPA alternative is used to manufacture fluoroelastomers rubbers that are used to protect some products from heat and environmental stressors. These products are generally labelled as chemically inert

and safe for human use. The data presented in this thesis demonstrates that BPAF is consistently the most toxicologically potent BPA alternative chemical tested in breast epithelial cells of cancerous and primary origin. Indeed, BPAF consistently activated the ER α at lower concentrations than BPA, an EDC banned in some consumer products like baby bottles in Canada and the European Union. Evidence from multiple cell lines (e.g., Mokra *et al.*, 2018; Ferreira *et al.*, 2022; Feng *et al.*, 2016) and animals (Mentor *et al.*, 2020; Moreman *et al.*, 2017; Li *et al.*, 2016) support that exposure to BPAF is associated with negative outcomes. However, while BPAF is ER α active in MCF-7 cells, its exact mechanisms in HMECs and other cell lines are unclear. Therefore, future mechanistic studies of BPA alternatives should prioritize BPAF.

6) Screen BPA alternative chemicals in breast organoids

Three-dimensional cell culture better approximates human tissues than traditional two-dimensional techniques. In particular, primary mammary cells cultured in three-dimensions (i.e., breast organoids) preserve key characteristics found in human breast tissues that are lost when cultured in two dimensions (Rosenbluth *et al.*, 2020). Techniques for culturing breast organoids from donor tissues are publicly available (Dekkers *et al.*, 2021). In addition, there is evidence that exposure to BPA and some alternatives negatively impact breast organoids

(Winkler *et al.*, 2022). However, breast organoids have not been used for HTTr chemical screening of BPA alternative chemicals, and it is unclear how mixtures of BPA alternatives affect breast organoids. Therefore, future applications of the HTTr framework presented should incorporate primary breast organoids to screen BPA alternatives.

7) Use multi-omics to gain deeper insights into the mechanisms of action of BPA alternatives

Transcriptomics, along with genomics, are the most commonly used ‘omics’ technologies for regulatory toxicological chemical screening. Evidence from various models supports that BPA, as well as some BPA alternatives, affect the lipidome (e.g., Hong *et al.*, 2022) and proteome (e.g., C. H. Kim *et al.*, 2021), (e.g., Sun *et al.*, 2023; Schaffert *et al.*, 2021). However, the multi-omics testing strategies for chemical screening of BPA alternatives has not been established. Therefore, there is a need to develop holistic approaches to chemical screening that couple multiple omics technologies. This will provide deeper insights into mechanisms than currently afforded by HTTr by associating changes in gene expression to biological molecules like lipids and/or proteins. Importantly, this will enable researchers to develop points of departure using lipid and/or protein expression data to supplement tPODs for chemical potency ranking exercises.

5.4 Concluding Remarks

The ability of BPA to act as a synthetic estrogen mimic has been known since the early 20th century. Despite this, it is widely used to manufacture commercial and industrial products to the present day. Regulations imposed by jurisdictions to limit or ban the use of BPA in certain consumer products have facilitated the proliferation of hundreds alternative chemicals, often used in products labelled “BPA-free”. Humans are exposed to several BPA alternatives, often as mixtures, and their health effects are mostly unknown. The replacement of a chemical of concern for data-poor alternatives is common problem in regulatory toxicology, and conventional animal-based toxicological screening studies operate on a slower timescale than chemical manufacturing industries. NAMs are increasingly being incorporated into regulatory toxicology strategies in Canada and jurisdictions worldwide to address these knowledge gaps and reduce reliance on research animals. The work presented herein advances the use of NAMs for regulatory toxicology by establishing an *in vitro* HTTr framework to efficiently screen BPA alternative chemicals individually and as mixtures. This approach enables the identification of chemical mechanism of action by application of transcriptomic biomarkers and gene set enrichment analysis to HTTr data sets. Measures of potency derived from gene expression data facilitate chemical toxicological comparison and derivation of human exposure safety limits. Incorporation of NAMs outlined in my thesis will enable regulatory toxicology to

accelerate screening of data-poor BPA alternative chemicals to protect public health from potential chemical harms.

References

- Acevedo, J.M. *et al.* (2025) Temporal and geographic variability of bisphenol levels in humans: A systematic review and meta-analysis of international biomonitoring data. *Environ Res*, **264**.
- Addicks, G.C. *et al.* (2023) Per- and polyfluoroalkyl substances (PFAS) in mixtures show additive effects on transcriptomic points of departure in human liver spheroids. *Toxicological Sciences*, **194**, 38–52.
- Altenburger, R. *et al.* (2000) Predictability of the toxicity of multiple chemical mixtures to *Vibrio fischeri*: Mixtures composed of similarly acting chemicals. *Environ Toxicol Chem*, **19**, 2341–2347.
- Atlas, E. *et al.* (2003) A deletion mutant of heregulin increases the sensitivity of breast cancer cells to chemotherapy without promoting tumorigenicity. *Oncogene*, **22**, 3441–3451.
- Beal, M.A. *et al.* (2024) High-throughput transcriptomics toxicity assessment of eleven data-poor bisphenol A alternatives. *Environmental Pollution*, **361**.
- Bermudez, D.S. *et al.* (2010) Modeling the interaction of binary and ternary mixtures of estradiol with bisphenol a and bisphenol AF in an in vitro estrogen-mediated transcriptional activation assay (T47D-KBluc). *Toxicological Sciences*, **116**, 477–487.
- Biswas, D.K. *et al.* (2004) NF- κ B activation in human breast cancer specimens and its role in cell proliferation and apoptosis.
- Biswas, S.K. *et al.* (2022) The Mammary Gland: Basic Structure and Molecular Signaling during Development. *Int J Mol Sci*, **23**.

- Björnsdotter, M.K. *et al.* (2017) Bisphenol A and replacements in thermal paper: A review. *Chemosphere*, **182**, 691–706.
- Bone, A. *et al.* (2016) Evaluation and comparison of bisphenol A analog activity using ToxCast data (SOT).
- Bourdon, J.A. *et al.* (2013) Gene expression profiling to identify potentially relevant disease outcomes and support human health risk assessment for carbon black nanoparticle exposure. *Toxicology*, **303**, 83–93.
- den Braver-Sewradj, S.P. *et al.* (2020) Substitution of bisphenol A: a review of the carcinogenicity, reproductive toxicity, and endocrine disruption potential of alternative substances. *Crit Rev Toxicol*, **50**, 128–147.
- Camon, C. *et al.* (2024) Exploring the effects of estrogen deficiency and aging on organismal homeostasis during menopause. *Nat Aging*, **4**, 1731–1744.
- CEPA (1999) Canadian Environmental Protection Act, 1999 Canada.
- Cervantes, P.W. and Corton, J.C. (2021) A Gene Expression Biomarker Predicts Heat Shock Factor 1 Activation in a Gene Expression Compendium. *Chem Res Toxicol*, **34**, 1721–1737.
- Chen, D. *et al.* (2016) Bisphenol Analogues Other Than BPA: Environmental Occurrence, Human Exposure, and Toxicity - A Review. *Environ Sci Technol*, **50**, 5438–5453.
- Chen, P. *et al.* (2022) Role of estrogen receptors in health and disease. *Front Endocrinol (Lausanne)*, **13**.
- Chen, S. *et al.* (2018) Fastp: An ultra-fast all-in-one FASTQ preprocessor. In, *Bioinformatics*. Oxford University Press, pp. i884–i890.
- Chen, Y. *et al.* (2022) Endocrine disrupting chemicals in the environment: Environmental sources, biological effects, remediation techniques, and perspective. *Environmental Pollution*, **310**.
- Cho, E. *et al.* (2021) Development and validation of the TGx-HDACi transcriptomic biomarker to detect histone deacetylase inhibitors in human TK6 cells. *Arch Toxicol*, **95**, 1631–1645.

- Cimmino, I. *et al.* (2020) Potential mechanisms of bisphenol a (BPA) contributing to human disease. *Int J Mol Sci*, **21**, 1–22.
- Clusan, L. *et al.* (2023) A Basic Review on Estrogen Receptor Signaling Pathways in Breast Cancer. *Int J Mol Sci*, **24**.
- Colorado-Yohar, S.M. *et al.* (2021) Concentrations of bisphenol-A in adults from the general population: A systematic review and meta-analysis. *Science of the Total Environment*, **775**.
- Corrales, J. *et al.* (2015) Global Assessment of Bisphenol A in the Environment: Review and Analysis of Its Occurrence and Bioaccumulation. *Dose Response*, **13**, 1559325815598308–1559325815598308.
- Corton, J.C., Matteo, G., *et al.* (2024) A 50-gene biomarker identifies estrogen receptor-modulating chemicals in a microarray compendium. *Chem Biol Interact*, **394**.
- Corton, J.C., Ledbetter, V., *et al.* (2024) A transcriptomic biomarker predictive of cell proliferation for use in adverse outcome pathway-informed testing and assessment. *Toxicological Sciences*.
- Corton, J.C. *et al.* (2025) Characterization of a gene expression biomarker predictive of hypoxia-inducible factor-1 modulation. *Chem Biol Interact*, 111595.
- Corton, J.C. *et al.* (2019) Identification of potential endocrine disrupting chemicals using gene expression biomarkers. *Toxicol Appl Pharmacol*, **380**, 114683.
- Crizer, D.M. *et al.* (2021) Benchmark Concentrations for Untargeted Metabolomics Versus Transcriptomics for Liver Injury Compounds in in Vitro Liver Models. *Toxicological Sciences*, **181**, 175–186.
- Crump, K.S. (1984) A new method for determining allowable daily intakes. *Toxicological Sciences*, **4**, 854–871.
- Cunha, S.C. *et al.* (2012) Determination of bisphenol A and bisphenol B in canned seafood combining QuEChERS extraction with dispersive liquid-liquid microextraction followed by gas chromatography-mass spectrometry. *Anal Bioanal Chem*, **404**, 2453–2463.

- Danforth, D.N. (2021) The role of chronic inflammation in the development of breast cancer. *Cancers (Basel)*, **13**.
- Deceuninck, Y. *et al.* (2019) Quantitative method for conjugated metabolites of bisphenol A and bisphenol S determination in food of animal origin by Ultra High Performance Liquid Chromatography-Tandem Mass Spectrometry. *J Chromatogr A*, **1601**, 232–242.
- Dekkers, J.F. *et al.* (2021) Long-term culture, genetic manipulation and xenotransplantation of human normal and breast cancer organoids. *Nat Protoc*, **16**, 1936–1965.
- Deng, L. *et al.* (2020) The role of ubiquitination in tumorigenesis and targeted drug discovery. *Signal Transduct Target Ther*, **5**.
- Diamanti-Kandarakis, E. *et al.* (2009) Endocrine-disrupting chemicals: an Endocrine Society scientific statement. *Endocr Rev*, **30**, 293–342.
- Dobin, A. *et al.* (2013) STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
- Escher, B.I. *et al.* (2020) Cytotoxicity burst? Differentiating specific from nonspecific effects in tox21 in vitro reporter gene assays. *Environ Health Perspect*, **128**, 1–10.
- Farmahin, R. *et al.* (2017) Recommended approaches in the application of toxicogenomics to derive points of departure for chemical risk assessment. *Arch Toxicol*, **91**, 2045–2065.
- Feng, Y. *et al.* (2016) Effects of bisphenol analogues on steroidogenic gene expression and hormone synthesis in H295R cells. *Chemosphere*, **147**, 9–19.
- Ferreira, R. *et al.* (2022) Bisphenols A, F, S and AF trigger apoptosis and/or endoplasmic reticulum stress in human endometrial stromal cells. *Toxicology*, **478**.
- Fridriksdottir, A.J. *et al.* (2015) Propagation of oestrogen receptor-positive and oestrogen-responsive normal human breast cells in culture. *Nat Commun*, **6**.
- Fuentes, N. and Silveyra, P. (2019) Estrogen receptor signaling mechanisms. *Adv Protein Chem Struct Biol*, **116**, 135–170.

- Gaidatzis, D. *et al.* (2015) QuasR: Quantification and annotation of short reads in R. *Bioinformatics*, **31**, 1130–1132.
- García-Córcoles, M.T. *et al.* (2018) Determination of bisphenols with estrogenic activity in plastic packaged baby food samples using solid-liquid extraction and clean-up with dispersive sorbents followed by gas chromatography tandem mass spectrometry analysis. *Talanta*, **178**, 441–448.
- Geens, T. *et al.* (2012) A review of dietary and non-dietary exposure to bisphenol-A. *Food and Chemical Toxicology*, **50**, 3725–3740.
- Generali, D. *et al.* (2006) Hypoxia-inducible factor-1 α expression predicts a poor response to primary chemoendocrine therapy and disease-free survival in primary human breast cancer. *Clinical Cancer Research*, **12**, 4562–4568.
- Goldinger, D.M. *et al.* (2015) Endocrine activity of alternatives to BPA found in thermal paper in Switzerland. *REGULATORY TOXICOLOGY AND PHARMACOLOGY*, **71**, 453–462.
- Government of Canada (2020) Technical Consultation: Proposed Subgrouping of Bisphenol A (BPA) Structural Analogues and Functional Alternatives.
- Guo, Q. *et al.* (2024) NF- κ B signaling in therapy resistance of breast cancer: Mechanisms, approaches, and challenges. *Life Sci*, **348**.
- Gys, C. *et al.* (2020) Biomonitoring and temporal trends of bisphenols exposure in Japanese school children. *Environ Res*, **191**.
- Hannan, F.M. *et al.* (2023) Hormonal regulation of mammary gland development and lactation. *Nat Rev Endocrinol*, **19**, 46–61.
- Hardy, A. *et al.* (2017) Update: use of the benchmark dose approach in risk assessment. *EFSA Journal*, **15**.
- Harrill, J.A. *et al.* (2021) High-Throughput Transcriptomics Platform for Screening Environmental Chemicals. *Toxicol Sci*, **181**, 68–89.
- Harrill, J.A. *et al.* (2024) Signature analysis of high-throughput transcriptomics screening data for mechanistic inference and chemical grouping. *Toxicological Sciences*.

- Health Canada and Environment and Climate Change Canada (2023) Draft strategy to replace, reduce or refine vertebrate animal testing under the Canadian Environmental Protection Act, 1999. .
- Henriksen, A.D. *et al.* (2020) Bisphenol a exposure in utero disrupts hypothalamic gene expression particularly genes suspected in autism spectrum disorders and neuron and hormone signaling. *Int J Mol Sci*, **21**, 1–16.
- Hong, X. *et al.* (2022) Lipidomic biomarkers: Potential mediators of associations between urinary bisphenol A exposure and colorectal cancer. *J Hazard Mater*, **427**.
- Hua, H. *et al.* (2018) Mechanisms for estrogen receptor expression in human cancer. *Exp Hematol Oncol*, **7**.
- Hua, S. *et al.* (2009) Genomic Antagonism between Retinoic Acid and Estrogen Signaling in Breast Cancer. *Cell*, **137**, 1259–1271.
- Jackson, A.C. *et al.* (2020) Identification of novel activators of the metal responsive transcription factor (MTF-1) using a gene expression biomarker in a microarray compendium. *Metallomics*, **12**, 1400–1415.
- Jennings, P. *et al.* (2013) An overview of transcriptional regulation in response to toxicological insult. *Arch Toxicol*, **87**, 49–72.
- Al Jishi, T. and Sergi, C. (2017) Current perspective of diethylstilbestrol (DES) exposure in mothers and offspring. *Reproductive Toxicology*, **71**, 71–77.
- Johnson, K.J. *et al.* (2022) A Transformative Vision for an Omics-Based Regulatory Chemical Testing Paradigm. *Toxicological Sciences*, **190**, 127–132.
- Judson, R. *et al.* (2016) Analysis of the effects of cell stress and cytotoxicity on in vitro assay activity across a diverse chemical and assay space. *Toxicological Sciences*, **152**, 323–339.
- Keminer, O. *et al.* (2020) A tiered high-throughput screening approach for evaluation of estrogen and androgen receptor modulation by environmentally relevant bisphenol A substitutes. *Science of the Total Environment*, **717**.

- Keshavarz-Maleki, R. *et al.* (2021) Bisphenol-A in biological samples of breast cancer mastectomy and mammoplasty patients and correlation with levels measured in urine and tissue. *Sci Rep*, **11**.
- Khalili Sadrabad, E. *et al.* (2023) Bisphenol A release from food and beverage containers – A review. *Food Sci Nutr*, **11**, 3718–3728.
- Kim, C.H. *et al.* (2021) Bisphenol a exposure changes the transcriptomic and proteomic dynamics of human retinoblastoma y79 cells. *Genes (Basel)*, **12**, 1–18.
- Kim, S. *et al.* (2021) Urinary concentrations of bisphenol mixtures during pregnancy and birth outcomes: The MAKE study. *Int J Environ Res Public Health*, **18**.
- Kitamura, S. *et al.* (2005) Comparative study of the endocrine-disrupting activity of bisphenol A and 19 related compounds. *Toxicological Sciences*, **84**, 249–259.
- Kojima, H. *et al.* (2019) Profiling of bisphenol A and eight its analogues on transcriptional activity via human nuclear receptors. *Toxicology*, **413**, 48–55.
- Korunes, K.L. *et al.* (2022) A gene expression biomarker for predictive toxicology to identify chemical modulators of NF- κ B. *PLoS One*, **17**, 1–24.
- Krewski, D. *et al.* (2010) Toxicity testing in the 21st century: A vision and a strategy. *J Toxicol Environ Health B Crit Rev*, **13**, 51–138.
- Kupersmidt, I. *et al.* (2010) Ontology-Based Meta-Analysis of Global Collections of High-Throughput Public Data. **5**.
- Kuruto-Niwa, R. *et al.* (2005) Estrogenic activity of alkylphenols , bisphenol S , and their chlorinated derivatives using a GFP expression system. **19**, 121–130.
- Lee, A. V. *et al.* (2015) MCF-7 Cells - Changing the Course of Breast Cancer Research and Care for 45 Years. *J Natl Cancer Inst*, **107**.
- Lee, H. *et al.* (2023) Mixture Effects of Bisphenol A and Its Structural Analogs on Estrogen Receptor Transcriptional Activation. *Toxics*, **11**.
- Li, H.H. *et al.* (2019) TGx-DDI, a Transcriptomic Biomarker for Genotoxicity Hazard Assessment of Pharmaceuticals and Environmental Chemicals. *Front Big Data*, **2**, 1–14.

- Li, J. *et al.* (2016) Gestational and lactational exposure to bisphenol AF in maternal rats increases testosterone levels in 23-day-old male offspring. *Chemosphere*, **163**, 552–561.
- Li, Y.Y. *et al.* (2013) Down-Regulation of C9orf86 in Human Breast Cancer Cells Inhibits Cell Proliferation, Invasion and Tumor Growth and Correlates with Survival of Breast Cancer Patients. *PLoS One*, **8**.
- Liao, C. and Kannan, K. (2013) Concentrations and profiles of bisphenol a and other bisphenol analogues in foodstuffs from the united states and their implications for human exposure. *J Agric Food Chem*, **61**, 4655–4662.
- Lim, L.Y. *et al.* (2009) Mutant p53 mediates survival of breast cancer cells. *Br J Cancer*, **101**, 1606–1612.
- Liu, B. *et al.* (2019) Association of bisphenol A and its substitutes, bisphenol F and bisphenol S, with obesity in United States children and adolescents. *Diabetes Metab J*, **43**, 59–75.
- Love, M.I. *et al.* (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*, **15**, 1–21.
- Lucarini, F. *et al.* (2020) Exposure to new emerging bisphenols among young children in Switzerland. *Int J Environ Res Public Health*, **17**, 1–15.
- Lucarini, F. *et al.* (2023) Simultaneous Quantification of 16 Bisphenol Analogues in Food Matrices. *Toxics*, **11**.
- Ma, Y. *et al.* (2019) The adverse health effects of bisphenol A and related toxicity mechanisms. *Environ Res*, **176**.
- Mandrah, K. *et al.* (2017) A dispersive liquid-liquid microextraction based on solidification of floating organic droplet followed by injector port silylation coupled with gas chromatography-tandem mass spectrometry for the determination of nine bisphenols in bottled carbonated be. *J Chromatogr A*, **1528**, 10–17.
- Manzoor, M.F. *et al.* (2022) An insight into bisphenol A, food exposure and its adverse effects on health: A review. *Front Nutr*, **9**.

- Matteo, G. *et al.* (2025) High-throughput transcriptomics analysis of equipotent and human relevant mixtures of BPA alternatives reveal additive effects in vitro. *Archives of Toxicology*, **99** (9):3707-3719.
- Matteo, G. *et al.* (2023) In vitro transcriptomic analyses reveal pathway perturbations, estrogenic activities, and potencies of data-poor BPA alternative chemicals. *Toxicological Sciences*, **191**, 266–275.
- Meier, M.J. *et al.* (2024) Progress in toxicogenomics to protect human health. *Nat Rev Genet.*
- Mentor, A. *et al.* (2020) Bisphenol AF and bisphenol F induce similar feminizing effects in chicken embryo testis as bisphenol A. *Toxicological Sciences*, **178**, 239–250.
- La Merrill, M.A. *et al.* (2020) Consensus on the key characteristics of endocrine-disrupting chemicals as a basis for hazard identification. *Nat Rev Endocrinol*, **16**, 45–57.
- Mesnage, R. *et al.* (2022) Transcriptome analysis of human mammary epithelial cells treated with bisphenol A and bisphenol A analogue mixtures reveals major alterations in multiple cellular pathways. *Public Health and Toxicology*, **2**, 1–10.
- Mesnage, R. *et al.* (2017) Transcriptome profiling reveals bisphenol a alternatives activate estrogen receptor alpha in human breast cancer cells. *Toxicological Sciences*, **158**, 431–443.
- Moerbeek, M. *et al.* (2004) A Comparison of Three Methods for Calculating Confidence Intervals for the Benchmark Dose. *Risk Analysis*, **24**, 31–40.
- Mokra, K. *et al.* (2018) Low-concentration exposure to BPA, BPF and BPAF induces oxidative DNA bases lesions in human peripheral blood mononuclear cells. *Chemosphere*, **201**, 119–126.
- Mölder, F. *et al.* (2021) Sustainable data analysis with Snakemake. *F1000Res*, **10**, 33.
- Moreman, J. *et al.* (2017) Acute Toxicity, Teratogenic, and Estrogenic Effects of Bisphenol A and Its Alternative Replacements Bisphenol S, Bisphenol F, and Bisphenol AF in Zebrafish Embryo-Larvae. *Environ Sci Technol*, **51**, 12796–12805.

- Moreno-Gómez-Toledano, R. *et al.* (2023) Combination of Bisphenol A and Its Emergent Substitute Molecules Is Related to Heart Disease and Exerts a Differential Effect on Vascular Endothelium. *Int J Mol Sci*, **24**.
- National Toxicology Program (2018) NTP Research Report on National Toxicology Program Approach to Genomic Dose-Response Modeling: Research Report 5.
- Ni, L. *et al.* (2023) Recent Advances in Sources, Migration, Public Health, and Surveillance of Bisphenol A and Its Structural Analogs in Canned Foods. *Foods*, **12**.
- Niu, Y. *et al.* (2021) Bisphenol Analogues and Their Chlorinated Derivatives in Breast Milk in China: Occurrence and Exposure Assessment. *J Agric Food Chem*, **69**, 1391–1397.
- OECD (2007) Test No. 440: Uterotrophic Bioassay in Rodents: A short-term screening test for oestrogenic properties. In, *OECD Guidelines for the Testing of Chemicals*. OECD Publishing, Paris.
- OECD (2018) Test No. 443: Extended One-Generation Reproductive Toxicity Study. In, *OECD Guidelines for the Testing of Chemicals*. OECD Publishing, Paris.
- OECD (2021) Test No. 455: Performance-Based Test Guideline for Stably Transfected Transactivation In Vitro Assays to Detect Estrogen Receptor Agonists and Antagonists. In, *OECD Guidelines for the Testing of Chemicals*. OECD Publishing, Paris.
- OECD (2012) Test No. 457: BG1Luc Estrogen Receptor Transactivation Test Method for Identifying Estrogen Receptor Agonists and Antagonists. In, *OECD Guidelines for the Testing of Chemicals*. OECD Publishing, Paris.
- OECD (2024) Test No. 493: Performance-Based Test Guideline for Human Recombinant Estrogen Receptor (hrER) In Vitro Assays to Detect Chemicals with ER Binding Affinity. OECD Publishing, Paris.
- Oshida, K. *et al.* (2015) Screening a mouse liver gene expression compendium identifies modulators of the aryl hydrocarbon receptor (AhR). *Toxicology*, **336**, 99–112.

- Owczarek, K. *et al.* (2018) Binary mixtures of selected bisphenols in the environment: Their toxicity in relationship to individual constituents. *Molecules*, **23**.
- Park, C. *et al.* (2020) The mixture effects of bisphenol derivatives on estrogen receptor and androgen receptor. *Environmental Pollution*, **260**.
- Paul Friedman, K. *et al.* (2020) Utility of in Vitro Bioactivity as a Lower Bound Estimate of in Vivo Adverse Effect Levels and in Risk-Based Prioritization. *Toxicological Sciences*, **173**, 202–225.
- Pelch, K.E. *et al.* (2019) Characterization of Estrogenic and Androgenic Activities for Bisphenol A-like Chemicals (BPs): In Vitro Estrogen and Androgen Receptors Transcriptional Activation, Gene Regulation, and Binding Profiles. *Toxicological Sciences*, **172**, 23–37.
- Pellacani, D. *et al.* (2019) Transcriptional regulation of normal human mammary cell heterogeneity and its perturbation in breast cancer. *EMBO J*, **38**.
- Peshdary, V. *et al.* (2021) Transcriptomic pathway and benchmark dose analysis of Bisphenol A, Bisphenol S, Bisphenol F, and 3,3',5,5'-Tetrabromobisphenol A in H9 human embryonic stem cells. *Toxicology in Vitro*, **72**.
- Petersen, O.W. *et al.* (1987) Frequency and Distribution of Estrogen Receptor-positive Cells in Normal, Nonlactating Human Breast Tissue¹.
- Qin, T. *et al.* (2021) Epigenetic Alteration Shaped by the Environmental Chemical Bisphenol A. *Front Genet*, **11**.
- Qin, X.Y. *et al.* (2012) Effects of bisphenol A exposure on the proliferation and senescence of normal human mammary epithelial cells. *Cancer Biol Ther*, **13**, 296–306.
- Reardon, A.J.F. *et al.* (2023) From vision toward best practices: Evaluating in vitro transcriptomic points of departure for application in risk assessment using a uniform workflow. *Frontiers in Toxicology*, **5**.
- Reardon, A.J.F. *et al.* (2021) Potency Ranking of Per- And Polyfluoroalkyl Substances Using High-Throughput Transcriptomic Analysis of Human Liver Spheroids. *Toxicological Sciences*, **184**, 154–169.

- Reeves, K.W. *et al.* (2018) Bisphenol-A in breast adipose tissue of breast cancer cases and controls. *Environ Res*, **167**, 735–738.
- Regueiro, J. and Wenzl, T. (2015) Determination of bisphenols in beverages by mixed-mode solid-phase extraction and liquid chromatography coupled to tandem mass spectrometry. *J Chromatogr A*, **1422**, 230–238.
- Reina-Pérez, I. *et al.* (2022) The Mixture of Bisphenol-A and Its Substitutes Bisphenol-S and Bisphenol-F Exerts Obesogenic Activity on Human Adipose-Derived Stem Cells. *Toxics*, **10**.
- Rochester, J.R. and Bolden, A.L. (2015) Bisphenol S and F: A systematic review and comparison of the hormonal activity of bisphenol a substitutes. *Environ Health Perspect*, **123**, 643–650.
- Rogers, L.D. (2021) What does clarity-bpa mean for Canadians? *Int J Environ Res Public Health*, **18**, 1–9.
- Rooney, J.P. *et al.* (2020) Mining a human transcriptome database for chemical modulators of NRF2. *PLoS One*, **15**, 1–26.
- Rosenbluth, J.M. *et al.* (2020) Organoid cultures from normal and cancer-prone human breast tissues preserve complex epithelial lineages. *Nat Commun*, **11**.
- Rosenmai, A.K. *et al.* (2014) Are structural analogues to bisphenol a safe alternatives? *Toxicological Sciences*, **139**, 35–47.
- Ross-Innes, C.S. *et al.* (2010) Cooperative interaction between retinoic acid receptor- α and estrogen receptor in breast cancer. *Genes Dev*, **24**, 171–182.
- Rowan-Carroll, A. *et al.* (2021) High-Throughput Transcriptomic Analysis of Human Primary Hepatocyte Spheroids Exposed to Per- And Polyfluoroalkyl Substances as a Platform for Relative Potency Characterization. *Toxicological Sciences*, **181**, 199–214.
- Ryan, N. *et al.* (2016) Moving toward integrating gene expression profiling into high-throughput testing: A gene expression biomarker accurately predicts estrogen receptor α modulation in a microarray compendium. *Toxicological Sciences*, **151**, 88–103.

- dos Santos, A. *et al.* (2022) Knockdown of CKAP2 Inhibits Proliferation, Migration, and Aggregate Formation in Aggressive Breast Cancer. *Cancers (Basel)*, **14**.
- Schaffert, A. *et al.* (2021) Alternatives for the worse: Molecular insights into adverse effects of bisphenol a and substitutes during human adipocyte differentiation. *Environ Int*, **156**.
- Shi, Y. *et al.* (2010) The metal-responsive transcription factor-1 protein is elevated in human tumors.
- Shioda, T. *et al.* (2013) Expressomal approach for comprehensive analysis and visualization of ligand sensitivities of xenoestrogen responsive genes. *Proc Natl Acad Sci U S A*, **110**, 16508–16513.
- Singh, M. *et al.* (2024) Tetra methyl bisphenol F: another potential obesogen. *Int J Obes*, **48**, 923–933.
- Soto, A.M. *et al.* (2017) Evidence of Absence: Estrogenicity Assessment of a New Food-Contact Coating and the Bisphenol Used in Its Synthesis. *Environ Sci Technol*, **51**, 1718–1726.
- Soto, A.M. *et al.* (1995) The E-SCREEN Assay as a Tool to Identify Estrogens: An Update on Estrogenic Environmental Pollutants. *Environ Health Perspect*, **Suppl 7**, 113–122.
- Stephens, M. (2017) False discovery rates: A new deal. *Biostatistics*, **18**, 275–294.
- Stucki, A.O. *et al.* (2022) Use of new approach methodologies (NAMs) to meet regulatory requirements for the assessment of industrial chemicals and pesticides for effects on human health. *Frontiers in Toxicology*, **4**.
- Sun, F. *et al.* (2023) BPA and its alternatives BPF and BPAF exaggerate hepatic lipid metabolism disorders in male mice fed a high fat diet. *Science of the Total Environment*, **867**.
- Tang, S.C. *et al.* (2022) The E3 ligase COP1 promotes ER α signaling and suppresses EMT in breast cancer. *Oncogene*, **41**, 173–190.

- Thomas, R.S. *et al.* (2013) Incorporating new technologies into toxicity testing and risk assessment: Moving from 21st century vision to a data-driven framework. *Toxicological Sciences*, **136**, 4–18.
- Thomas, R.S. *et al.* (2019) The next generation blueprint of computational toxicology at the U.S. Environmental protection agency. *Toxicological Sciences*, **169**, 317–332.
- US Environmental Protection Agency (2012a) Benchmark Dose Technical Guidance.
- US Environmental Protection Agency (2012b) Endocrine Disruptor Screening Program Universe of Chemicals and General Validation Principles.
- US Environmental Protection Agency (USEPA) (2014) Bisphenol A Alternatives in Thermal Paper, final report Washington, DC.
- Vandenberg, L.N. (2021) Endocrine disrupting chemicals and the mammary gland. In, *Advances in Pharmacology*. Academic Press Inc., pp. 237–277.
- Vandenberg, L.N. *et al.* (2012) Hormones and endocrine-disrupting chemicals: Low-dose effects and nonmonotonic dose responses. *Endocr Rev*, **33**, 378–455.
- Verheijen, M.C.T. *et al.* (2022) R-ODAF: Omics data analysis framework for regulatory application. *Regulatory Toxicology and Pharmacology*, **131**, 105143.
- Wan, Y. ping *et al.* (2023) Ten bisphenol analogues in Chinese fresh dairy milk: high contribution ratios of conjugated form, importance of enzyme hydrolysis and risk evaluation. *Environmental Science and Pollution Research*, **30**, 88049–88059.
- Wang, H. *et al.* (2023) Large-scale biomonitoring of bisphenol analogues and their metabolites in human urine from Guangzhou, China: Implications for health risk assessment. *Chemosphere*, **338**.
- Wang, L. *et al.* (2021) Transcriptional coregulator NUPR1 maintains tamoxifen resistance in breast cancer cells. *Cell Death Dis*, **12**.
- Wang, T. *et al.* (2023) A new glance at autophagolysosomal-dependent or -independent function of transcriptional factor EB in human cancer. *Acta Pharmacol Sin*, **44**, 1536–1548.

- Weng, Y.I. *et al.* (2010) Epigenetic influences of low-dose bisphenol A in primary human breast epithelial cells. *Toxicol Appl Pharmacol*, **248**, 111–121.
- Winkler, J. *et al.* (2022) Bisphenol A replacement chemicals, BPF and BPS, induce protumorigenic changes in human mammary gland organoid morphology and proteome.
- Wörthmüller, J. *et al.* (2023) MAGI1 Prevents Senescence and Promotes the DNA Damage Response in ER+ Breast Cancer. *Cells*, **12**.
- Xian, Y. *et al.* (2017) Dispersive micro solid phase extraction (DMSPE) using polymer anion exchange (PAX) as the sorbent followed by UPLC–MS/MS for the rapid determination of four bisphenols in commercial edible oils. *J Chromatogr A*, **1517**, 35–43.
- Yersal, O. and Barutca, S. (2014) Biological subtypes of breast cancer: Prognostic and therapeutic implications. *World J Clin Oncol*, **5**, 412–424.
- Zeng, H. *et al.* (2016) Systematic identification of Ctr9 regulome in ER α -positive breast cancer. *BMC Genomics*, **17**, 1–14.
- Zhang, S. *et al.* (2024) Protein neddylation and its role in health and diseases. *Signal Transduct Target Ther*, **9**, 85.
- Zhang, Y. *et al.* (2024) O-GlcNAcylation of MITF regulates its activity and CDK4/6 inhibitor resistance in breast cancer. *Nat Commun*, **15**.
- Zhang, Z.G. *et al.* (2019) KDM5B promotes breast cancer cell proliferation and migration via AMPK-mediated lipid metabolism reprogramming. *Exp Cell Res*, **379**, 182–190.
- Zhao, H. *et al.* (2021) Inflammation and tumor progression: signaling pathways and targeted intervention. *Signal Transduct Target Ther*, **6**.
- Zhi, S. *et al.* (2024) Hypoxia-inducible factor in breast cancer: role and target for breast cancer treatment. *Front Immunol*, **15**.
- Zhu, J. *et al.* (2014) The atypical ubiquitin ligase RNF31 stabilizes estrogen receptor α and modulates estrogen-stimulated breast cancer cell proliferation. *Oncogene*, **33**, 4340–4351.

Zhu, J. *et al.* (2020) The ubiquitin ligase RNF181 stabilizes ER α and modulates breast cancer progression. *Oncogene*, **39**, 6776–6788.

Ziegler, Y. *et al.* (2019) Suppression of FOXM1 activities and breast cancer growth in vitro and in vivo by a new class of compounds. *NPJ Breast Cancer*, **5**.

Appendix A: A 50-gene biomarker identifies estrogen receptor-modulating chemicals in a microarray compendium

Modified from: Corton JC, Matteo G, Chorley B, Liu J, Vallanat B, Everett L, Atlas E, Meier MJ, Williams A, Yauk CL. A 50-gene biomarker identifies estrogen receptor-modulating chemicals in a microarray compendium. *Chemico-Biological Interactions*, 2024 May 1:394:110952.

A.1 Abstract

High throughput transcriptomics (HTTr) profiling has the potential to rapidly and comprehensively identify molecular targets of environmental chemicals that can be linked to adverse outcomes. We describe here the construction and characterization of a 50-gene expression biomarker designed to identify estrogen receptor (ER) active chemicals in HTTr datasets. Using microarray comparisons, the genes in the biomarker were identified as those that exhibited consistent directional changes when ER was activated (4 ER agonists; 4 ESR1 gene constitutively active mutants) and opposite directional changes when ER was suppressed (4 antagonist treatments; 4 ESR1knockdown experiments). The biomarker was evaluated as a predictive tool using the Running Fisher algorithm by comparison to annotated gene expression microarray datasets including

those evaluating the transcriptional effects of hormones and chemicals in MCF-7 cells. Depending on the reference dataset used, the biomarker had a predictive accuracy for activation of up to 96%. To demonstrate applicability for HTTr data analysis, the biomarker was used to identify ER activators in a set of 15 chemicals that are considered potential bisphenol A (BPA) alternatives examined at up to 10 concentrations in MCF-7 cells and analyzed by full-genome TempO-Seq. Using benchmark dose (BMD) modeling, the biomarker genes stratified the ER potency of BPA alternatives consistent with previous studies. These results demonstrate that the ER biomarker can be used to accurately identify ER activators in transcript profile data derived from MCF-7 cells.

A.2 Introduction

New approach methodologies (NAMs) that efficiently provide information about chemical hazards without using whole animals are needed to accelerate the pace of chemical risk assessments [1]. Only a small fraction of the thousands of chemicals in U.S. commerce (<https://www.epa.gov/tsca-inventory>) can be thoroughly evaluated for human safety using current animal toxicity testing approaches. NAMs are defined as any technology, methodology, approach or combination thereof that can be used to provide information on chemical hazard and risk that avoids the use of intact animals [1] and can encompass many different types and combinations of in vitro bioactivity studies, in silico modeling of bioactivities and exposure predictions, and chemoinformatics. There is an

increasing impetus for development and implementation of NAMs, given that the United States Environmental Protection Agency (US-EPA) has made a commitment to gradually phase out mammalian toxicity testing [2].

Technological advancements in gene expression assays have now put within practical reach in vitro high-throughput transcriptomics (HTTr) for NAMs-based hazard characterization of environmental chemicals. In the EPA ToxCast screening program, HTTr strategies are currently being used to replace the battery of individual ToxCast screening assays with targeted sequencing techniques [3,4]. In contrast to individual assays, HTTr technologies have the advantage of examining the effects of environmental chemicals on essentially all pathways active in cell models, many of which are not examined by any of the ToxCast assays. A major challenge is how to interpret the gene expression profiles to identify the molecular targets of chemicals and then to make predictions of potential adversity.

A number of approaches have been used to interpret the HTTr profiles, and they include comparison of profiles to those of reference chemicals with known bioactivity [5] or pathway analysis [6]. While these approaches are useful to identify the network of all potential pathways and targets that are perturbed by chemical exposure, these approaches do not necessarily lead to predictions with known accuracies. Gene expression biomarkers (also referred to as signatures) have emerged as an alternative approach to predict specific targets of chemicals. Biomarkers consist of sets of genes

known or predicted to be regulated by a particular factor or cellular process [7]. To identify chemical modulators of the factor being predicted, the biomarker gene expression pattern is compared to gene expression profiles derived from human cells exposed to chemicals using computational techniques including correlation analysis [8]. Information about the spectrum of factors/pathways altered by the chemicals can then be linked to predictions of adversity through the network of adverse outcome pathways [7].

An increasing number of gene expression biomarkers have been developed to predict the modulation of transcription factors in human cell lines. One of the earliest gene expression biomarkers described was the TGx-DDI biomarker that identifies relevant vs. nonrelevant DNA damaging chemicals [9]. This biomarker has been extensively characterized [10,11] and is currently undergoing review by the Food and Drug Administration to be used as a tool to identify potential DNA damaging agents in human TK6 cells. A number of biomarkers developed from DNA microarray data have been described that can identify chemical and genetic modulation in human cells lines including interactions with androgen receptor [12], metal-induced transcription factor 1[13], the oxidant-induced transcription factor Nrf2 [14], heat shock factor 1 [15] and NF-kB [16]. Additionally, a biomarker that identifies histone deacetylase inhibitors (the TGx-HDACi biomarker) has been used to screen for HDAC inhibitors [17]. Most of these biomarkers have been developed in part because of the availability of the large quantity of gene expression data that exists in commercial and public repositories. This data includes

reference chemicals with known target activities as well as studies in which the factor to be predicted has been genetically activated or suppressed.

In the present study, we built and characterized a biomarker that can be used in HTTr studies to identify estrogen receptor (ER) modulators. ER, like other nuclear receptorfamily members, regulates target gene expression through well-defined mechanisms. The classical pathway includes ligand binding by agonists followed by direct DNA bindingto estrogen response elements (ERE) and modulation of gene regulation [18,19]. Nonclassical pathways include post-transcriptional modulation of ER through upstream activation of several kinase-dependent signaling pathways. In addition, estrogens also activate ER-dependent transactivation through ER interactions with Sp1, AP1, and other DNA-bound transcription factors [19,20]. We have previously described a 46-gene biomarker that can accurately predict both ER activation (agonism) or suppression (antagonism) in ER positive human breast cancer cells after chemical exposure [21]. While the biomarker worked well with gene expression profiles generated from microarray comparisons [22], our preliminary studies indicated that the biomarker did not optimally identify ER actives in profiles generated by targeted RNA-Seq, possibly because many of the biomarker genes were below the level of detection. We thus constructed and characterized a novel ER biomarker that included genes detected in HTTr data generated by full genome Templated Oligo-Sequencing (TempO-Seq) in MCF-7 cells. The ER biomarker includes 50 genes with consistent expression changes upon ER activation or ER suppression using both chemical treatments and perturbations of the

ESR1 gene. We show here that the biomarker accurately identified both chemical and genetic ER activators as well as suppressors in our analyses.

A.3 Methods

Overall strategy for identification of ER-dependent genes and their use as a biomarker to screen for ER active chemicals. A summary of the methods used in this study is outlined in Figure A.1. Differentially expressed genes (DEGs) from chemical or genetic perturbations of MCF-7 cells were used to identify biomarker genes. These perturbations included exposure to known ER activators (bisphenol A, 17 β -estradiol (E2), estrone, nonylphenol), ER antagonists or selective ER modulators (SERMs) (bazedoxifene, fulvestrant (also known as ICI 182780), raloxifene), and genetic perturbations that lead to increased (ESR1 constitutively active mutants) or decreased (knockdown) expression/activity of ESR1. Genes that exhibited consistent directional changes across the 16 perturbations and that are expressed in MCF-7 cells using TempO-seq were selected as biomarker genes. The biomarker includes fold-change values associated with each gene, derived from the average differences in expression across treatment by the 4 agonists. The biomarker genes were characterized using Ingenuity Pathway Analysis (IPA) for canonical pathway enrichment and potential upstream transcription factors that regulated the genes. Identification of primary targets of ER α was carried out through analysis of chromatin immunoprecipitation coupled with

DNA sequencing (ChIP-Seq) datasets. This analysis led to the identification of a potential 50-gene biomarker.

Associated fold-change values for the 50-gene biomarker were uploaded into BaseSpace Correlation Engine (BSCE) and then, using the Running Fisher test compared to every gene list (also called biosets) in the database (~140,000 biosets). For the overlapping genes between the biomarker and individual biosets, the method determines whether those overlapping genes are significantly regulated in a similar or opposite manner. Biosets which exhibited expression of biomarker genes that are positively correlated with the biomarker would be predicted to exhibit ER activation. Biosets which exhibited expression of biomarker genes that are negatively correlated to the biomarker would be predicted to exhibit ER suppression. Available information about each bioset was extracted from BSCE and used to populate a spreadsheet of experimental parameters. Biosets generated from experiments in the human breast cancer cell line, MCF-7, were used to characterize the biomarker accuracy. The ability of the biomarker to identify ER actives in MCF-7 cells using TempO-Seq was determined using a dataset generated after exposure to bisphenol A alternatives. It should be noted that while the biomarker was developed using perturbations in MCF-7 cells in which ER α is the predominant subtype [[23], [24], [25]], we designated the final gene list the “ER biomarker” as our methods may also identify ER β activators, given the similarities in compound activation of the two subtypes [26]. Determination of whether the ER biomarker could identify ER β activators will require a separate analysis.

Use of a gene expression microarray experiment compendium. As described previously [22], data from BaseSpace Correlation Engine (BSCE) (<https://www.illumina.com/products/by-type/informatics-products/basespace-correlation-engine.html>) was used to build a spreadsheet of gene expression comparisons (called biosets) from experiments carried out using human cell lines and tissues. This compendium included the study accession information, bioset name, cell line, tissue, chemical name and when available chemical concentration and treatment time. Of the biosets included, ~13,400 biosets were derived from experiments on chemically-treated cells. Methods used to derive the filtered gene lists are described in detail previously [8]. In short, gene expression data were processed using BSCE protocols to generate filtered gene lists with expression fold-change values for each bioset; information about the methods used to normalize and identify DEGs have been described previously [8]. Gene lists were filtered to include only genes which had an absolute fold change magnitude of ≥ 1.2 and $p < 0.05$.

Construction of the ER biomarker. Lists of statistically filtered genes from BSCE were used to derive a consensus gene expression biomarker for ER. Biosets used to create the biomarker are described in Table A.1. These biosets were selected because they included many structurally diverse and well-known agonists and antagonists or SERMs. Additionally, biosets were included in which the ESR1 gene was genetically modulated. These included 4 biosets from experiments in which the wild-type or constitutively active ESR1 were expressed and 4 biosets in which cells were treated with

si- or shRNAtargeted to decrease the expression of ESR1. The top 5000 genes with the greatest degree of overlap in expression between all 16 biosets were identified using the “meta-analysis” function in BSCE, and all data were exported. The meta-analysis function identifies those genes that maximally overlap in expression changes between the biosets examined. The genes in the biomarker were selected using the following steps. First, those genes which exhibited consistent expression behavior across the 4 agonists were selected. For this filter, the genes had to consistently exhibit either up or down regulation in 2 or more of the 4 comparisons. The resulting gene list was then compared with the gene profiles of cells treated with the 3 antagonist/SERMs from 4 biosets. The genes that exhibited opposite directional changes to the agonist profiles in at least 2 of the 4 biosets were selected. The gene selection process was repeated with the biosets in which ESR1 was genetically activated and then again with the 4 biosets in which ESR1 was suppressed using the same filtering criteria. Like the chemical comparisons, to pass these thresholds, the gene had to exhibit either up or down regulation in 2 or more of the 4 comparisons. For the knockdown experiments, the fold-change in ESR1 expression was -3.02 (GSE18431), -2.21, -1.74, -1.25 (GSE10061), -4.79, -1.6 (GSE37820), and -2.4, -1.63, -1.42, -1.32 (GSE10890) (the ESR1 gene is represented by multiple probe sets on the Affymetrix arrays). The genes were further filtered for those that were consistently expressed in TempO-seq whole transcriptome data from a screen of ~1500 ToxCast chemical samples using MCF-7 cells. Consistently expressed genes were found to have at least one probe with mean count ≥ 5 in TempO-seq data generated for 95% or more of

the tested chemicals. Lastly, the average fold-change of the genes was determined across the 4 agonist comparisons. The final list consisted of 50 genes. The average fold-change values and gene abbreviations were imported into BSCE without any further filtering.

Ingenuity pathway analysis. The ER biomarker genes were analyzed using the canonical pathway and upstream analysis functions of Ingenuity Pathway Analysis (IPA, Qiagen Bioinformatics, Redwood City, California). IPA was used to calculate the significance of overlap of the biomarker genes with canonical pathways and upstream transcription factors using a right-tailed Fisher's Exact test, yielding a list of p-values describing the probability of the overlap between the ER biomarker gene list and the IPA pathway gene lists. Upstream analysis used the number of DEGs to predict upstream regulators of the biomarker genes.

Identification of interactions between ER and ER biomarker genes using ChIP-Seq data. Analysis using human-based ChIP-Seq data was carried out to identify biomarker genes directly regulated by ER. Specifically, we utilized the Target Gene function analysis of ChIP-Atlas (<https://chip-atlas.org/> [27]), which accesses over 144,000 publicly available experiments saved as Sequence Read Archives (SRA). Briefly, human genome(hg38) archives were searched for ESR1 ChIP experiments focusing on ± 10 kB of transcriptional start sites (TSS) of protein-encoding genes based on reFlat data (<http://hgdownload.cse.ucsc.edu/downloads.html>). The BEDTools (ver. 2.17.0) “window”

command was used to search for target genes among peak-call data and the TSS library. BED4-formatted ESR1 peak-call data of each SRA were determined by MAC2 scores (MACS2 Q-value $< 1 \times 10^{-5}$) [28]. A total of 972 ESR1-specific SRA datasets from 27 sample types were found. If a gene intersected with multiple peaks of a single SRA, the highest MACS2 score was chosen for gene scoring. Binding scores were converted from MACS2 Q-values ($-10 \times \log_{10}[\text{MACS2 Q-value}]$) and a score of 100 was considered a significant peak in the gene window for an analyzed SRA.

Comparison of the biomarker to database biosets. The ER biomarker was compared to biosets in the BSCE database using the Running Fisher algorithm. This method provides an assessment of the statistical significance of the overlapping genes between the biomarker and each bioset by assessing correlation and providing a summary p-value. A complete description of the Running Fisher test was previously described [8]. The p-values of the pair-wise comparisons were exported and converted to $-\log(\text{p-values})$, with negative values used to indicate negative correlation between the biomarker and the bioset. Biosets with $|\log(\text{p-value})| \geq 4$ were considered significant based on prior studies using this threshold [29,30]. Biosets that were positively correlated with the biomarker ($-\log(\text{p-value}) \geq 4$) were predicted to exhibit activation of ER, while biosets that were negatively correlated ($-\log(\text{p-value}) \leq -4$) were predicted to exhibit suppression of ER. It should be noted that we are referring to activation or suppression of ER as opposed to agonism or antagonism to appropriately include chemical-

independent ER modulators such as changes in the abundance or function of the ESR1 gene itself.

Comparison of biomarker predictions to OECD reference chemicals. A list of ER reference chemicals was derived from the Organisation for Economic Co-operation and Development (OECD) TG457 BG1 guidance document [31]. Of the 45 reference chemicals in the document, there were 19 chemicals represented by 122 biosets in the MCF-7 compendium, including 2 strong agonists (E2, diethylstilbestrol), 2 moderate agonists (alpha-estradiol, estrone), 4 weak agonists (bisphenol A, bisphenol B, daidzein, genistein), 4 very weak agonists (4-nonylphenol, apigenin, chrysin, kaempferol), and 7 inactives (corticosterone, cycloheximide, flutamide, haloperidol, ketoconazole, reserpine, spironolactone).

Comparison of ER biomarker predictions to the ToxCast ER model agonist scores. The predictive capabilities of the 50-gene ER biomarker model were compared to that of the ToxCast ER computational model [32] for chemicals that were examined in MCF-7 cells. Area under the curves (AUC) for the agonist predictions were acquired from Supplemental File 2 in this study [32]. To determine estrogenic activity for the compound, we used the same cutoff ($AUC \geq 0.0501$) applied in earlier studies [32,33]. In addition, Browne et al. removed any chemicals from their analysis that were designated as “inconclusive” ($0 < AUC < 0.0501$) [33]. Prior studies with the 46-gene ER biomarker showed that using a cutoff of the Running Fisher algorithm $p\text{-value} \leq 10^{-4}$ after a

Benjamini Hochberg correction of $\alpha = 0.001$ resulted in a balanced accuracy of ER activation or ER suppression of 94% or 93%, respectively [21] and thus this cutoff was used in the present analysis. The values for predictive accuracy were calculated as follows: sensitivity (TP rate) = $TP/(TP + FN)$; specificity (TN rate) = $TN/(FP + TN)$; positive predictive value (PPV) = $TP/(TP + FP)$; negative predictive value (NPV) = $TN/(TN + FN)$; balanced accuracy = (sensitivity + specificity)/2; where FN = false negative and FP = false positive.

Comparison between ER α biomarker predictions and the Tox21 ER α quantitative high-throughput screening (HTS) assays. ER α biomarker predictions were compared to two sets of assays that were carried out by the NIH National Center for Advancing Translational Sciences Chemical Genomics Center as part of their chemical screening program [34]. Data was taken directly from the following websites on Nov. 30, 2016: <https://ntp.niehs.nih.gov/sandbox/tox21-curve-visualization/> and <https://ntp.niehs.nih.gov/sandbox/tox21-activity-browser/>. The data was derived from the reporter gene assay, ER-luc, run in both agonist mode and agonist mode in the presence of 10 nM fulvestrant in a quantitative HTS format. The VM7 cell line used was originally thought to be the human ovarian cancer cell line BG1, but is now recognized to be a variant of MCF-7 cells (<https://ntp.niehs.nih.gov/iccvam/methods/endocrine/bg1luc/bg1luc-vm7luc-june2016-508.pdf>; accessed Sept. 2, 2019). The cell line was stably transfected with an estrogen-responsive luciferase reporter gene plasmid (pGudLuc7ere) containing the estrogen

responsive element (ERE) and luciferase reporter gene [35]. Results of the assays were filtered prior to comparison with the ER α biomarker to exclude chemicals:

- with marginal activity (i.e., hit calls of 0.5),
- classified as resulting in weak or noisy data,
- assayed in duplicate and gave inconsistent results, i.e., were positive in one assay but not the other,
- that did not meet purity standards as specified in <https://tripod.nih.gov/tox21/samples> to eliminate substances with suboptimal F-graded purity (Fns, Fc, and F) when either 4-month purity or 0-month purity information was available (when both were available, 4-month purity information was used),
- that exhibited activity that significantly overlapped with concentrations that induced cytotoxicity. Chemicals flagged as cytotoxic were identified as described previously [36].

In addition, chemicals were excluded if the concentration of the chemical in the microarray study was below the EC50 value in the HTS assay and was not active using the biomarker approach. This criterion excluded experiments that were examined at concentrations of chemical likely to be insufficient for inducing activation of ER α . Chemicals were designated as ER activators if they were positive in both assays and were designated as negative if they were negative in both assays. The 8 biosets from

chemical treatments used to create the biomarker (the training set) were not included in the analysis.

The accuracy analysis was carried by two methods. In the first method, true positive biosets were used individually, assuming that the conditions of exposure would always lead to activation of ER. However, there were a number of studies in which ER was activated only at a subset of concentrations or times of exposure. Thus, the second method examined the biosets by study. If any bioset in that study was positive for ER activation, then the chemical in that study was called positive. Accuracy determinations were carried out as described above.

ER biomarker predictions in MCF-7 cells treated with BPA and alternatives. For detailed experimental methods see Matteo et al. [37]. Briefly, MCF-7 cells were cultured in complete media and transferred to phenol red-free media containing charcoal stripped serum for 48 h to remove estrogens. Cells were then exposed to BPA or 15 alternative chemicals (Table A.2) across 10 concentrations (0.0005, 0.001, 0.01, 0.1, 0.5, 1, 5, 10, 50, 100 μ M), as well as positive (17 β -estradiol), negative (dexamethasone), and solvent (DMSO) controls, for 48 h. Samples were lysed for TempO-Seq and gene expression was measured as per manufacturer's instructions. Gene expression data are available on NCBI's Gene Expression Omnibus (GEO) under the accession number GSE211183.

For the ER biomarker genes, normalized read counts were shifted by 0.5 and then log₂transformed using the counts function in the DESeq2 package and imported into

BMDEExpress v2.3 [38]. Benchmark concentration (BMC) analysis of ER biomarker genes was conducted in accordance with recommendations outlined in the US National Toxicology Program (NTP) Research Report on National Toxicology Program Approach to Genomic Dose-Response Modeling [39]. Test chemicals with positive ER classifications were included for BMC modeling with cytotoxic concentrations removed. Pre-filtering was conducted using the Williams trend test retaining features with a permutation p-value <0.05 (with 500 permutations) with a fold change cut-off of 1.5. Genes passing the pre-filters were fit to linear, restricted power, 2^o polynomial and the exponential models (2, 3, 4, and 5). The best fit model was selected using the Akaike Information Criterion(AIC) value. A benchmark response (BMR) of 1 standard deviation was used to define the BMC. Post filtering of BMCs was conducted. Genes with goodness of fit p-values greater than 0.1, a BMC/BMC lower-limit (BMCL) ratio <20 , a BMC upper-limit (BMCU)/BMCL ratio <40 and BMC $<$ the highest concentration were retained. Confidence intervals for the median BMC was obtained using the bootstrap method. For each gene, 100 bootstrap samples were generated, bootstrapping the residuals from the full model. The bootstrapped data were imported into BMDEExpress v2.3 with the same filtering criteria and model selection as in the previous analysis. As bootstrap samples are independent, genes were randomly selected to generate 2000 bootstrap experiments based on each gene's inclusion probability. Inclusion probabilities were estimated by the relative frequency of that gene passing all the filtering criteria. If a gene was selected, then a BMC was randomly selected from the list of BMCs for that gene. For each

bootstrap experiment, the median BMC was estimated. From the resulting bootstrap distribution, 95% percentile confidence intervals were obtained for the median.

For the 50 ER biomarker genes, normalized enrichment scores (NES) were calculated on a per sample basis. This was done by normalizing the data to controls. Gene set enrichment analysis was then conducted using the `fgseaSimple` function with 500 permutations from the `fgsea` R package [40]. Pre-filtering was conducted using the Williams trend test with 500 permutations using the identical models in all previous analyses; however, no post filtering was conducted on the NES BMCs.

A.4 Results

A.4.1. Assembly and characterization of a 50-gene ER biomarker

To assemble a biomarker predictive of ER modulation, gene expression comparisons (biosets) were utilized from MCF-7 cells, a cell line which expresses ER α as the major ER subtype [[23], [24], [25]]. As described in the Materials and Methods, the biomarker was built using biosets from MCF-7 cells perturbed by: 1) treatment with chemical ER agonists or antagonists; 2) genetic modification to overexpress wild-type or constitutively active ESR1; or 3) knockdown of expression of the ESR1 gene. The details of the weight of evidence approach for selecting the genes is described in the Methods, and the biosets used to make the biomarker are described in Table A.1. A total of 50 genes (39 with increased expression and 11 with decreased expression) were identified

that exhibited consistent regulation across the treatments (Figure A.2A, upper panel). A number of the identified genes are well-known targets of ER, including PGR [41], CXCL12 [42], EGR3[43], and PRSS23 [44]. There was only a 13 gene overlap in the genes between this 50-gene biomarker and the previously described 46-gene ER biomarker. The full list of genes in the 50-gene biomarker is found in Table S1 where it is compared to that of the 46-gene ER biomarker.

A.4.2 Comparison of the biomarker to individual biosets used to create the biomarker

The Running Fisher algorithm [8], a fold-change rank-based pattern matching method, was used to assess modulation of ER. The method calculates the significance of the correlation for the genes that overlap between the biomarker and biosets in the database. Using this approach, the ER biomarker was examined for correlation to the 16 biosets used to identify the biomarker genes (Figure A.2A, lower panel). As expected, the biosets from agonist-treated cells exhibited statistically significant positive correlation to the biomarker ($p\text{-values} \leq 10^{-9}$), and the biosets from treatments with the 3 antagonists exhibited significant negative correlation to the biomarker ($p\text{-values} \leq 10^{-11}$) (Figure A.2A, lower panel). Likewise, overexpression of wild-type or constitutively-active mutants led to consistent positive correlation ($p\text{-values} \leq 10^{-8}$), and the treatments that decreased expression of the ESR1 gene led to negative correlation ($p\text{-values} \leq 10^{-8}$).

A.4.3 Functional classification of the biomarker genes

The 50 ER biomarker genes were examined for upstream analysis function and canonical pathway enrichment by IPA. The upstream analysis function of IPA identified a number of transcription factors that were predicted to regulate the biomarker genes (Figure A.2B, left). Based on the significance of the overlap, the most highly ranked transcription factor was ESR1 followed by PGR and ESR2, encoding the progesterone receptor and ER β , respectively. The top 15 canonical pathways enriched with the biomarker genes included “Estrogen Receptor Signaling” and pathways not directly linked to ER signaling (Figure A.2B, right). The complete set of IPA results are found in Table S2 and Table S3.

The biomarker genes were examined for evidence that they are direct targets of ER regulation using published ChIP-Seq experiments, which identified direct interactions between ER α and the promoter regions of biomarker genes in MCF-7 cells. Forty-nine out of 50 genes of the ER α biomarker had at least 72 datasets (7.4% of total datasets in the ChIP-Atlas database using the ER antibody) that exhibited a significant ER α binding site (MACS score >100 within 10 kB of a transcriptional start site of a biomarker gene). Twelve (24%) of these genes exhibited significant binding in over 50% of the identified datasets (KLF10, MAPT, RET, SNX24, PDZK1, CA12, RAPGEFL2, TPD53L1, IER3, IGFBP4, GREB1, and RARA). Only MYB did not contain a significant binding site. Of the

18,672 genes in the human genome analyzed, only 2.4% (443) of the genes exhibited significant ER α binding in over 50% of the ESR1 ChIP-seq datasets, therefore it is unlikely the biomarker genes exhibited binding regions by chance (Chi-square statistic with Yates correction of 89.5 with a p-value < 0.00001). Overall, the results of the ChIP-Seq studies indicate that all but the MYB gene are likely under direct transcriptional control of ER. The ChIP-Seq analysis is detailed in Table S4.

A.4.4 The biomarker identifies ER modulators in ER-positive breast cancer cell lines but not other cell lines

The ability of the 50-gene biomarker to identify conditions that are known or suspected to modulate ER was examined in a compendium of biosets derived from human cells including a large number of comparisons in MCF-7 cells. We also determined if the biomarker developed using MCF-7 cell experiments could be used in conjunction with gene expression profiling in other human cell lines. Biomarker behavior was first examined across E2-treated breast cancer cell lines with known ER activity. The biomarker was able to identify significant ER activation in 89 out of 91 biosets from E2-treated MCF-7 cells (Figure A.3A). The two biosets that did not exhibit activation were from cells treated with 10 nM E2 for 24 h (from GSE42619 and GSE48989). We examined 57 biosets from other human cells treated with E2 (Figure A.3B, left). All 5 biosets from 3 ER α positive cell lines (MDA-MB-134, SUM44PE, ZR-75-1) treated with E2 showed

significant activation. In contrast, only 3 of the 14 biosets from the ER α positive T47D cells exposed to E2 showed significant ER α activation. Although these cells are generally ER-responsive, the very high levels of progesterone receptor expressed in T47D cells may inhibit E2-induced gene expression [45,46]. It should be noted that this cell line was used in the ToxCast ER model to integrate E2-induced growth into the assay set [32]. None of the 4 biosets from E2-treated ER α -negative human breast cancer cell lines MDA-MB-231 and SK-BR-3 showed significant ER activation.

The biomarker was also evaluated to determine responsiveness in cell lines derived from tissues other than breast. ER activation was observed after E2 treatment in 5 of 11 biosets from endometrium-derived cell lines (Figure A.3B, middle). Most of these biosets were generated using the Ishikawa cell line, which expresses ER α and ER β [48]. None of the other biosets derived from 13 human cell types exhibited ER activation (Figure A.3C, right). Lack of activation could be due to either little or no expression of ER α /ER β , or that ER α is expressed but regulates a different set of genes in these tissues. Thus, the biomarker appears to be most consistently responsive to E2 in MCF-7 cells. With the exception of some ER-positive human breast cancer cell lines, the biomarker does not appear to be useful for screening chemicals in other cell lines.

The ability of the biomarker to identify treatments that would inhibit ER in a cell selective manner was determined. There were 16 biosets from MCF-7 cells treated with the anti-estrogen fulvestrant or the selective ER modulators bazedoxifene (BZA) or

raloxifene (RLF) (Figure A.3D, left). All but 2 exhibited $-\log(\text{p-value}) \geq 4$ that would indicate ER suppression. Treatment with fulvestrant or RLF in 4 other human non-breast cell lines (B-cell lymphoma, HL60, Ishikawa, PC3) in 11 biosets did not result in significant suppression. Additionally, the effect of hormone starvation on ER in MCF-7 cells was examined. Out of the 4 biosets from MCF-7 cells comparing hormone starvation to nonstarved cells, 3 exhibited significant ER suppression and the other approached significance (Figure S1).

In addition to E2 and classical ER agonists and antagonists, the effects of a large number of chemicals on ER modulation was examined by mining publicly available data. From the CMAP 2.0 study, there were 3442 biosets representing three cell lines (HL60, MCF-7, PC3) each exposed to ~1200 chemicals (Figure A.3E). None of the 2205 biosets from HL60 or PC3 cells exhibited significant ER activation or suppression. In contrast, there were 64 and 35 biosets that exhibited ER activation or suppression, respectively, in MCF-7 cells. In the DeAbrew et al. study [47], 34 chemicals were used to expose 4 cell lines (HepG2, HepaRG, Ishikawa, MCF-7) for a total of 407 biosets (Figure A.3F). None of the biosets from HepG2, HepaRG, or Ishikawa cells exhibited significant activation or suppression, while there were 36 and 3 biosets that exhibited significant activation or suppression of ER, respectively, in MCF-7 cells. These studies indicate that the 50-gene ER biomarker is potentially useful for screening chemicals in MCF-7 cells but not other cell types.

A.4.5 The biomarker identifies genetic perturbations of the ESR1 gene

Given that the biomarker was built based in part on genetic perturbation of the ESR1 gene, we determined whether or not the biomarker could identify perturbations of the ESR1 gene itself. We identified 16 and 6 biosets in the database in which ER was predicted to be activated or suppressed, respectively. (These biosets did not include those used to create the biomarker.) The biosets predicted to activate ER included overexpression of wild-type ESR1 or constitutively active mutants of ESR1 that possessed mutations in the hinge or ligand binding domain. Out of the 16 tested, 11 exhibited ER activation and one bioset approached significance ($-\log(p\text{-value}) = 3.52$) (Figure A.4, left). All of the biosets that did not exhibit significant ER activation were from one study (GSE51448). One of the mutants that did not have significant ER activation was S463P; in the original study, this mutant exhibited lower constitutive ER activation than the other mutants tested [49]. Gene expression was also examined in MCF-7 cells after the ESR1 gene was knocked down using siRNA or shRNA methods. Out of the 6 biosets, 4 exhibited significant suppression of ER and one approached significance ($-\log(p\text{-value}) = -3.15$) (Figure A.4, right). These findings indicate that the biomarker may be a useful tool to identify genes that when over- or under-expressed (or when mutated) affect the activity of ER.

A.4.6 Determination of the predictive accuracy of the ER biomarker

The biomarker was evaluated for the ability to identify chemicals with known ER activity. In the following studies, we focused only on those microarray comparisons generated in MCF-7 cells. A list of ER reference chemicals was taken from the OECD TG457 BG1 guidance document [50]. Of the 45 reference chemicals listed in the document, there were 19 chemicals represented by 122 biosets in the MCF-7 compendium. The chemicals examined included 2 strong agonists (E2, diethylstilbestrol), 2 moderate agonists (alpha-estradiol, estrone), 4 weak agonists (bisphenol A, bisphenol B, daidzein, genistein), 4 very weak agonists (4-nonylphenol, apigenin, chrysin, kaempferol), and 7 inactives or nonactivators (corticosterone, cycloheximide, flutamide, haloperidol, ketoconazole, reserpine, spironolactone). For each class, the range of $-\log(p\text{-values})$ of the biosets were examined (Figure A.5A). In general, the median $-\log(p\text{-value})$ s for the different classes followed the order: strong > moderate ~ weak > very weak >> inactive. The order indicates that the significance of the correlation can be a rough approximation of the strength of the chemical as an ER agonist. However, this is dependent on whether the concentration of the chemical examined in the study achieved maximum activation of ER.

To determine if the strength of the chemical as an ER activator would be reflected in comparing concentration-response between chemicals, the relationships between $-\log(p\text{-value})$ and concentration for 4 chemicals were examined in MCF-7 cells exposed

for 48 h (from GSE50705 [51]) (Figure A.5B). Each of the 4 chemicals exhibited concentration-dependent increases in the correlation as reflected by increases in the $-\log(\text{p-value})$ s with concentration of chemical. While the two weak agonists (BPA and genistein) exhibited similar concentration response, our approach stratified the two strong agonists showing that DES was clearly more potent than E2.

The biomarker predictions were compared to those of the ToxCast ER model [32]. For 112 overlapping chemicals examined in 15 studies, Figure A.5C shows the comparison between the $-\log(\text{p-value})$ s from the 50-gene ER biomarker and the ToxCast ER model AUC. In this analysis we are assuming that the results of the ToxCast ER model are the truth to which the predictions of the ER biomarker are compared. Most of the compounds (79) had no activity as assessed by both methods (i.e., AUC of < 0.05 and $|\log(\text{P value})| < 4$) (Figure A.5C, black points). Both approaches identified 18 chemicals that were classified as ER activators (Figure A.5C, green points). There were 9 chemicals that were predicted to activate ER in MCF-7 cells as assessed using the biomarker but were inactive in the ToxCast ER model (Figure A.5C, green points) and included 4-hydroxytamoxifen, 4-nonylphenol, farnesol, flutamide, methotrexate, oxybenzone, retinoic acid, tamoxifen, and theobromine. There were 6 chemicals that were active in the ToxCast ER model but not in the 50-gene ER model (Figure A.5C, yellow points). The chemicals included apigenin, chrysin, nordihydroguaiaretic acid, prednisone, propylthiouracil, and quercetin. Using the ToxCast ER model as the reference dataset, an accuracy test for predictions of activation using the ER 50-gene biomarker gave a

sensitivity of 0.75, a specificity of 0.90, a positive predictive value of 0.67, a negative predictive value of 0.93, and an overall balanced accuracy of 0.82. The biosets used in the analysis are described in Table S5.

Differences in the predictions from the ToxCast ER model and those using the biomarker were examined in greater detail. The false positives were those 9 chemicals that had activity predicted by the biomarker but were inactive in the HTS ER model. Tamoxifen and its metabolite 4-hydroxytamoxifen (4-HT) were found to be activators of ER α ; both chemicals can act as partial agonists in some cellular contexts [52]. We had previously shown that oxybenzone activates ER in trans-activation assays [22]. There is evidence that 4-nonylphenol activates ER [53]. The compounds farnesol, flutamide, methotrexate, and retinoic acid were all from one previous study [47]. We previously found that flutamide and other compounds examined in this study (phenobarbital, di-2-ethylhexyl phthalate, clofibrate) were positive using the 46-gene biomarker, but in follow-up independent studies and using RNA-Seq studies and the biomarker approach, we could not confirm their estrogenicity [22]. Further work is necessary to determine if these compounds are true positives or not. If these 4 compounds from the De Abrew et al. study are removed as well as the compounds that are known to activate ER in some contexts but not others (tamoxifen, 4-HT, 4-nonylphenol, oxybenzone), there is one false positive (theobromine) resulting in an accuracy of 0.87.

All of the false negatives were from the CMAP 2.0 study evaluated at one concentration that ranged from 1 to 23.4 μM , clearly below the maximum concentration used in the ToxCast ER model assays. Apigenin and chrysin approached the statistical cutoff ($-\log(\text{p-value}) = 2.3$ and 2.1 , respectively). It should be noted that there are additional weaknesses in the use of the CMAP 2.0 data set for this application. Statistically-significant gene lists were derived using a t-test comparison between one treated replicate vs. multiple control replicates at only one time point and one dose level ($\sim 10\text{--}20\ \mu\text{M}$ for most chemicals). As described in a recent paper, 5 strong ER activators from the CMAP collection were identified using the 46-gene biomarker approach but could not be confirmed using other methods [22]. Using an approach encompassing multiple replicates of a full concentration-response spanning up to cytotoxic concentrations is most appropriate as an initial screen to identify potential positives that can be more completely analyzed in follow-up studies. This approach has been implemented by EPA [6], as well as Health Canada [37] and the National Toxicology Program ([54]).

The 50-gene ER biomarker results were compared to the results of a screen conducted by the National Center for Advancing Translational Sciences (NCATS) that examined the ER α activity of $\sim 10,500$ chemicals (~ 8300 unique) in the Tox21 library of compounds. The screen was carried out in VM7 cells containing a stably integrated luciferase gene under control of 4 concatemerized EREs (ER-luc) [35]. ER α agonist assays examined in the absence or presence of 10 nM fulvestrant were measured in separate assays after 18–24 h of treatment. Positives and negatives from the Tox21

assays were defined as those that exhibited consistent activity in both assays. After filtering for chemicals that overlapped with those in the MCF-7 compendium, 91 chemicals in 168 biosets from 21 studies were examined. The results of the comparisons are shown in Figure A.5D. Using the Tox21 data as the reference data set, the 50-gene ER biomarker had a sensitivity of 0.93, a specificity of 0.98, a positive predictive value of 0.97, a negative predictive value of 0.96 and a balanced accuracy of 0.96. Because there were multiple biosets that examined the same chemical, which biases the results, we reevaluated the sensitivity predictions based on a single bioset representing one chemical. If there were multiple biosets representing one chemical, the bioset with the highest $-\log(p\text{-value})$ was selected for the analysis. Using this approach there were 91 chemicals in 91 biosets from 21 studies. Again, using the Tox21 data as the reference data set, the 50-gene ER biomarker had a sensitivity of 0.87, a specificity of 0.97, a positive predictive value of 0.87, a negative predictive value of 0.97 and a balanced accuracy of 0.92. There were 2 false positives (pentetic acid, methotrexate) and 2 false negatives (ethisterone, apigenin) both of which approached significance ($-\log(p\text{-value}) = 3.0$ and 2.3 , respectively) indicating if the concentration used in the microarray study was increased, then statistical significance may be reached. Thus, the ability of the 50-gene biomarker to identify ER agonists in microarray comparisons was excellent. The biosets used in the analysis are described in Table S6.

A.4.7 Application of the biomarker to predictions in HTTr data: identification and ranking of BPA alternatives

The 50 gene ER biomarker was used to assess the estrogenicity of BPA and 15 BPA alternative chemicals (Table A.2) in MCF-7 cells exposed to the chemicals for 48 h. Quality control/assurance metrics and transcriptomic analyses from this study were published previously [37]. The biomarker correctly identified the positive control (17 β -estradiol) as ER active and the negative control (dexamethasone) as inactive. Of the 16 chemicals, the biomarker identified nine as ER active (Table A.3) including BPAF, BPC, BPA, 4,4'-BPF, BPAP, BPS, 2,4'-BPF, P201, and 2,4'-BPS. These chemicals were identified as ER active if they had at least one concentration called significant ($-\log(p\text{-value}) \geq 4$) by the Running Fisher Test.

Two approaches were used to derive transcriptomic points of departure (tPODs) for potency ranking of chemicals for ER α activation (Figure A.6A): a) median BMC of biomarker genes with bootstrapped confidence intervals; and b) BMC modeling of Normalized Enrichment Scores (NES) for the biomarker gene set. The order of potency of the median BMC with bootstrapped confidence interval was (from most to least potent): BPAF, BPC, BPA, 4,4'-BPF, BPAP, BPS, 2,4'-BPF, P201, 2,4'-BPS. The order of potency from the NES BMC modeling was (from most to least potent): BPAF, BPA, BPC, 4,4'-BPF, BPAP, BPS, 2,4'-BPF, P201, 2,4'-BPS. The BMCs derived from NES modeling were generally more conservative, in some cases (e.g., BPA and BPS) by one order of

magnitude. Overall, these two approaches produced almost identical potency rankings and were highly correlated (Figure A.6B).

The main difference between the transcriptomic points of departure was that the median ER biomarker BMC ranked BPC as the second most potent chemical tested, whereas BPA was the second most potent using the BMC of NES approach. However, the confidence intervals on the BMCs were similar and overlapping between the two approaches, indicating no significant differences in potency calls. The exception to this was BPA. Here, the NES approach yielded a very steep slope at low concentrations that did not fit the BMC models well, leading to very large confidence intervals (Figure S2). Thus, there is less certainty overall in the BMC calculated for BPA using the NES approach.

A.5 Discussion

Profiling chemicals by HTTr provides an alternative to typical HTS assays, such as those run in the ToxCast screening program [4], and is also proposed as a potential replacement for some animal tests that are used to identify chemical targets important to the EDSP [55]. In our past studies, we developed [21] and applied [22,56,57] [55] a 46-gene expression biomarker that is highly predictive of ER activity in gene expression profiles generated by microarrays. This biomarker was built using profiles from MCF-7 cells treated with 7 agonists and 3 antagonists. However, preliminary studies in our lab

indicated that the 46-gene biomarker was not useful to interrogate HTTr data generated using the TempO-Seq platform. Additionally, our group has shown that building gene expression biomarkers using both chemical biosets as well as an approach that takes advantage of genetic manipulations in which the target gene itself is modulated in expression or function results in biomarkers with improved predictivity and specificity.

In the present study, we describe the development and characterization of a 50-gene biomarker that could be used to screen large HTTr datasets to identify ER actives that follows these improved practices for biomarker derivation. Like our previous effort with the 46-gene ER biomarker, the 50-gene ER biomarker includes genes that are regulated in a consistent manner in microarray biosets by ER chemical agonists and antagonists. Unlike the 46-gene biomarker, the 50-gene biomarker was built using a genetic strategy utilizing biosets in which either the wild-type ESR1 or constitutively active ESR1 mutants were overexpressed or in which ESR1 expression was knocked down. Even though there was an overlap in the chemicals used to create both biomarkers, there was only a 13-gene overlap. Our dual genetic and genomic approach to identify genes that could be used to predict ER activity is similar to past approaches used in identifying other characterized biomarkers. For example, the gene expression biomarker predictive of identifying chemicals that activate or suppress the androgen receptor (AR) utilized biosets in which AR expression was knocked down or AR was constitutively active. These genetic modulations along with biosets in which AR was chemically activated or suppressed resulted in a biomarker with high predictive accuracy [58]. Other biomarkers

that predict effects in human cells that incorporated a genetic strategy included those that predict modulation of NFkB [16], Nrf2 [14] and heat shock factor 1 [15]. This integration of genetic effects into biomarker construction was used in earlier efforts to build biomarkers that predict transcription factor activity in the mouse liver; the predictive accuracies for 6 biomarkers derived from these liver experiments were all greater than 91% [29,30,[58], [59], [60], [61]]. Thus, while accurate biomarkers can potentially be developed using only chemical comparisons (such as the 46-gene ER biomarker), incorporating a genetic strategy provides additional confidence that the genes identified and used in the biomarker comprise a gene set that is specific for the factor being predicted. Given the need for a large set of well characterized gene expression biomarkers to interpret HTTr data, future studies could be focused on utilizing batteries of cell lines knocked out or knocked down for the (transcription) factor to be predicted. This “bottom-up” strategy to build biomarkers with high predictive accuracy would complement the “top-down” approach that is currently being used in most studies to link gene expression changes with chemical hazards.

As in our other biomarker studies, we used a performance-based approach to evaluate specificity and sensitivity of the 50-gene ER biomarker and compared performance to existing EDSP test methods used for or considered for regulatory decision-making, i.e. the ToxCast ER model based on 18 HT assays [32]. We were able to take advantage of multiple sets of well-studied ER reference chemicals examined for gene expression changes in MCF-7 cells that are found in BSCE. Our analyses focused

on ER agonism, as most ER-active environmental chemicals have been identified as agonists and is consistent with the EDSP Tier 1 battery of assays that only are capable of detecting agonism (summarized in [32]). Across a range of structures and potencies, the 50-gene ER biomarker was able to roughly stratify ER activity based on the significance of the correlation to a set of 19 OECD chemicals including those that were classified as “weak” and “very weak”. The biomarker predictions were compared to those of the ToxCast ER model [32]. For 112 overlapping chemicals examined in 15 studies, the 50-gene biomarker had a predictive accuracy of 82–87% depending on what chemicals were included in the test set. Lastly, comparing the predictions to chemicals with consistent results from two trans-activation assays run as part of the Tox21 chemical screening program, the predictive accuracies were 96%. Because some chemicals were represented by more than one bioset, accuracy was recalculated using only one bioset per chemical resulting in a predictive accuracy of 92%. Overall, the 50-gene biomarker was excellent at identifying ER chemical activators. It should also be noted that the biomarker could identify conditions in which the ESR1 gene itself was perturbed. The biomarker exhibited significant correlation to 16 out of 22 of the biosets in which ESR1 was genetically activated or suppressed. Thus, the 50-gene biomarker could be used as a screening tool to identify genes that affect ER activity.

Given the increasing use of targeted RNA-seq data as a method to interrogate gene expression, we sought to determine whether our approach using the 50-gene ER biomarker could accurately identify ER active chemicals in a dataset generated in cultured

breast cancer cells and analyzed using TempO-Seq. The dataset consisted of MCF-7 cells exposed to BPA and 15 alternative chemicals across 10 concentrations for 48 h, as well as positive and negative controls [37]. Most BPA alternative chemicals selected were structurally homologous to BPA (i.e., a bisphenol), with some differences in aryl substituents and/or adjoining atom, as well as two sulphonyl urea alternative chemicals (BTUM and P201). The 50-gene biomarker identified BPA and eight alternative chemicals as ER active, consistent with previous results using the 46-gene biomarker [37]. The results were generally in line with previous publications (see [62,63]. However, P201 and 2,4'-BPS have previously been reported inactive for ER transactivation [63] and, to our knowledge, 2,4'-BPF has no reported ER activation ability. Further, our work demonstrates how transcriptional points of departure (tPODs) can be derived for potency ranking by using either the median gene BMC or the BMC of modeled NES for the biomarker genes. These two analyses demonstrated nearly identical potency rankings, although the latter was generally more conservative. While these data support that the 50-gene ER biomarker can be used to identify and compare ER activity of chemicals in HTTr screens, the current focus of the lab is to establish the performance of the biomarker in larger datasets including hundreds of possible ER active chemicals. Overall, the results indicate the value of biomarkers in general for rapidly screening transcriptomic data sets to identify chemical mechanisms of action and for potency analyses.

Our approach to identifying ER modulators has a number of limitations typical of gene expression biomarkers that have been previously characterized. In comparisons of

the biomarker to biosets derived from various cell lines treated with E2 or ER antagonists/SERMs, we found that the biomarker can only be used in certain cell lines that contain a functional ER. For example, the biomarker could identify E2 treatments in MDA-MB-134, SUM44PE, and ZR-75-1 cells but not consistently in another human breast cancer cell line, T47D cells. The biomarker also did not appear to be useful for screening in 14 other cell lines tested, including Ishikawa cells. It appears that to identify treatments that activate ER in other cell lines, biomarkers would have to be developed de novo from profiles in the cell line in which screening would take place. This would potentially allow the identification of SERMs that would not necessarily be identified in MCF-7 cells. For example, while tamoxifen is an ER α antagonist in MCF-7 cells, tamoxifen is an ER α agonist in the uterus and derived cell lines. Thus, it is envisioned that to more fully identify ER modulators, a battery of gene expression biomarkers would have to be created to identify those chemicals that induce cell context-dependent activity.

One of the most formidable challenges, though not unique to our system, is the fact that MCF-7 cells have limited capacity to identify chemicals that may be biotransformed to active or inactive metabolites. Inclusion of immobilized metabolic enzymes derived from hepatic S9 fractions in alginate microspheres attached to 96-well peg lids and then coupled to an ER trans-activation assay decreased the number of false positives and false negatives [64]. While the 50-gene ER biomarker model can readily identify chemicals that modulate ER through classical binding mechanisms, it is not known if the method can identify chemicals that activate ER indirectly through growth

factor or cytokine-dependent mechanisms [65]. Lastly, it should be noted that even though chemicals are identified that activate ER, the chemicals should be further characterized for human relevance including assessment of absorption, distribution, metabolism and excretion and potency; using in vitro to in vivo extrapolation models can facilitate determination of whether ER could be activated in human tissues given presumably low levels of exposure.

In conclusion, the 50-gene ER biomarker model accurately predicts ER bioactivity of in vitro reference chemicals identified through ToxCast or Tox21 screening approaches. This transcriptomic biomarker model could potentially be used as an alternative to the EDSP Tier 1 ER binding and ER trans-activation assays. While the 46-gene biomarker was able to reasonably replicate the results of the uterotrophic assay [55], the 50-gene biomarker would have to be characterized relative to the same chemical reference set. The ER biomarker approach has an advantage over the ToxCast ER model in that ER activity can be assessed in one assay versus numerous (between 4 and 18 depending on the model) and the ability to couple analyses with other characterized biomarkers. It will be important to repeat our findings using a HTTr dataset with a greater number of chemicals, performed in concentration-response format using an (targeted) RNA-Seq approach generated using a similar set of conditions unlike the heterogeneous set of studies used in the present analysis. Hopefully, these conditions will allow calibration of the model to incorporate the degree of correlation as an indirect measure of the strength of ER activation, currently not possible with a one chemical – one

concentration approach that represented most of the chemicals in the compendium. A dose-response analysis of 4 ER agonists clearly showed a concentration-dependent increase in correlation that paralleled the strength of ER activation assessed by a trans-activation assay [51]. As detailed in a United States Federal Register Notice (<https://www.federalregister.gov/articles/2015/06/19/2015-15182/use-of-high-throughput-assays-and-computational-tools-endocrine-disruptor-screening-program-notice>), four assays in the EDSP Tier I battery could potentially be replaced by in vitro ER assays based on the ability of the assays used in the Judson et al. model [66] to accurately predict uterotrophic results in mice and rats [32]. Thus, the methods developed here could not only be used as a more streamlined alternative to the ER ToxCast assays but could also provide a general strategy for the identification of ER modulators that would meet the needs of a number of EDSP stakeholders. This approach will help EPA to reduce reliance on animal-based assays and identify chemicals with the greatest potential for ER bioactivity that may be high priority candidates for further screening and testing.

A.6 Tables and Figures

Table A.1: Biosets used to create the 50-gene ER biomarker.

Bioset Name From BSCE	Category	Factor	Hours of Chemical Treatment	Public id
Breast adenocarcinoma MCF7 cells + 6hr 10uM bisphenol A vs DMSO	Chemical Agonist	Bisphenol A	6	GSE69851
Breast cancer MCF7 cells control siRNA 48hr - 1 nM estradiol additional 4hr _vs_ vehicle	Chemical Agonist	Estradiol	4	GSE57935
MCF7 cells + estrone, 14.8uM vs DMSO vehicle	Chemical Agonist	Estrone	6	GSE5258
Breast cancer MCF7 cells treated with 25uM p-nonylphenol for 48hr _vs_ controls	Chemical Agonist	Nonylphenol	48	GSE50705
MCF-7 breast cancer cells +24hr 100 nM bazedoxifene _vs_ ethanol vehicle controls	Chemical Antagonist/SERM	Bazedoxifene	24	GSE35428
MCF7 cells + fulvestrant, 0.01µM vs DMSO vehicle	Chemical Antagonist/SERM	Fulvestrant	6	GSE5258
MCF-7 breast cancer cells +24hr 100 nM ICI 182 780 _vs_ ethanol vehicle controls	Chemical Antagonist/SERM	ICI 182780	24	GSE35428
MCF7 cells + raloxifene, 7.8uM vs DMSO vehicle	Chemical Antagonist/SERM	Raloxifene	6	GSE5258
Breast MCF7/BUS cells 16hr with DMSO - overexpressing <i>ESR1</i> wildtype vs control vector	<i>ESR1</i> Activation	Wild-type		GSE115564
Breast cancer MCF7 cells 24hr untreated - <i>ESR1</i> D538G knock-in vs wildtype	<i>ESR1</i> Activation	D538G		GSE112243
Breast cancer MCF7 cells normal media - expressing ER S436 mutant 48hr vs GFP	<i>ESR1</i> Activation	S436P		GSE51448
Breast cancer MCF7 cells treated 24hr with vehicle - <i>ESR1</i> mutant (Y537S) _vs_ wildtype controls	<i>ESR1</i> Activation	Y537S		GSE89888
MCF7 breast cancer cell line <i>ESR1</i> shRNA _vs_ control shRNA	<i>ESR1</i> Suppression	<i>ESR1</i> shRNA		GSE18431

MCF7 breast cancer cells transfected 72hr with estrogen receptor a siRNA _vs_ untransfected	<i>ESR1</i> Suppression	<i>ESR1</i> siRNA		GSE10061
Breast cancer MCF7 cell line - <i>ESR1</i> siRNA _vs_ control siRNA	<i>ESR1</i> Suppression	<i>ESR1</i> siRNA		GSE37820
MCF7 with siRNA disrupted <i>ESR1</i> _72hr _vs_ siRNA controls	<i>ESR1</i> Suppression	<i>ESR1</i> siRNA		GSE10890

Table A.2: Chemicals examined in HTTr data from the Matteo et al. study [37].

Chemical Name	Abbreviation	CASRN	DTXSID
2,2-Bis(4-hydroxyphenyl)propane	BPA	80-05-7	DTXSID7020182
Bis(4-hydroxyphenyl)methane	4,4'-BPF	620-92-8	DTXSID9022445
1,1-Bis(4-hydroxyphenyl)-1-phenylethane	BPAP	1571-75-1	DTXSID5051444
4,4'-Sulfonyldiphenol	BPS	80-09-1	DTXSID3022409
2,2-Bis[4-(glycidyoxy)phenyl]propane	BADGE	1675-54-3	DTXSID6024624
4,4'-Dichlorodiphenyl sulfone	DDS	80-07-9	DTXSID9024986
4,4'-(Hexafluoroisopropylidene)diphenol	BPAF	1478-61-1	DTXSID7037717
Dimethyl sulfoxide	DMSO	67-68-5	DTXSID2021735
17 β -estradiol	E2	50-28-2	DTXSID0020573
Dexamethasone	Dex	50-02-2	DTXSID3020384
4-[[4-(2-Propen-1-yloxy)phenyl]sulfonyl]phenol	BPS-MAE	97042-18-7	DTXSID10629100
4,4'-Bis(p-toluenesulfonylamino)carbonylamino)diphenylmethane	BTUM	15188-2-81-4	DTXSID50872911
4-((4-Isopropoxyphenyl)sulfonyl)phenol	D-8	95235-30-6	DTXSID2058706
4-Methyl-N-[[[3-[(4-methylphenyl)sulfonyl]oxy]phenyl]amino]carbonyl]benzenesulfonamide	Pergafast201 (P201)	23293-8-43-1	DTXSID90872910
2,4'-Dihydroxydiphenyl sulfone	2,4'-BPS	5397-34-2	DTXSID9052183
4-(4-hydroxy-3-prop-2-enylphenyl)sulfonyl-2-prop-2-enylphenol	TGSA	41481-66-7	DTXSID9047598

2,2-Bis(4-hydroxy-3-methylphenyl)propane	BPC	79-97-0	DTXSID8047890
2,4'-Dihydroxydiphenylmethane	2,4'-BPF	2467-03-0	DTXSID2044890
4-((4-(Benzyloxy)phenyl)sulfonyl)phenol	BPS-MPE	63134-33-8	DTXSID4058655

Table A.3: Benchmark concentration modeling and ER α predictions using the 50 gene transcriptomic biomarker.

Chemical	ER α Activation Prediction	Genes fitting BMCs	Bootstrapped Median BMC	BMCL	BMCU	BMC of NES	BMCL	BMCU
BPAF	Active	31	0.063	0.023	0.092	0.014	0.0061	0.032
BPA	Active	11	0.42	0.12	0.92	0.041	0.00012	0.13
BPC	Active	34	0.17	0.092	0.42	0.11	0.052	0.21
4,4'-BPF	Active	36	0.47	0.17	1.1	0.34	0.17	0.73
BPAP	Active	21	0.65	0.25	1.9	0.46	0.31	0.91
BPS	Active	33	1.4	0.65	4.3	0.77	0.4	1.6
2,4'-BPF	Active	23	2.6	0.88	6.2	3.1	2.3	4.7
P201	Active	23	6.8	2.1	47	12	8.6	22
2,4'-BPS	Active	14	22	2.4	95	37	28	56
BADGE	Inactive							
D-8	Inactive							
TGSA	Inactive							
BPS-MAE	Inactive							
BPS-MPE	Inactive							
BTUM	Inactive							
DDS	Inactive							

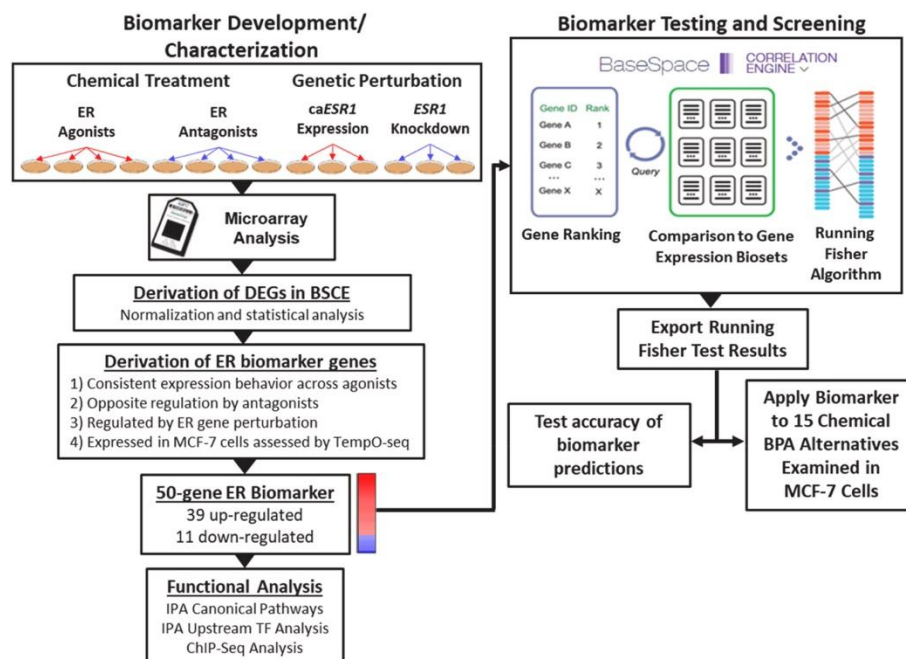


Figure A.1: Biomarker development and strategy to identify ER active compounds in HTTr datasets. Left, ER biomarker development. Differentially expressed genes (DEGs) from exposures of MCF-7 cells to known ER activators, ER suppressors including one anti-estrogen and two SERMs, and genetic perturbations that lead to increased or decreased expression/activity of ESR1 were used to identify biomarker genes identified as those consistently up- or down-regulated by the chemical exposures or genetic perturbations including constitutively active ER (caER) mutants. Post hoc analysis of genes in the biomarker was performed using Ingenuity Pathway Analysis (IPA) for canonical pathway enrichment and potential upstream transcription factor regulators. Identification of primary targets of ER α was carried out through queries of chromatin immunoprecipitation coupled with DNA sequencing (ChIP-Seq) studies of ER α targets. Right, biomarker testing and use in screening. The biomarker was compared to microarray datasets derived from MCF-7 cell line studies archived in a human microarray compendium to determine predictive accuracy for identifying ER activators. The biomarker was also compared to a HTTr dataset to determine the potency of bisphenol A alternatives as ER activators.

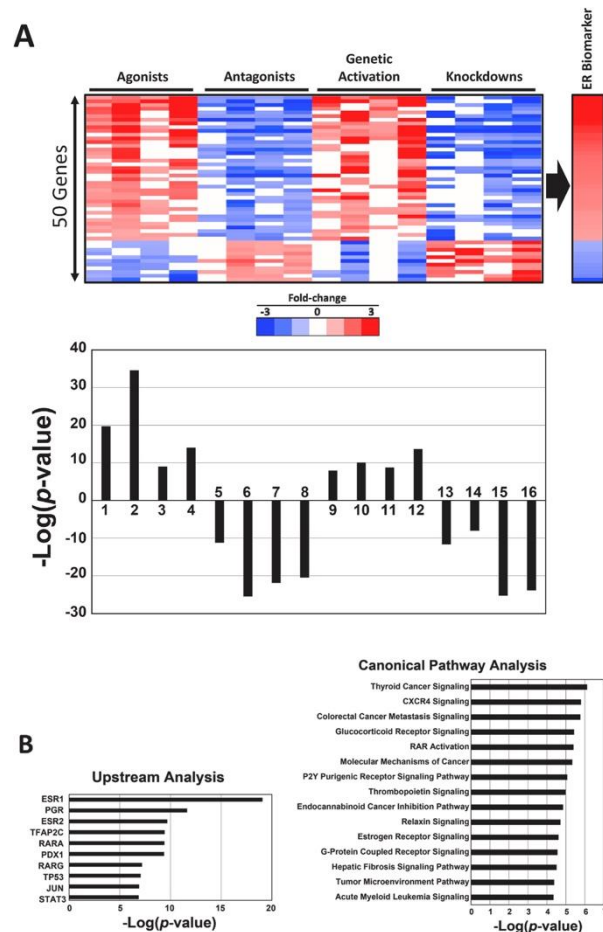


Figure A.2: Assembly and characterization of a 50-gene ER biomarker. A. Assembly of the ER biomarker. (Top) Lists of statistically significant genes from MCF-7 cells after chemical or genetic perturbations listed in Table A.1 were filtered as described in the Methods. The expression of the genes with consistent behavior across the perturbations are shown. Fold-change values for each gene were averaged across the 4 agonist treatments to give fold-change values for each gene in the biomarker (B). The resulting ER biomarker consisted of 50 genes. (Bottom) Assessment of the correlation between the ER biomarker and the 16 biosets used to create the biomarker. The ER biomarker was compared with the biosets using the Running Fisher test [8]. The significance of the correlation is indicated by the $-\log(p\text{-value})$. Biosets with positive or negative correlation to the biomarker are indicated as positive or negative numbers, respectively, and would be predicted to be those that activate or suppress ER, respectively. B. Ingenuity Pathway Analysis of the ER biomarker genes. (Left) Transcription factors predicted to regulate the biomarker genes using the upstream regulator analysis. (Right) Canonical pathways that significantly overlap with the biomarker genes. Biomarker genes were compared to the genes in the canonical pathway lists from IPA.

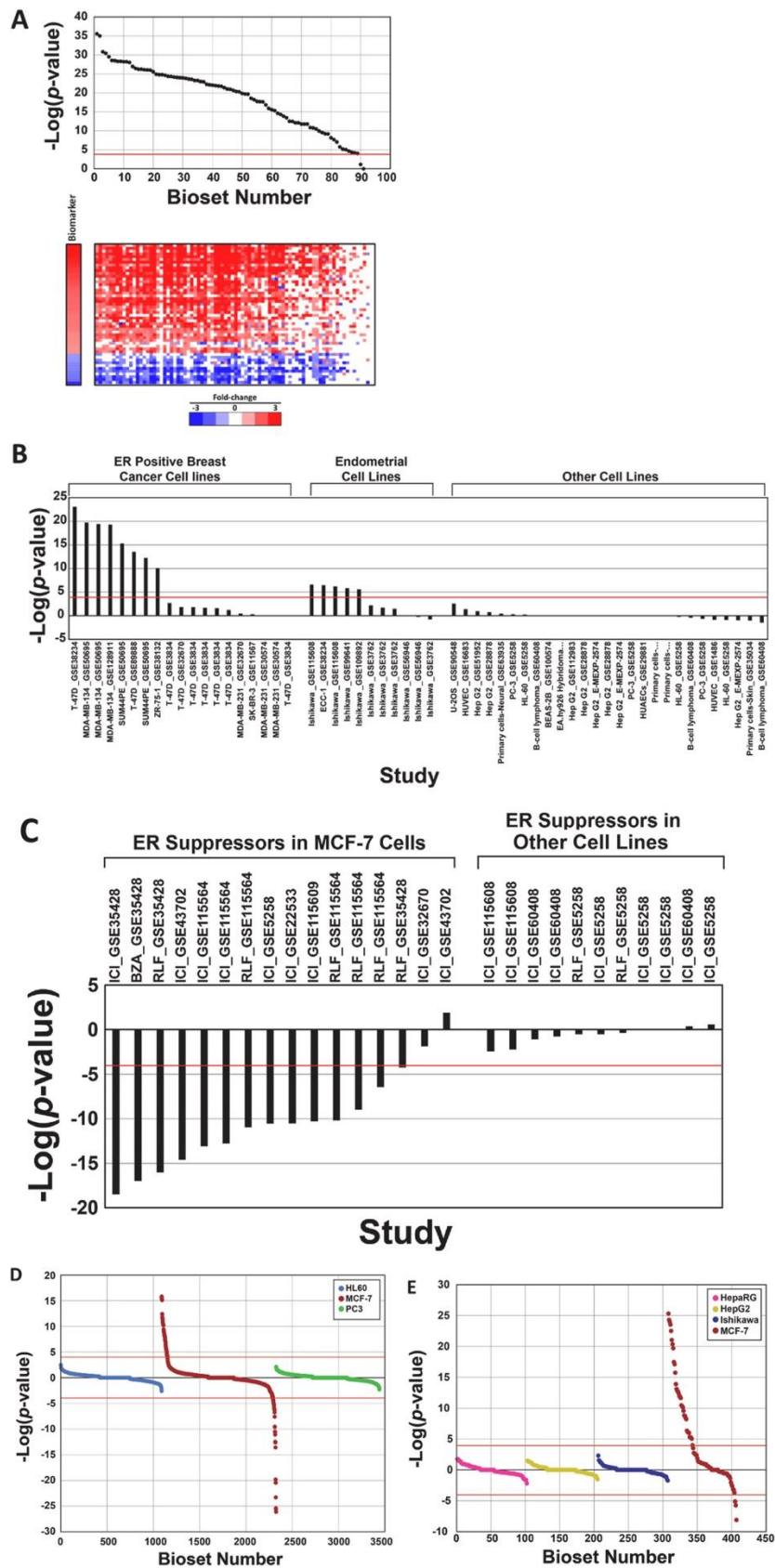
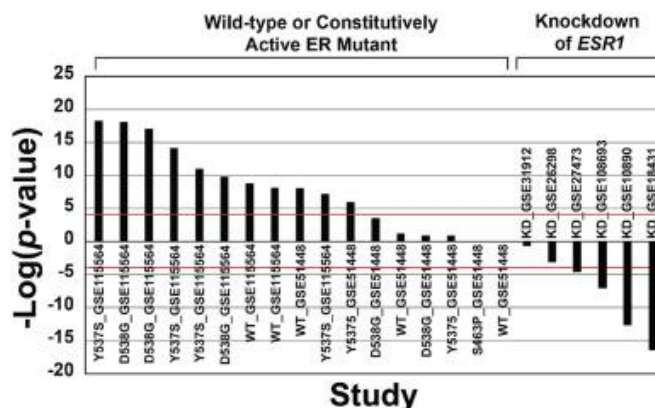


Figure A.3: The biomarker identified ER active chemicals in ER positive breast cancer cells. A. The biomarker identified E2 treatments in MCF-7 cells. The biomarker was compared to 91 biosets in which MCF-7 cells were treated with E2. The top panel shows the $-\text{Log}(p\text{-value})$ s across the comparisons, and the bottom panel shows the expression changes of the biomarker genes after exposure to E2. B. The biomarker identified E2 treatments in ER-positive human breast and endometrial cancer cell lines but not other cell lines. The biomarker was compared to biosets from E2 treatments from the indicated cell lines. The GSE accession number refers to the archived dataset in Gene Expression Omnibus. C. The biomarker identified ER antagonists/SERM comparisons in MCF-7 cells but not other cell lines. The biomarker was compared to biosets in which the cells were treated with the indicated antagonist or SERM in MCF-7 cells or other cell lines. The GSE accession number refers to the archived dataset in Gene Expression Omnibus. ICI, fulvestrant; BZA, bazedoxifene; RLF, raloxifene. D. Comparison of the biomarker to the CMAP dataset. The 3442 biosets representing ~ 1200 chemicals examined in three cell lines is shown. Only MCF-7 cells exhibited $-\log(p\text{-value}) \geq 4$. E. Comparison of the biomarker to the De Abrew et al. [47] dataset. The 407 biosets representing 34 chemicals examined in four cell lines is shown. Only MCF-7 cells exhibited $-\log(p\text{-value}) \geq 4$.



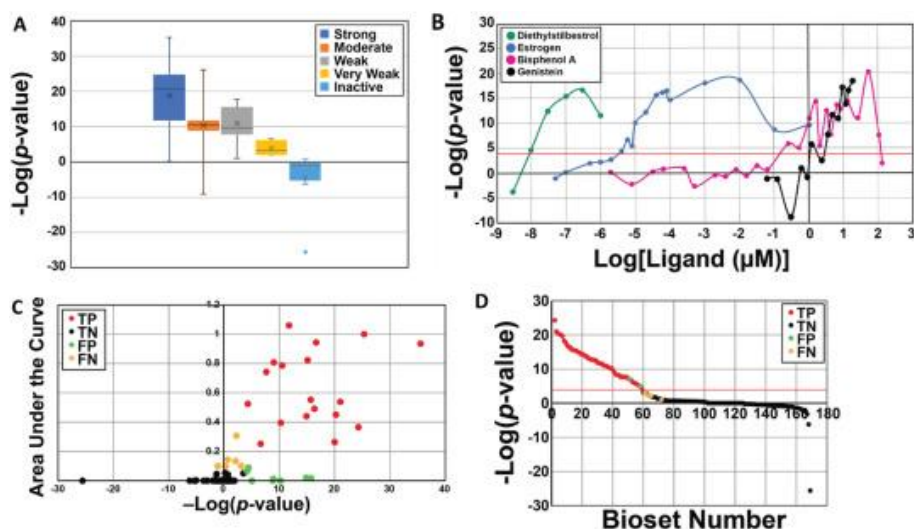


Figure A.5: Predictive accuracy of the ER biomarker. A. The range of $-\log(p\text{-value})$ s for chemicals classified as ER agonists by OECD. There were 19 chemicals represented by 122 biosets from the MCF-7 compendium that were examined. The figure shows the box-and-whisker plots for the range of $-\log(p\text{-value})$ s for each category of ER agonist. B. Relationships between concentration of 4 ER agonists and significance of the correlation to the 50-gene biomarker. Four chemicals were examined for transcriptional changes (from GSE50705). The figure shows that the significance of the correlation increases with increasing concentration of the chemicals as reflected by increases in $-\log(p\text{-value})$ s. C. Comparison of 50-gene ER biomarker $-\log(p\text{-value})$ s to predictions from the ToxCast ER model. For 112 chemicals, the significance of the correlations from the 50-gene ER biomarker were plotted against the area under the curve (AUC) derived from the ToxCast ER model. The chemicals that fall into different categories are indicated. TP, true positives; TN, true negatives; FP, false positives; FN, false negatives. D. Comparison of the 50-gene ER biomarker $-\log(p\text{-value})$ s to results from the Tox21 ER trans-activation screen. The overlapping chemicals were classified as shown.

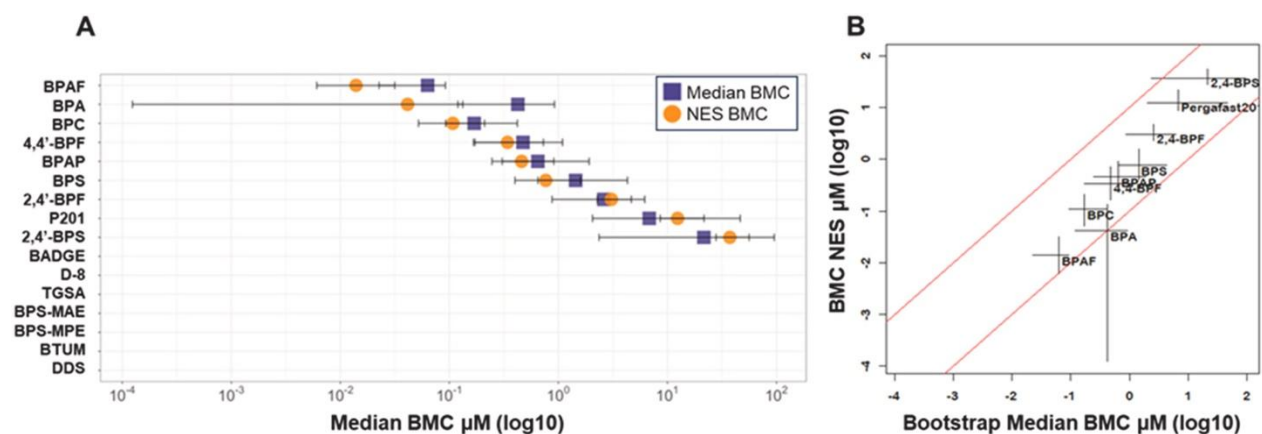


Figure A.6: Comparison of tPODs derived from the 50 gene ER biomarker. MCF-7 cells ($n = 3\text{--}4$ per concentration) were exposed to BPA and 15 alternative chemicals at a range of concentrations ($0.0005\text{--}100\text{ }\mu\text{M}$) for 48 h. The chemicals were rank ordered from most (BPAF) to least (2,4'-BPF) potent based on the NES BMC. A. Bootstrapped median benchmark concentrations with confidence intervals (BMCL and BMCU) and benchmark concentration of normalized enrichment scores with confidence intervals (BMCL and BMCU). B. Logarithmic scatterplot with red 1:1 line.

Appendix B: From vision toward best practices: Evaluating *in vitro* transcriptomic points of departure for application in risk assessment using a uniform workflow

Modified from: Reardon AJF, Farmahin R, Williams A, Meier MJ, Addicks GC, Yauk CL, Matteo G, Atlas E, Harrill J, Everett LJ, Shah I, Judson R, Ramaiahgari S, Ferguson SS, Barton-Maclaren TS. From vision toward best practices: Evaluating *in vitro* transcriptomic points of departure for application in risk assessment using a uniform workflow. *Frontiers of Toxicology*, 2023 May 23;5:1194895.

B.1 Abstract

The growing number of chemicals in the current consumer and industrial markets presents a major challenge for regulatory programs faced with the need to assess the potential risks they pose to human and ecological health. The increasing demand for hazard and risk assessment of chemicals currently exceeds the capacity to produce the toxicity data necessary for regulatory decision making, and the applied data is commonly generated using traditional approaches with animal models that have limited context in terms of human relevance. This scenario provides the opportunity to implement novel, more efficient strategies for risk assessment purposes. This study aims to increase

confidence in the implementation of new approach methods in a risk assessment context by using a parallel analysis to identify data gaps in current experimental designs, reveal the limitations of common approaches deriving transcriptomic points of departure, and demonstrate the strengths in using high-throughput transcriptomics (HTTr) to derive practical endpoints. A uniform workflow was applied across six curated gene expression datasets from concentration-response studies containing 117 diverse chemicals, three cell types, and a range of exposure durations, to determine tPODs based on gene expression profiles. After benchmark concentration modeling, a range of approaches was used to determine consistent and reliable tPODs. High-throughput toxicokinetics were employed to translate *in vitro* tPODs (μM) to human-relevant administered equivalent doses (AEDs, mg/kg-bw/day). The tPODs from most chemicals had AEDs that were lower (i.e., more conservative) than apical PODs in the US EPA CompTox chemical dashboard, suggesting *in vitro* tPODs would be protective of potential effects on human health. An assessment of multiple data points for single chemicals revealed that longer exposure duration and varied cell culture systems (e.g., 3D vs. 2D) lead to a decreased tPOD value that indicated increased chemical potency. Seven chemicals were flagged as outliers when comparing the ratio of tPOD to traditional POD, thus indicating they require further assessment to better understand their hazard potential. Our findings build confidence in the use of tPODs but also reveal data gaps that must be addressed prior to their adoption to support risk assessment applications.

B.2 Introduction

Every year, new substances are introduced into the global marketplace with limited toxicity information, while presently registered chemicals that are existing on the market are continually used and re-purposed into a myriad of products available to consumers. Thus, higher-throughput techniques are required to effectively contextualize and predict the hazard potential posed by chemicals to facilitate risk assessment activities. A large number of substances, including but not limited to, those on Canada's Domestic Substances List, present a challenge for chemicals management programs, particularly when inadequate data are available. Historically, data-driven assessments used animal (principally rodent) models from traditional standardized protocols designed to address acute (immediate effects from single dose), as well as sub-chronic and chronic (longer-term effects from repeated dose) toxicological assessments (Collins et al., 2008; Barile, 2013). Regulatory agencies, including those within Canada, are taking steps to implement new approach methods (NAMs) and evolving scientific approaches to address the limitations of traditional methods (e.g., hindrance on data generation as a result of time, cost, and labour-intensive practices) as well as ethical concerns of animal use, as part of the paradigm shift to next-generation risk assessment strategies (Bhuller et al., 2021).

Transcriptomics provides a high-throughput means of producing large datasets covering a broad range of molecular responses to potentially hazardous substances. Transcriptomic technologies have been in use for over two decades in molecular biology

and recent innovations have enhanced their specificity and dynamic range to enable implementation in risk assessment activities. Application of transcriptomics for understanding toxicology (toxicogenomics) includes the analysis and interpretation of changes in gene expression caused by exposure to potentially hazardous substances to explain their prospective adverse effects (Chepelev et al., 2015; Moffat et al., 2015; Farmahin et al., 2017; Johnson et al., 2020; Krewski et al., 2020). High-throughput transcriptomics (HTTr) enables rapid evaluation of global changes in gene expression profiles in cell culture models to identify potential chemical toxicities (Harrill et al., 2021). Efficient use of HTTr is supported by the availability of computational pipelines that process large transcriptomic datasets and can be uniformly applied across numerous chemicals and exposure levels (Mezencev and Subramaniam, 2019; Verheijen et al., 2022). Benchmark dose (BMD) modeling has been applied to derive transcriptomic points of departure (tPODs) in a manner, that is, analogous to the production and application of apical PODs using traditional approaches (Thomas et al., 2013; Moffat et al., 2015; Webster et al., 2015; Farmahin et al., 2017). Given the involvement of gene expression pathways in cellular regulation, transcriptomic changes may provide opportunities for inference to a variety of contexts related to risk assessment. This includes quantifying potency, informing mode of action/mechanistic information, and characterizing adverse effects that support weight of evidence approaches to evaluate chemicals (North and Vulpe, 2010; Bourdon-Lacombe et al., 2015).

Recently, a logic framework was proposed to examine the potential to develop and apply transcriptomic methods to refine, or even replace, the current risk assessment paradigm that relies on traditional apical PODs (Johnson et al., 2022). This framework supports the current shift away from the identification of specific critical effect endpoints in animal models toward establishing conservative tPODs that are sufficiently protective to meet the contemporary needs of regulatory agencies (Stucki et al., 2022). This logic framework aligns with current efforts to facilitate efficient chemical screening using *in vitro* HTTr.

The aim of the current work was to demonstrate the capability of HTTr to derive protective tPODs by applying a uniform analysis across a diverse chemical space, building on the foundation of approaches from previous works deriving tPODs using various experimental designs (Farmahin et al., 2017; Pagé-Larivière et al., 2019; Ramaiahgari et al., 2019; Thomas, 2019; Ewald et al., 2022). This study was conducted for the purpose of building confidence in the application of these technologies by revealing the associated uncertainty and potential variability that corresponds to *in vitro* POD derivation using transcriptomic data. To achieve this, *in vitro* tPODs were compared to apical endpoints by employing *in vitro* to *in vivo* extrapolation (IVIVE) to generate an administered equivalent dose (AED). More specifically, we used high-throughput toxicokinetic (httk) modeling to translate tPODs (μM) to AEDs (mg/kg-bw/day). The AED provides a valuable endpoint to determine the human relevance of tPODs for the purpose of chemical prioritization and/or screening level risk assessment. They also provide

insight into the applicability of these, and similar NAM-based approaches, as potential endpoints for human health risk assessment. This work supports regulatory initiatives in the efforts to establish best practices and scientific confidence in the use of NAMs to produce protective human health relevant thresholds further inspiring the shift toward the reduction and replacement of animals for toxicity testing (van der Zalm et al., 2022).

B.3 Materials and methods

B.3.1 Study design

This investigation used data from multiple studies that were selected based on availability (i.e., published studies using multiple, publicly available transcriptomic datasets), specifically *in vitro* HTTr data analyzed using the TempO-Seq™ platform (BioSpyder Technologies, Inc., Carlsbad, CA), and refined using a bioinformatic pipeline. Study selection resulted in data for a total of 179 concentration-response experiments of various experimental designs, spanning 117 chemicals. An overview of these datasets is presented in Table B.1 and a detailed list of chemicals and experiments is available in the Supplementary Appendix S1. Further details on chemical preparation, working solutions, cell cultures and exposure, details on TempO-Seq library building, and an overview of QA/QC protocols with removal of designated outliers are available in the original publications (Ramaiahgari et al., 2019; Buick et al., 2021; Harrill et al., 2021; Reardon et al., 2021; Rowan-Carroll et al., 2021; Matteo et al., 2022). tPODs were derived using a

variety of approaches as described below. The tPODs were subject to IVIVE using a widely available program (httk R-package, v2.2.1) to account for the large scope of chemicals (Pearce et al., 2017) before comparison with high quality data extracted from commonly used regulatory databases (i.e., data from previous assessments using traditionally-derived PODs). An overview of the steps from collection of raw/processed data to the final comparison of AEDs and apical PODs is detailed below and depicted in a schematic workflow (Figure B.1).

B.3.2 Datasets and exposure conditions

Six datasets were identified and selected for evaluation from previously published works as well as a published OECD case study on integrated approaches to testing and assessment (IATA) conducted by Health Canada (listed in Table B.1). Within each dataset it was indicated when studies used human whole transcriptome (~20,000 probes) kits, or reduced coverage using a subset of genes with the S1500+ panel (~3,000 probes). Recent work evaluated the robustness of the S1500+ platform and demonstrated that this template is an acceptable surrogate for the whole transcriptome (Lee et al., 2021). A brief overview of the experimental designs (by cell model) has been provided below but readers are referred to the respective published studies for more detailed descriptions. The results herein were interpreted in the same manner regardless of the platform used.

Human breast cancer cells (Michigan Cancer Foundation-7; MCF-7) were depleted of estrogen for 48 h prior to exposure for 2 days to 16 bisphenols and bisphenol alternatives at ten concentrations ranging from 0.0005 to 100 μ M alongside dimethyl sulfoxide (DMSO) solvent controls. The experiment included a positive control (17 β -estradiol, range 0.0001–10 nM) and a non-estrogenic control (dexamethasone, range 0.0001–1 μ M) (Matteo et al., 2022; OECD. Series on Testing and Assessment 373, 2022). A second dataset was produced with MCF-7 cells that were treated for 6 hours (0.25 days) with 44 different substances at concentrations between 0.03 and 100 μ M (and DMSO solvent controls) alongside three chemicals (genistein, sirolimus, and trichostatin A) at a single concentration. These three chemicals were included for reference purposes but since a single concentration was used for each of them, we were not able to use them for concentration-response modeling (Harrill et al., 2021). Both experiments used human whole transcriptome TempO-Seq™ profiling.

Human HepaRG™ cells in differentiated (Hepa-D) and non-differentiated (proliferated, Hepa-P) states were exposed to 25 chemicals for 4 days (at 10 concentrations with half-log spacing) and assessed using the S1500+ platform (Ramaiahgari et al., 2019). A second set of human HepaRG™ cells were exposed to 12 potentially genotoxic (i.e., DNA damage inducing) chemicals for 2 days over a range of concentrations specific to each chemical and assessed using the S1500+ platform (Buick et al., 2021) and time-matched to solvent controls.

Primary human liver cell spheroids (3D spheroids; from 10 different human liver donors) were exposed to 4 PFAS (PFBS, PFOS, PFDS, and PFOA), including exposures of 1, 4, 10, and 14 days over a range of concentrations from 0.02 to 100 μ M using the S1500+ platform alongside time-matched DMSO solvent controls (Rowan-Carroll et al., 2021). The 3D spheroid data was expanded to 23 per- and polyfluoroalkyl substances (PFAS) but limited to 1 and 10 day exposures based on previous work demonstrating optimal exposure of these chemicals assessed using the S1500+ platform (Reardon et al., 2021).

B.3.3 Data processing and handling

All included data were obtained in the form of count tables (read counts for each probe in each sample) generated using the Templated Oligo detection assay (TempO-Seq™, from Biospyder) that were processed and subjected to QA/QC analysis as outlined in the original publications. Previous analysis of TempO-Seq™ data from a variety of publicly available datasets had found that choice of aligners and normalization methods used to process data did not significantly influence expression outcomes, demonstrating the robust nature of results produced using this target RNA-seq platform (Everett et al., 2022). Thus, our meta-analysis of multiple transcriptomic datasets used the original pre-processed data that were produced from independent bioinformatic pipelines. In several of the studies (Buick et al., 2021; Matteo et al., 2022), a general bioinformatics pipeline

designed to process transcriptomic data for regulatory applications was used for data pre-processing, known as the Omics Data Analysis Framework for Regulatory Application (<https://github.com/R-ODAF/>) (Verheijen et al., 2022). Other studies used the US EPA pilot pipeline (https://github.com/USEPA/htrpl_pilot) (Ramaiahgari et al., 2019; Harrill et al., 2021), or other custom analysis pipelines (i.e., TempO-SeqR, v3.0, provided by BioSpyder for aligning TempO-Seq data, alongside custom data pre-processing steps in R) (Reardon et al., 2021; Rowan-Carroll et al., 2021; OECD. Series on Testing and Assessment 373, 2022).

All the datasets in this study used a similar strategy to process data, starting with FASTQ files generated from the sequencing results of a TempO-Seq™ workflow, and ending with a tabular matrix. In this tabular matrix, the genes are represented in rows and samples are represented in columns, and the values in the matrix are the output (i.e., counts) from the alignment step. The steps in handling the data from these experiments include: 1) data quality assessment, 2) study-wide alignment quality controls; and 3) downstream applications to derive a tPOD. The first step, pre-processing, deals primarily with assessing the quality of the high-throughput sequencing data used in the study and creating the initial count matrix. Study-wide QC strategies were based largely on those described in Harrill et al., 2021, which aim to eliminate low-quality samples based on alignment to reference sequences and the distribution of gene counts in individual samples compared to all treatment-related samples in the dataset (i.e., calculating quality metrics such as alignment rates, panel coverage in terms of the number of active probes,

etc., from the count matrix). Finally, downstream applications include calculating gene-level, pathway-level, or signature-level BMCs and subsequently deriving a tPOD, the results of which are filtered based on several statistical rules (e.g., R-ODAF criteria) at the per-gene or per-probe level. In this meta-analysis, we reanalyzed the datasets under consideration to compare methods for tPOD derivation.

B.3.4 Derivation of *in vitro* points of departure using transcriptomic concentration-response data

tPODs were derived from *in vitro* concentration response data using BMC modeling prior to conversion to AEDs. The workflow below describes the steps in the methodology developed to derive AEDs from original processed data. Prior to importing into BMDEExpress software (v2.3) for analysis, the count matrix was imported to a DESeq2 object, and the size factor was normalized to log2 counts per million (CPM) as a separate function for each sample concentration (Yang et al., 2007; Phillips et al., 2019). A single project file (.bm2) was created using a consolidated listing of chemicals that was inclusive of all experimental conditions for analysis within BMDEExpress. It is important to note that concentration data (generally with a minimum of three concentrations and solvent controls) are required to identify the BMC of responsive genes within the data. Here, a minimum of five concentrations were included along with a solvent control. The BMC modeling was performed separately for each experimental model for each test chemical

and only used those samples (i.e., replicate concentrations) that passed all QC filters as defined by the criteria within their respective studies. Williams Trend tests (Williams, 1971) were applied to filter out probes that did not show a concentration-response (i.e., probes passing filters had p-values <0.05). Additional filtering was also applied to remove probes that did not achieve a fold-change (FC) of at least 1.5 in at least one concentration. To calculate BMC values for probes, the best-fit curve was selected from a series of possible models including Power, Linear, Polynomial 2, and Exponential 2, 3, 4, and 5 models. Best-fit models for probes were selected based on a nested chi-square test cut-off of 0.05 to select among the linear and polynomial models that was followed by the lowest Akaike Information Criterion that estimates the quality of each model relative to the other models. A full description of the modeling parameters is available in the BMDEExpress2 published online documentation (US EPA BMDEExpress2, n.d.; Phillips et al., 2019). Additional parameters applied for modeling included: restrict power equal or greater than one; maximum iterations of 250; confidence interval of 0.95, and benchmark response factor of one standard deviation (i.e., BMR of 1 SD). Post-filtering criteria included removing BMCs with a goodness-of-fit test p-value less than 0.1, a ratio of BMC upper bound (BMCU) divided by the BMC lower bound (BMCL) greater than 40; and removing BMCs that were greater than the highest exposure concentration. Probe IDs representing select genes passing all filtering criteria were converted to their corresponding Entrez Identifiers and were carried forward to be used for tPOD derivation.

We evaluated seven approaches to calculate tPODs falling under two separate tracks, using concentration-response modeling performed with log2 CPM data using BMDEExpress. The first track used the distribution of genes (e.g., percentiles, numbered rank, and mode), while the second used the median gene BMC value of the lowest (i.e., most sensitive) gene set from a selection of available, open-source, curated, and peer-reviewed pathway databases. All of the included methods from both tracks are described in more detail below and density plots for each chemical meeting the minimum criteria to derive selected tPODs are provided in Supplementary Appendix S2.

B.3.4.1 The fifth percentile

The fifth percentile gene was calculated using the BMC of the gene closest to the fifth percentile value of the BMC distribution. It represents the lower bound fifth percentile of all filtered gene BMCs commonly employed in previous work using the US EPA Toxicity Forecasting database (Paul Friedman et al., 2019) to derive *in vitro* PODs and has been employed within previous studies (Reardon et al., 2021; Rowan-Carroll et al., 2021) to derive tPODs [i.e., represented mathematically as $\text{floor}(0.05 \times \# \text{ BMCs})$]. The fifth percentile is a conservative metric value for the tPOD that targets the most responsive genes that may contribute to the toxicological response.

B.3.4.2 The first mode

Several previous studies used genes from the first mode of the BMD frequency distributions (Qutob et al., 2018; Farmahin et al., 2019; Pagé-Larivière et al., 2019; Alcaraz et al., 2021). In the current study, density estimation was used with forward, backward, and centered differencing to estimate the first and second derivatives. The first mode was defined as the first point at which the first derivative changes from positive to negative with a negative second derivative (second derivative test).

B.3.4.3 The 25th lowest ranked gene BMC

All genes with BMCs were ranked from lowest to highest. The 25th rank-ordered gene was set as the threshold and the BMC of this gene was used to indicate the concentration where a defined change in the transcriptome has occurred (Reardon et al., 2021; Matteo et al., 2022). Those chemicals that were unable to produce a representative value for this tPOD (i.e., did not have at least 25 responsive genes with BMCs) were identified as “inactive” based on this approach and alternatives methods were considered for tPOD derivation.

B.3.4.4 The LCRD

The lowest consistent response dose (LCRD) was performed as described in Crizer et al. (2021). The LCRD is identified as the lowest BMC in a rank order of gene

BMCs where all subsequent ratio values from adjacently ranked BMCs are within 1.66 (thereby eliminating unrealistically low gene BMCs that may be biological noise). BMCs in the ranked group are declared the consistent response group of BMCs (CRGB) as defined by Crizer et al. (2021), because all sequential BMCs in this group have at least one BMC, that is, within $1/4$ log difference in value. The lowest BMC in the CRGB is then identified as the LCRD. It has been recognized that this method may be sensitive to a few exceptionally low BMCs (i.e., deriving an estimate that is overly conservative), particularly for an extrapolated BMC below the lowest dose. Here, a modification of the LCRD is presented where the BMCs were grouped as stated above except the LCRD was the lowest dose from the largest CRGB. This eliminates the issue of small groups ($n = 2$ or 3) of genes with low BMCs being defined as the LCRD.

B.3.4.5 The lowest gene set

This approach to tPOD derivation was defined by the United States National Toxicology Program (now the Division of Translational Toxicology: DTT) (National Toxicology Program, 2018). In the current study, to capture the most sensitive gene set, genes and their associated BMC values were matched to their corresponding gene sets using three well-known public databases: 1) Gene Ontology (GO) Biological Process (Harris et al., 2008); 2) Reactome Pathways (Fabregat et al., 2017); and 3) Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al., 2017). Gene sets that

contained at least three genes with a BMC representing at least 5% of the gene set (based on total annotated genes) were selected. The lowest gene set approach was established to extract tPODs using genomic dose-response modeling, that is, linked to meaningful biological change while reducing the influence of background (i.e., noise) as the gene sets define biological functions/processes (National Toxicology Program, 2018). However, it should be noted that the function of the selected gene set is not considered when interpreting the data.

B.3.5 Conversion of transcriptomic points of departure to administered equivalent doses

IVIVE was used to convert derived tPODs (μM) to AEDs (mg/kg-bw/day), a more aligned metric for comparison with apical PODs from *in vivo* data. Each derived tPOD was converted to an AED using reverse dosimetry and available data within htkk. The AED is the theoretical dose required to reach a given steady-state plasma concentration (C_{ss} ; Wetmore et al., 2015). In brief, the htkk package version 2.2.1 (Pearce et al., 2017) in R (R Core Team: R Foundation for Statistical Computing, 2020) was used for IVIVE. To predict the C_{ss} in the current study, the three-compartment steady-state toxicokinetic model (“3compartmentss”) (modified from Wetmore et al., 2012; 2015) was used. The parameters required for this model are intrinsic hepatic clearance and plasma protein binding. Full absorption by gut is assumed when data on the fraction of compound

absorbed by the gut was not available (Wetmore et al., 2012). The htk package also provides tools to perform Monte Carlo sampling that represents the inter-individual variability within the population (Ring et al., 2017). A similar approach to Paul Friedman et al. (2019) was used within the current work, wherein the C_{ss} was calculated using the “calc_mc_oral_equivalent” function in htk with the default parameters and output using the value depicted by the 95th quantile.

Then, the resultant C_{ss} was used to calculate the AED (mg/kg-bw/day) using Eq. 1:

$$AED \left(\frac{mg}{kg \cdot day} \right) = bioactivity \text{ concentration } (\mu M) \times \frac{1 \frac{mg}{kg}}{C_{ss} (\mu M)} \frac{day}{day}$$

B.3.6 Comparison of administered equivalent dose to corresponding *in vivo* data

For the purpose of the current study, *in vivo* data was collected from available databases for comparison with *in vitro* derived endpoints (e.g., AEDs). Data was extracted from the EPA CompTox (CompTox Chemicals Dashboard (epa.gov), as well as databases from the publicly available ECHA REACH dossier (Understanding REACH - ECHA (europa.eu), the OECD QSAR Toolbox (The OECD QSAR Toolbox - OECD), and dossiers available through Health Canada (Chemical substances - Canada.ca). The

lowest value available from oral repeat dose (sub-chronic and chronic), developmental toxicology and reproductive toxicology studies was selected as the apical POD (Table B.2). An expanded list was considered in cases of chemicals with an available lowest apical POD from multiple study types (Supplementary Appendix S3). The lowest apical POD from *in vivo* data was selected from available endpoints, including the no-observed-adverse-effect-level (NOAEL); lowest-observed-adverse-affect-level (LOAEL); no-observed-effect-level (NOEL); no-effect-level (NEL); lowest-observed-effect-level (LOEL); or the lowest-effect-level (LEL).

Under circumstances where a designated NOAEL or LOAEL was not available for select chemicals, an alternative value was selected to represent the *in vivo* endpoint. For aflatoxin B1 (AFB1), the lowest outcome available was the lower-bound BMD (BMDL). For cyclophosphamide, only endpoints from studies of carcinogenicity were available, resulting in a designated “cancer unit” used to represent the apical POD. For eugenol, the highest no effect level (HNEL) was established as the lowest apical POD for comparison. A concerted effort was made to capture the largest number of chemicals by collecting data using commonly available regulatory databases.

B.4 Results and discussion

B.4.1 A transcriptomic workflow for a diverse chemical space

Transcriptomic technologies provide information on gene expression and the initiation of molecular changes that occur prior to the development of apical effects. Mounting evidence supports that these alterations can be used to establish molecular-based PODs that are human health-protective in the absence of predicting a specific hazard (Johnson et al., 2022). Such a transcriptomic effect level would ideally identify exposure concentrations equal to or below those causing critical effects associated with adverse outcome pathways (Ankley et al., 2010). Previous studies using short-term animal exposure data have demonstrated that the use of both *in vitro* and *in vivo* derived tPODs are comparable or even more sensitive than apical PODs derived using regulatory guideline studies (Bhat et al., 2013; Thomas et al., 2013; Johnson et al., 2020; LaRocca et al., 2020). To accelerate the adoption of NAMs, frameworks to increase confidence for application in the regulatory decision-making process have been proposed that include the key elements of “fitness for purpose, human biological relevance, technical characterization, data integrity and transparency, and independent review” (van der Zalm et al., 2022). The meta-analysis described herein applies these concepts to increase scientific confidence in the use of tPODs, and their corresponding AEDs derived from *in vitro* data to further demonstrate support that this approach results in equal or greater protection of human health.

The workflow includes datasets that have previously undergone data processing which allows for the derivation of tPODs from a range of approaches (Figure B.1). The recently published R-ODAF pipeline has been used for previous datasets generated at

Health Canada. Although this pipeline is proposed as a “baseline reference” for data processing prior to BMC modeling (Verheijen et al., 2022), it is acknowledged that best practices have yet to be developed for analyses in the field of toxicogenomics, particularly for application in risk assessment frameworks. It has been argued that flexibility is required because methods and approaches should ideally be fit-for-purpose (Buesen et al., 2017). Practically, it is unlikely that a universal tPOD would be considered sufficient across all potential outcomes, or as a single best practice, and a range of approaches across diverse *in vitro* datasets should be considered. Integration of new methods into regulatory decision-making requires best practices and interpretation procedures to establish alignment and consistency of use for human health risk assessment. To work toward this goal, a subset of tPODs was identified based on data-distribution (e.g., percentiles, modes, and ranked genes), data-driven (e.g., LCRD), and/or pathway-related (e.g., gene set) approaches. Furthermore, applying a “baseline reference” approach such as the R-ODAF to as many datasets as possible alleviates some of the difficulty involved in comparing studies undertaken at various times and/or under different conditions. The current work used HTTr datasets with a defined workflow to investigate a variety of tPOD derivation methods that identify a point of concerted molecular change from exposure to a diverse set of environmental chemicals.

B.4.2 A general comparison of approaches

When examining all of the included approaches to tPOD derivation, bisphenols and bisphenol alternatives were the most potent chemicals but had the tPODs with the highest variability across all datasets (Figure B.2); a comprehensive list of all tPODs for all chemicals within their defined datasets is provided in Supplementary Appendix S4. The overall median $\log_{10}$ value for all tPODs within the bisphenols group of exposed MCF-7 cells was $-0.33 \mu\text{M}$ (Figure B.2A), compared to values of 0.90, 0.91, 0.99, and $1.74 \mu\text{M}$ from investigations of PFAS (Figure B.2B), or data from Ramaiahgari et al., (Figure B.2D), Harrill et al., (Figure B.2E), and Buick et al., (Figure B.2C), respectively. The degree of concordance between tPODs was evaluated by calculating the difference between the lowest (minimum) and highest (maximum) tPOD for each chemical (i.e., the higher the value of difference the lower the agreement between tPODs). Comparing between datasets, the median $\log_{10}$ difference from least to most agreement was $1.87 > 1.67 > 1.26 > 1.07 > 0.84 \mu\text{M}$ for the bisphenols case study, Ramaiahgari et al., HC PFAS studies, Harrill et al., and Buick et al., respectively (Figure B.2). Thus, the bisphenols dataset had the largest median log difference across tPODs spanning multiple orders of magnitude (Figure B.2A). Concordance between different tPODs on the same chemical within the Ramaiahgari et al. dataset was also quite low, but unlike bisphenols this was attributed to only a few outlier chemicals with individual tPODs derived using the fifth percentile that were orders of magnitude lower than tPODs derived using other approaches (Figure B.2D). In contrast, other datasets such as the PFAS studies and those chemicals evaluated from Buick et al., yielded concordant and consistent tPODs

regardless of the metric used (Figures B.2B, C). This small range between tPODs of PFAS and those from Buick et al. is likely a consequence of the study design. Specifically, the use of a limited range of exposure concentrations that results in tighter groupings of derived tPODs. Overall, the data revealed the fifth percentile to be lower, in some cases by more than an order of magnitude, than the median of all tPODs for select chemicals, including Bis4CPS (Figure B.2A); PFPeA (Figure B.2B); aspirin, DMN, CPZ, sucrose, cyclophosphamide, PB, levofloxacin (Figure B.2D), and fulvestrant (Figure B.2E). Except for select bisphenol alternatives, the general agreement for the majority of observed tPODs across datasets of diverse chemicals reinforces the robustness and pursuit for the practical application of transcriptomic data for chemical potency ranking, grouping and risk assessment.

Although numerous approaches may be considered to derive tPODs for chemicals across a diverse space, specific approaches may over- or underestimate chemical potency. It was observed that tPODs derived using the fifth percentile produced high potency rankings for sucrose and other chemicals used as low hazard reference chemicals that have been rarely associated with liver toxicity (caffeine, levofloxacin, and aspirin) (Ramaiahgari et al., 2019). Furthermore, we note that percentile gene BMCs and the extent of transcriptional change (e.g., the total number of responsive genes) may be influenced by the top dose/concentration. For example, the fifth percentile tPOD for cyclophosphamide was 0.013 μ M (derived from 14 genes fitting BMC models) in the Ramaiahgari et al., dataset using HepaRG cells that were exposed up to a top

concentration of 300 μM . In contrast, the tPOD for this same chemical was 804.3 μM from the Buick et al. dataset using the same cell type where exposure ranges were up to 10,000 μM , thereby producing a more bioactive response (394 genes fitting BMCs). Although previously found to be effective when deriving a tPOD from a wealth of assays using ToxCast (Paul Friedman et al., 2019), deriving a tPOD at the single gene level (such as the fifth percentile) with transcriptomic data may produce inaccurate results (i.e., significant over- or underestimations), and should be interpreted with caution.

Alternatives to the fifth percentile include approaches with requirements and filters to reduce the potential for mischaracterizing chemical potency that includes, but is not limited to, the BMC distribution requiring at least 25 genes, or a gene set with a minimum of three genes and/or 5% of a pathway. The 25th ranked gene tPOD requires a minimum amount of biological activity as described by 25 concentration-responsive genes for tPOD derivation, decreasing the likelihood of chemicals with low bioactivity being identified as having the potential for toxicity. Although interpreted as an arbitrary value, the 25th ranked gene BMC in previous work provided consistent potency rankings of PFAS along with representing a toxicological response in approximately 0.1% of genes in the genome (Reardon et al., 2021). The LCRD tPOD “identifies the most sensitive non-outlier feature, that is, the plausibly representative lowest dose level where a consistent response in biological features is observed”, to identify a point of toxicological relevance (Crizer et al., 2021). Here, although there are a few exceptions, as an alternative approach to the fifth percentile the LCRD was generally the lowest and most conservative tPOD based on the

distribution of BMCs across all chemicals, making it a promising candidate for deriving protective tPODs. In contrast to this, the first mode was predominantly the least conservative estimate in our study, consisting of the lowest proportion of tPODs. This approach requires the presence of a mode and, thus, a sufficient extent of biological activity. Previous work examining responses to 1060 chemicals across a battery of 815 *in vitro* assay endpoints suggested that the first mode generally corresponded to a disruption of specific biomolecular targets or pathways (e.g., receptors or enzymes) and generalized disruption of cellular machinery (Judson et al., 2016). However, observed effects at higher concentrations (i.e., higher modes) are often characterized by a larger number of affected pathways compared to the first mode, and represent dysregulation of cellular machinery that leads to cell stress and cytotoxicity. Thus, selecting the first mode as a tPOD provides an estimate at which initial molecular events are triggered.

As highlighted, there are multiple options when considering a means of obtaining a tPOD and the general agreement observed among chemicals for most datasets underscores the robust nature of gene expression data. For ease of comparison, the results, advantages, and disadvantages of all of the aforementioned approaches have been tabulated (Table B.3). In summary, deriving tPODs using percentiles (e.g., the fifth) could lead to mischaracterization of a chemicals potency when interpreting the results; whereas other approaches such as the 25th ranked gene, the LCRD, the first mode, and the lowest gene set (further described below) provide viable alternatives to derive relevant and protective tPODs to facilitate risk assessment activities.

B.4.3 Conversion to human relevant exposure values

The integration of IVIVE within the workflow allowed for determination of a dose that converts the tPOD (μM) to an estimate of the surrogate bioactivity POD (i.e., the AED in mg/kg-bw/day). Chemicals with AEDs (mg/kg-bw/day) were ranked by relative potency (Figure B.3). The list was reduced from the 117 chemicals with tPODs (Figure B.2) to 54 chemicals that could be modeled in the htk R-package (Figure B.3). The bisphenols and alternatives group (specifically BPA and BPAF) that were previously ranked as some of the highest and most potent chemicals based on derived tPODs were no longer among the most highly potent and instead were within the middle of the relative chemical potency ranking. Furthermore, the PFAS group that was not particularly potent based on tPODs (Figure B.2) included 7 of the top 12 most potent chemicals when ranked by their AEDs (Figure B.3). Other select chemicals (2DD-glucose, TGSA, Bis4CPS, and urea) had apical PODs from animal data but AEDs from BMC distributions or gene set approaches could not be produced. These chemicals were subsequently excluded from further analysis.

The conversion of tPODs to AEDs had two primary impacts on the results and interpretation. First, it reduced the number of chemicals that could be analyzed because of a lack of data availability for IVIVE. Second, AED derivation altered the relative potency ranking of chemicals. In silico models are currently being employed to predict the

parameters needed to broaden the application of NAM-based PODs toward “characterisation, validation and reporting of Physiologically Based Kinetic (PBK) models for regulatory purposes” under the OECD (OECD, 2021). These models allow for the determination of a C_{ss} and are a critical step in creating relevant exposure estimates from *in vitro* data that could be applied in the context of human health risk assessment. Here, the AED was employed as a quantitative estimate to serve as a protective human-relevant effect level that was anticipated to be lower than potential apical adverse outcomes. This process has been previously employed to determine protective estimates for potentially hazardous chemicals using data derived from the US EPA toxicological forecast (ToxCast) database (Paul Friedman et al., 2019; Health Canada, 2021). The primary advantage of combining IVIVE with human cell models or tissues is that a correction factor for sources of uncertainty related to interspecies differences from animal-derived data may not be necessary for developing regulatory values (Bos et al., 2020). Furthermore, although it is beyond the intent of the current work, the httk platform also has the potential to accommodate input of parameters from *in silico* predictions such as for the fraction of chemical not bound to protein and intrinsic hepatic clearance, to model the C_{ss} for chemicals not listed within the httk chemical library.

The implementation of IVIVE using *in silico* models and corresponding model assumptions also have potential limitations. In our study, a lack of available data on the parameters necessary for IVIVE reduced the overall sample size (i.e., number of chemicals included in our study); this is particularly problematic when evaluating “data-

poor” chemicals that are also outside the applicability domain of models used to provide predicted input parameters. The AEDs produced using the PFAS data were influenced by IVIVE resulting in their increased potency ranking. The generalized parameters and assumptions of IVIVE models may not have adequately captured their toxicokinetic properties and disposition, resulting in overestimation of their chemical potency (e.g., prolonged half-life) (Fenton et al., 2021). Thus, although there is some inherent uncertainty associated with htk models due to the (conservative) assumptions necessary to allow for minimal data input and high-throughput data processing, it has been demonstrated to be a useful approach to provide protective AEDs from *in vitro* data. Here, htk was used to convert tPODs to AEDs to determine practical, human-relevant effect levels that would allow risk assessors to make informed decisions, even when traditional animal data was not available. Ongoing efforts at Health Canada, as part of international collaborations, are working to characterize the impact of different IVIVE models for reverse dosimetry, *in vitro* disposition models, as well as parameterization of models using various data streams, on the refinement to the AED estimate. Such work will increase confidence in the use of IVIVE approaches and provide guidance on when higher tier models should be considered based on chemical space and context of use.

The estimate from the lowest and most conservative approach for each individual chemical was identified for both the BMC distribution (blue markers, Figure B.3) and gene set level (magenta markers, Figure B.3) approaches were carried forward to evaluate the extent of correlation between different tracks (Figure B.4). Overall, there was a degree of

agreement (within one order of magnitude of perfect agreement) for the majority of chemicals (~90%), even considering differences in experimental design including exposure time, model type and approach used to derive tPOD [Pearson correlation, $r = 0.98$ ($p < 0.0001$)]. Of these chemicals, select data points (from BPA, PFOA, fenpyroximate, cyanazine, 4-hydroxytamoxifen, and amiodarone hydrochloride) were more than an order of magnitude outside the range of agreement (labeled in Figure B.4), and fulvestrant was significantly (greater than two orders of magnitude) outside the range of agreement. For BMC distribution level values, over the chemical space, the lowest AED for the majority of substances was the LCRD (45 of 54), followed by the fifth percentile (6 of 54), and the remaining three represented by the first mode and 25th ranked gene. Thus, the AEDs of over 90% of concentration-response experiments were based on the LCRD and only a small number of chemicals were represented by the fifth percentile (atrazine, BPA, caffeine, fulvestrant, PFDA, and vinclozolin). As previously discussed, the fifth percentile is subject to bias from the concentration exposure range; this may explain fulvestrant being significantly outside the range of agreement.

Overall, the results demonstrate a high degree of concordance between BMC distribution and gene set level AEDs produced using a diverse set of approaches. The application of htk generated human relevant AEDs from *in vitro* derived tPODs to produce practical estimates for potentially hazardous and data-poor chemicals.

B.4.4 Comparing *in vitro* to *in vivo* derived data

The AED for each chemical was selected from the lowest value of either the BMC distribution (Figure B.5A) or lowest gene set (Figure B.5B) for comparison with apical PODs derived using *in vivo* data from regulatory databases (Table B.2). The AEDs (black) were generally lower than apical PODs from developmental (red), reproductive (green) and repeat dose (blue) studies (Figure B.5). Using the BMC distribution, the lowest AEDs were derived using the LCRD, fifth percentile, first mode, or 25th ranked gene approaches for 45, 6, 2, and 1 of 54 substances, respectively, wherein for most (47 chemicals), there was an observed pattern of lower tPOD-based AEDs than apical PODs (Figure B.5A). Another commonly employed method of tPOD derivation uses the median value of the lowest/most-sensitive gene set as a measure of potency (National Toxicology Program, 2018). AEDs derived from the lowest gene set found that most of the observed AEDs (45 of 50) followed the same pattern and were also lower than apical PODs (Figure B.5B). Most AEDs from the lowest/most sensitive gene set were derived from tPODs using the gene ontology (GO) database (33 of 50), followed by REACTOME (14 of 50), and KEGG (3 of 50) (Figure B.5B). A general pattern of lower/more conservative AEDs was observed, suggesting that for the majority of chemicals, NAM-based AEDs using either distribution or gene set based approaches are at least equal to, or more conservative than conventional apical endpoint PODs selected for risk assessment purposes.

The LCRD was the most consistent approach from the BMC distribution and GO was the consistent data source for the lowest median gene set level approach amongst the multiple methods of tPOD derivation examined. As previously discussed, the LCRD is a recently developed data-driven approach to derive PODs using transcriptomic data. Our results indicate this approach to be well-suited to derive conservative tPODs across a diverse chemical space. Previous work using functional enrichment and predictive modeling to identify lowest median genes sets showed that results differed when assessing equivalent pathways from different databases; thus, database choice is a significant factor when using pathway-centric approaches (Mubeen et al., 2019). The size of the gene set database may also be a significant factor, as the number of pathways present will influence results (Mubeen et al., 2022). Our finding that the GO database produced the highest proportion of gene set AEDs was likely attributed to its size, with a larger number of available human annotations compared to REACTOME and KEGG. An optimal database for gene sets has yet to be established. Thus, at present we recommend including multiple databases to ensure that a broad biological space is queried and that a conservative AED is derived.

AEDs may be used to provide a protective dose below which toxicity is not expected to occur in support of chemical screening. Previously, AEDs derived using the fifth percentile of values from a large set of *in vitro* assays using available ToxCast data were compared with apical PODs from *in vivo* data to derive a practical risk-based metric for prioritization and assessment activities; namely, the bioactivity exposure ratio (BER)

(Paul Friedman et al., 2019; Health Canada, 2021). During a formal risk assessment, the most appropriate critical effect level from animal studies is typically selected after a thorough review that includes a study quality evaluation, wherein the lowest value recorded as the apical POD may not always be the most fitting estimate for comparison with estimates of exposure. The rationale for exclusion may be based on limitations in the study design, the quality of the data or reporting, or may not consider a particular sub-population that represents the exposure scenario of interest. For example, the NOAEL from a developmental or reproductive toxicity study may be most suitable when evaluating the risk for women during pregnancy and their developing children but in some cases may also be generalized to include women of childbearing age or men. Here, a comparative assessment indicated that transcriptomic based AEDs were equally or more protective than the majority of apical PODs, regardless of the study type considered (e.g., repeated dose, reproductive, or developmental toxicity). Such AEDs may be applied in the same manner to create human relevant BERs, reinforcing the effectiveness of NAM-based approaches to create protective, human relevant PODs for consideration in risk assessment.

Additional chemicals had a sufficient number of modeled genes to derive BMCs from distribution level approaches (53, 62, and 100 total genes with BMCs for cyproconazole, atrazine, and simazine, respectively, Figure B.5A) but not gene set level approaches (Figure B.5B). There is likely an insufficient number of these genes with BMCs that could be mapped to a gene set in each of the KEGG, GO, or REACTOME

databases after application of parameters and filters. Thus, although our study had a somewhat limited sample size because of data availability, the results included both focused datasets with numerous chemicals from a single class (e.g., PFAS and bisphenol alternatives) as well as datasets with substances from broader classes (e.g., Ramaiahgari et al., and Harrill et al.). These findings support observations that transcriptomic AEDs derived from a variety of approaches provide a conservative endpoint for evaluating chemical potencies and emphasize the importance of including multiple databases when mapping genes for AED derivations that use gene sets.

B.4.5 Identifying outliers using ratios

A $\text{Log}_{10}\text{Ratio}$ was used to compare AEDs from *in vitro* data with apical PODs from traditional data (adapted from Paul Friedman et al., 2019). The ratio was produced using the log_{10} (mg/kg-bw/day) units and calculated by the difference between NAM-based AEDs and traditional PODs with the following equation (Eq. 2).

$$\text{Log}_{10}\text{Ratio} = \text{Log}_{10}\text{POD}_{\text{Traditional}} - \text{Log}_{10}\text{AED}_{\text{NAM}}$$

The majority of values for the $\text{Log}_{10}\text{Ratio}$ were > 0 , with NAM-based AEDs being more conservative than traditional PODs (Figure B.6). Using BMC distribution approaches, 28 of 35 chemicals (corresponding to 60 of 72 included data points, or 83% of the dataset) had a median $\text{Log}_{10}\text{Ratio} > 0$ indicating that these AEDs were lower than apical PODs (Figure B.6A). Using gene set level approaches, 27 of 32 chemicals

(corresponding to 56 of 66 included data points, or 85% of the dataset) had AEDs lower than apical PODs (Figure B.6B).

In contrast, those chemicals that were flagged as outliers (with a $\text{Log}_{10}\text{Ratio} < 0$) included three chemicals identified as triazine herbicides (simazine, atrazine, and cyanazine), as well as B[a]P, eugenol, cyclophosphamide, and AFB1 (Figure B.6). A detailed description for each of these chemicals, including the *in vitro* AED, the comparative apical POD, and a rationale to define those chemicals as outliers is detailed in the supplementary information (Annex A). In brief, atrazine, cyanazine, and simazine were included in previous work as triazine herbicides for their capacity to inhibit photosystem-II that were not intended to be active within the cell model of interest (MCF-7) and, as a result, were outside of the applicability domain for the target cell culture system resulting in non-conservative tPODs (Harrill et al., 2021). B[a]P has numerous routes of exposure for humans that results in developmental, reproductive, and immunological toxicity *in vivo* (US EPA, 2017) that may not be captured within the current *in vitro* model (HepaRG). The higher *in vitro* POD for eugenol than apical PODs may also be attributed to the limited applicability domain of the *in vitro* models used to derive the tPOD and subsequent AED for this chemical. Cyclophosphamide requires metabolic activation, primarily through the liver, in order to induce varying degrees of toxicity (Ayash et al., 1992; Moghe et al., 2015; Groehler et al., 2016). AFB1 also requires bioactivation in order to enact hepatocarcinogenesis resulting in AED estimates that were higher (i.e., less conservative) than epidemiological PODs from incidences of human liver cancer

(EFSA, 2007). Although the HepaRG cell models used within the *in vitro* studies are considered to be more metabolically active than more simple *in vitro* human hepatocyte models (Lübberstedt et al., 2011), they likely do not maintain a high enough metabolic capacity in conventional 2D static culture models to induce the level of toxicity observed with cyclophosphamide or AFB1 *in vivo*. This limitation could be resolved by using a more sophisticated model such as a liver spheroid or microtissue, or via the addition of a metabolic component (e.g., an S9 fraction) to provide a more accurate *in vitro* tPOD and corresponding AED.

The Log₁₀Ratio provides an efficient means for evaluating the effectiveness of *in vitro* PODs to derive conservative estimates for numerous substances across a broad chemical space. Overall, we found that the lowest and most conservative value from the range of derived AEDs was lower (i.e., protective) than available *in vivo* derived apical PODs for the same chemicals with a few exceptions (e.g., outlier chemicals). Within this context, the workflow could be applied in a tiered framework to highlight or flag potentially hazardous or problematic chemicals for further research, data generation or risk assessment.

B.4.6 Examining design considerations and uncertainty when applying transcriptomic data for risk assessment

B.4.6.1 Study design

Multiple HTTr data sets could be used to examine the effect of study design parameters on tPODs and AEDs. Specifically, increasing the duration of exposure and/or model complexity generally corresponded to increased chemical potency (Figure B.7). PFAS represented the majority of chemicals with varied data points (e.g., multiple time points and cell models), including carboxylates (PFHxA, PFHpA, PFOA, PFNA, PFDA, and PFUnA), sulfonates (PFBS, PFHxS, and PFOS), and a longer-chain perfluorosulfonamide (PFOSA). With a few exceptions, the PFAS showed agreement (within one order of magnitude) between multiple AEDs produced from experiments using these different designs. Outside of PFAS, AFB1, B[a]P, BPA, cyclophosphamide, and troglitazone from independent experiments produced multiple AEDs that, with the exception of B[a]P, were different by orders of magnitude. Considering all of these chemicals, there was a trend of increasing potency (i.e., a decreased AED) with increased exposure duration. For specific chemicals with available data, AEDs derived using more complex 3D liver spheroid models (e.g., PFOS) or HepaRG cells (e.g., troglitazone) were lower than AED from MCF-7 cell models. These outcomes suggest a need to consider uncertainty relating to the study design used, such as the exposure duration and model when applying these values in risk assessment applications.

B.4.6.2 Uncertainty relating to apical endpoints

Prior to discussion of the uncertainties related to *in vitro* tPOD and AED derivation, it is important to consider the main sources of uncertainty related to traditional approaches and conventional animal testing. Currently, PODs for health effects are derived using modeled apical measures such as NOAELs, LOAELs, and BMDs from relevant studies. Uncertainty factors are commonly applied to PODs to compensate for limitations, knowledge deficiencies, and uncertainties in the data. Such limitations have been previously discussed (Kimmel and Gaylor, 1988; Leisenring and Ryan, 1992; Haber et al., 2018). Sources of uncertainty include dose selection and dose spacing that are identified when defining a NOAEL or LOAEL using a limited range of doses within the experimental design. Moreover, the dose–response is not accounted for in NOAEL or LOAEL derivation as these estimates are derived based on the effect observed at a single dose. Among different studies the observed experimental response in animal models may vary, making inter-study comparisons challenging (Kimmel and Gaylor, 1988; Leisenring and Ryan, 1992; Haber et al., 2018). Recent work has attempted to quantify inherent biological and protocol variability of inter-laboratory results using curated reference data for acute oral rat LD50 that resulted in values that varied by approximately $\pm 0.24 \log_{10}$ units (mg/kg) (Karmaus et al., 2022). Pham et al., performed a statical evaluation and estimated a variance of 0.5–0.6 $\log_{10}$ units (mg/kg/day) for LEL and/or LOAEL values using critical effect level outcomes (e.g., target organ, clinical chemistry, or in-life observation) from *in vivo* studies within the publicly available EPA Toxicity Reference Database (Pham et al., 2019; 2021). Herein, the use of apical PODs provided a reference

point to compare and contrast the NAM-based effect levels; specifically, the results suggest that *in vitro* derived tPODs and AEDs are generally more conservative than apical PODs. Given the range of AEDs observed based on multiple studies within the current work, it is also important to recognize that similar ranges, and uncertainties exist for apical PODs and strategies are used to account for these within decision-making frameworks. The identification and quantitation of uncertainty are currently being investigated as part of on-going efforts to implement *in vitro* derived data into current and future risk assessment strategies.

B.4.6.3 Uncertainty relating to *in vitro* models to derive AEDs

Aspects of the design, including the *in vitro* experimental model used, as well as the selection of the dose range and spacing of the exposure are subject to uncertainty. Thus, aspects of the design may influence the magnitude of AEDs from derived tPODs that reflect the observed chemical potency. There are numerous factors that should be considered in order to provide a practical fit-for-purpose molecular-based dose estimate. One primary factor is the selection of an appropriate and representative *in vitro* model. For example, immortalized cancer derived cell lines are frequently used in toxicology experiments because they are widely available, easily cultured and facilitate reproducibility (i.e., the same cells can be used for all experiments). However, these cell lines may not reflect the response of non-cancer tissues. Primary cells directly derived

from humans retain tissue-specific characteristics but have low proliferative potential and are typically only used in a limited number of experiments (Liu et al., 2020). Immortalized cancer cell lines typically lack the capacity to metabolically activate chemicals to produce an accurate AED. The use of model mixtures including spheroids and microtissues consisting of primary cells derived from multiple donors may provide an effective means to capture more human-relevant responses. These complex, multi-donor models reduce any donor-specific biases in the data output. Furthermore, immortalized cell lines are generally maintained *in vitro* in a monolayer or in suspension, whereas microtissues using 3D models containing multiple cell types better reflect *in vivo* characteristics (Proctor et al., 2017; Reardon et al., 2021; Rowan-Carroll et al., 2021). Although complex spheroid models may be more suitable, they have limitations related to donor heterogeneity (i.e., sampled from a limited number of individuals), cost-efficiency, and availability (Fraczek et al., 2013; Zeilinger et al., 2016; Ruoff et al., 2020). Spheroid models were observed to be more sensitive for PFASs, but data within these more complex models was not available for the identified outlier chemicals that are known to require metabolic activation, such as cyclophosphamide, AFB1, and B[a]P.

In our study, the emphasis was on the derivation of protective tPODs that did not dive into the underlying mechanistic data to predict mode-of-action. However, it is acknowledged that the mechanism of toxicity can be an important consideration as specific tested cell lines and cell types have shown cell-dependent differential sensitivity to specific chemicals (Lawal and Ellis, 2010; Robert et al., 2014). Further to this,

immortalized and transformed cell lines that have abnormal/unstable karyotypes may not produce the transcriptomic responses that are consistent with expected responses in normal human tissues (Kleensang et al., 2016). Cell monocultures do not reflect the complexity of organisms or represent the heterogeneity of the human population. As mentioned previously, metabolism (biotransformation) of chemicals *in vivo* may be different from that observed in exposed cell lines (Wilk-Zasadna et al., 2015). Metabolic pathways and resultant active by-products or metabolites are often difficult to fully predict but are also important considerations for method development and integration that would be required in the context of the paradigm shift to non-animal testing and assessment approaches. Overall, there are a variety of potential uncertainties that must be brought to bear relating to the *in vitro* cell models used. These aspects of non-animal models are being widely studied in parallel to improve the understanding related to impact for use in the derivation of effect levels and to develop methods that will address the inherent challenges of *in vitro* models. Increasingly *in vivo*-relevant models will inevitably lead to increased confidence in NAM-based approaches in the future. In the interim and to advance this area, there is a continued need to evaluate the fit-for-purpose use of the NAM in the context of the information gap being addressed. In parallel, the way in which apical effects from animal studies are used for regulatory decision making in light of their respective uncertainties must be reflected upon.

Along with cell system and model, the selection of dose-range and dose-spacing were identified as sources of potential uncertainty. The range and spacing of exposure

concentrations of chemicals within the design can influence the potency of chemicals represented by derived tPODs and subsequent conversion to AEDs. Inaccurate selection of concentrations compromises the ability of *in vitro* models to reproduce the initiated changes as a result of exposures that are representative of *in vivo* biology, influencing the biological activity and initiation of a given mode of action. This consideration goes beyond the aim of the current work to produce conservative and protective estimates and may influence the outcomes of a predictive toxicology analysis aimed at providing a biological basis for the selection of tPODs.

B.5 Conclusion

Overall, this meta-analysis provides evidence that *in vitro* tPODs and AEDs, in the majority of cases, are equal to or more protective estimates when compared to those derived using traditional animal toxicity tests. The results support that there is a high level of correlation between the different approaches evaluated to derive tPODs. However, we caution the use of the fifth percentile for *in vitro* transcriptomics, because of the considerable influence from the top concentrations within these studies on these tPODs. The diversity of chemicals and experimental conditions within the dataset analyzed revealed sources of uncertainties for consideration when shifting to the use of NAM-based data in regulatory applications. Nonetheless, the workflow used here produced multiple AEDs that could be used in screening and prioritization to flag chemicals of greater

potential concern for further assessment. Our findings support the current movement transitioning to the use of non-animal test methods in chemical risk assessment, and align with global initiatives that include, but are not limited to, the inter-governmental accelerating the pace of chemical risk assessment (APCRA) collaboration, numerous omics-based initiatives within the OECD, and speciality NAM-based working groups within the Health and Environmental Sciences Institute (HESI). It was demonstrated that transcriptomics reliably detected perturbations in gene expression as a result of chemical exposure within *in vitro* models, outcomes that support the first principle in the proposed logic framework to incorporate omics-based data into the regulatory chemical testing and assessment paradigm (Johnson et al., 2022).

B.6 Tables and Figures

Table B.1: List of chemicals with available in vitro datasets.

Dataset	Chemical Name	Abbreviated Name	CASRN	Model(s)	Exposure (Days)
OECD 2022	2-(4-hydroxyphenyl)sulfonylphenol	2,4-BPS	5397-34-2	MCF-7	2
OECD 2022	2-[(4-Hydroxyphenyl)methyl]phenol	2,4-BPF	2467-03-0	MCF-7	2
OECD 2022	2-[4-(Benzyloxy)benzene-1-sulfonyl]phenol	BPS-MPE	63134-33-8	MCF-7	2
OECD 2022	4-(4-hydroxyphenyl)sulfonylphenol	4,4-BPS	80-09-1	MCF-7	2
OECD 2022	4-[(4-hydroxyphenyl)methyl]phenol	4,4-BPF	620-92-8	MCF-7	2
OECD 2022	4-[(4-Isopropoxyphenyl)sulfonyl]phenol	D8	95235-30-6	MCF-7	2

OECD 2022	4,4'-Bis(p-tolylsulfonylureido)diphenylmethane	BTUM	151882-81-4	MCF-7	2
OECD 2022	4,4'-Sulfonylbis[2-(prop-2-en-1-yl)phenol]	TGSA	41481-66-7	MCF-7	2
OECD 2022	4-{4-[(Prop-1-en-2-yl)oxy]benzene-1-sulfonyl}phenol	BPS-MAE	97042-18-7	MCF-7	2
OECD 2022	Bis (4-chlorophenyl) Sulfone	Bis4CPS	80-07-9	MCF-7	2
OECD 2022	Bisphenol A	BPA	80-05-7	MCF-7	2
OECD 2022	Bisphenol A diglycidyl ether	BADGE	1675-54-3	MCF-7	2
OECD 2022	Bisphenol AF	BPAF	1478-61-1	MCF-7	2
OECD 2022	Bisphenol AP	BPAP	1571-75-1	MCF-7	2
OECD 2022	Bisphenol C	BPC	14868-03-2	MCF-7	2
OECD 2022	Dexamethasone	Dex	50-02-2	MCF-7	2
OECD 2022	B-Estradiol	Estradiol	50-28-2	MCF-7	2
OECD 2022	Pergafast 201	Perg201	232938-43-1	MCF-7	2
Harrill et al.	3,5,3'-Triiodothyronine	Triiodothyronine	6893-02-3	MCF-7	0.25
Harrill et al.	4-Cumylphenol		599-64-4	MCF-7	0.25
Harrill et al.	4-Hydroxytamoxifen		68392-35-8	MCF-7	0.25
Harrill et al.	4-Nonylphenol (branched)	4-Nonylphenol	84852-15-3	MCF-7	0.25
Harrill et al.	Amiodarone hydrochloride	Amiodarone HCl	19774-82-4	MCF-7	0.25
Harrill et al.	Atrazine		1912-24-9	MCF-7	0.25
Harrill et al.	Bifenthrin		82657-04-3	MCF-7	0.25
Harrill et al.	Bisphenol A	BPA	80-05-7	MCF-7	0.25
Harrill et al.	Bisphenol B	BPB	77-40-7	MCF-7	0.25
Harrill et al.	Butafenacil		134605-64-4	MCF-7	0.25
Harrill et al.	Cladribine		4291-63-8	MCF-7	0.25
Harrill et al.	Clofibrate		637-07-0	MCF-7	0.25
Harrill et al.	Clomiphene citrate (1:1)	Clomiphene Cit	50-41-9	MCF-7	0.25
Harrill et al.	Cyanazine		21725-46-2	MCF-7	0.25
Harrill et al.	Cycloheximide		66-81-9	MCF-7	0.25
Harrill et al.	Cypermethrin		52315-07-8	MCF-7	0.25

Harrill et al.	Cyproconazole		94361-06-5	MCF-7	0.25
Harrill et al.	Cyproterone acetate	Cyproterone Ace	427-51-0	MCF-7	0.25
Harrill et al.	Farglitazar		196808-45-4	MCF-7	0.25
Harrill et al.	Fenofibrate		49562-28-9	MCF-7	0.25
Harrill et al.	Fenpyroximate (Z,E)		111812-58-9	MCF-7	0.25
Harrill et al.	Flutamide		13311-84-7	MCF-7	0.25
Harrill et al.	Fomesafen		72178-02-0	MCF-7	0.25
Harrill et al.	Fulvestrant		129453-61-8	MCF-7	0.25
Harrill et al.	Genistein		446-72-0	MCF-7	0.25
Harrill et al.	Imazalil		35554-44-0	MCF-7	0.25
Harrill et al.	Lactofen		77501-63-4	MCF-7	0.25
Harrill et al.	Lovastatin		75330-75-5	MCF-7	0.25
Harrill et al.	Maneb		12427-38-2	MCF-7	0.25
Harrill et al.	Nilutamide		63612-50-0	MCF-7	0.25
Harrill et al.	Perfluorooctanesulfonic acid	PFOS	1763-23-1	MCF-7	0.25
Harrill et al.	Perfluorooctanoic acid	PFOA	335-67-1	MCF-7	0.25
Harrill et al.	Prochloraz		67747-09-5	MCF-7	0.25
Harrill et al.	Propiconazole		60207-90-1	MCF-7	0.25
Harrill et al.	Pyraclostrobin		175013-18-0	MCF-7	0.25
Harrill et al.	Reserpine		50-55-5	MCF-7	0.25
Harrill et al.	Rotenone		83-79-4	MCF-7	0.25
Harrill et al.	Simazine		122-34-9	MCF-7	0.25
Harrill et al.	Simvastatin		79902-63-9	MCF-7	0.25
Harrill et al.	Sirolimus		53123-88-9	MCF-7	0.25
Harrill et al.	Tetrac		67-30-1	MCF-7	0.25
Harrill et al.	Thiram		137-26-8	MCF-7	0.25

Harrill et al.	Trichostatin A		58880-19-6	MCF-7	0.25
Harrill et al.	Trifloxystrobin		141517-21-7	MCF-7	0.25
Harrill et al.	Troglitazone		97322-87-7	MCF-7	0.25
Harrill et al.	Vinclozolin		50471-44-8	MCF-7	0.25
Harrill et al.	Ziram		137-30-4	MCF-7	0.25
Ramaiahgari et al.	Acetaminophen	APAP	103-90-2	HepaRG	4
Ramaiahgari et al.	Aflatoxin B1	AFB1	1162-65-8	HepaRG	4
Ramaiahgari et al.	Aspirin		50-78-2	HepaRG	4
Ramaiahgari et al.	Benzo(a)pyrene	B[a]P	50-32-8	HepaRG	4
Ramaiahgari et al.	Caffeine		58-08-2	HepaRG	4
Ramaiahgari et al.	Chenodeoxycholic acid	CDCA	474-25-9	HepaRG	4
Ramaiahgari et al.	Chlorpromazine	CPZ	50-53-3	HepaRG	4
Ramaiahgari et al.	Cyclophosphamide monohydrate	Cyclophosphamide	6055-19-2	HepaRG	4
Ramaiahgari et al.	Diphenhydramine hydrochloride	Diphenhydramine	147-24-0	HepaRG	4
Ramaiahgari et al.	Fenofibric acid	FFA	42017-89-0	HepaRG	4
Ramaiahgari et al.	Levofloxacin hydrate	Levofloxacin	138199-71-0	HepaRG	4
Ramaiahgari et al.	Menadione		58-27-5	HepaRG	4
Ramaiahgari et al.	N-Nitrosodimethylamine	DMN	62-75-9	HepaRG	4
Ramaiahgari et al.	Omeprazole	OMP	73590-58-6	HepaRG	4
Ramaiahgari et al.	Potassium chloride	KCl	7447-40-7	HepaRG	4
Ramaiahgari et al.	Rifampicin	RIF	13292-46-1	HepaRG	4
Ramaiahgari et al.	Ritonavir		155213-67-5	HepaRG	4
Ramaiahgari et al.	Rosiglitazone		122320-73-4	HepaRG	4
Ramaiahgari et al.	Sucrose		57-50-1	HepaRG	4

Ramaiahgari et al.	Tamoxifen		10540-29-1	HepaRG	4
Ramaiahgari et al.	Troglitazone		97322-87-7	HepaRG	4
Ramaiahgari et al.	Trovafloracin mesylate	Trovafloracin	147059-75-4	HepaRG	4
Ramaiahgari et al.	Valproic acid	VPA	99-66-1	HepaRG	4
Ramaiahgari et al.	Phenobarbital sodium	PB	57-30-7	HepaRG	4
PFAS	2H,2H,3H,3H-Perfluorooctanoic acid	5:3 Acid	914637-49-3	Spheroids	1, 10
PFAS	4:2 Fluorotelomer sulfonic acid	4:2 FtS	757124-72-4	Spheroids	1, 10
PFAS	6:2 Fluorotelomer alcohol	6:2 FtOH	647-42-7	Spheroids	1, 10
PFAS	6:2 Fluorotelomer phosphate monoester	6:2 monoPAP	57678-01-0	Spheroids	1, 10
PFAS	6:2 Fluorotelomer sulfonic acid	6:2 FtS	27619-97-2	Spheroids	1, 10
PFAS	8:2 Fluorotelomer alcohol	8:2 FtOH	678-39-7	Spheroids	1, 10
PFAS	8:2 Fluorotelomer phosphate monoester	8:2 monoPAP	57678-03-2	Spheroids	1, 10
PFAS	8:2 Fluorotelomer sulfonic acid	8:2 FtS	39108-34-4	Spheroids	1, 10
PFAS	Perfluorobutanesulfonic acid	PFBS	375-73-5	Spheroids	1, 4, 10, 14
PFAS	Perfluorobutanoic acid	PFBA	375-22-4	Spheroids	1, 10
PFAS	Perfluorodecanesulfonic acid	PFDS	335-77-3	Spheroids	1, 4, 10, 14
PFAS	Perfluorodecanoic acid	PFDA	335-76-2	Spheroids	1, 10
PFAS	Perfluoroheptanesulfonic acid	PFHpS	375-92-8	Spheroids	1, 10
PFAS	Perfluoroheptanoic acid	PFHpA	375-85-9	Spheroids	1, 10
PFAS	Perfluorohexanesulfonic acid	PFHxS	355-46-4	Spheroids	1, 10
PFAS	Perfluorohexanoic acid	PFHxA	307-24-4	Spheroids	1, 10
PFAS	Perfluorononanoic acid	PFNA	375-95-1	Spheroids	1, 10
PFAS	Perfluorooctanoic acid	PFOA	335-67-1	Spheroids	1, 4, 10, 14
PFAS	Perfluorooctanesulfonic acid	PFOS	1763-23-1	Spheroids	1, 4, 10, 14

PFAS	Perfluorooctanesulfonamide	PFOSA	754-91-6	Spheroids	1, 10
PFAS	Perfluoropentanoic acid	PFPeA	2706-90-3	Spheroids	1, 10
PFAS	Perfluorotetradecanoic acid	PFTeDA	376-06-7	Spheroids	1, 10
PFAS	Perfluoroundecanoic acid	PFUnA	2058-94-8	Spheroids	1, 10
Buick et al.	2-Deoxy-D-glucose	2DD-Glucose	154-17-6	HepaRG	2
Buick et al.	Aflatoxin B1	AFB1	1162-65-8	HepaRG	2
Buick et al.	Benzo(a)pyrene	B[a]P	50-32-8	HepaRG	2
Buick et al.	Cisplatin		15663-27-1	HepaRG	2
Buick et al.	Cyclophosphamide monohydrate	Cyclophosphamide	6055-19-2	HepaRG	2
Buick et al.	Cytosine arabinoside	Cyt arabinoside	147-94-4	HepaRG	2
Buick et al.	Eugenol		97-53-0	HepaRG	2
Buick et al.	Methyl methanesulfonate	M-mSulfonate	66-27-3	HepaRG	2
Buick et al.	N-Nitroso-N-ethylurea	N-Nitrosourea	759-73-9	HepaRG	2
Buick et al.	Propyl gallate		121-79-9	HepaRG	2
Buick et al.	Urea		57-13-6	HepaRG	2
Buick et al.	Zidovudine (azidothymidine)		30516-87-1	HepaRG	2

OECD 2022—Listing of Bisphenols from Health Canada OECD Case-study, Harrill et al.—Listing of chemicals from published dataset in Harrill et al. (2021), Ramaiahgari et al.—Listing of chemicals from published dataset in Ramaiahgari et al. (2019), PFAS—Listing of chemicals from published datasets in Reardon et al. (2021) and Rowan-Carroll et al. (2021), Buick et al.—Listing of chemicals from published dataset in Buick et al. (2021).

Table B.2: List of chemicals with available in vivo datasets.

	Name	CASRN	Animal	Study Type	Method	Effect Level (mg/kg-bw/day)
EPA ToxVal	4-Cumylphenol	599-64-4	Rat	Repeat Dose	NOAEL	50
	AFB1	1162-65-8	Human	Repeat Dose	BMDL01	0.000078
	Atrazine	1912-24-9	Mouse	Reproductive	NOEL	0.001
	Bifenthrin	82657-04-3	Rat	Developmental	NOAEL	1
	Bis4CPS	620-92-8	Rat	Repeat Dose	LOEL	20
	BPA	80-05-07	Rat	Developmental	NOAEL	0.015
	Cyanazine	21725-46-2	Rat	Repeat Dose	NOAEL	0.005
	Cyclophosphamide	6055-19-2		Repeat Dose	Cancer	0.57
	Cyproterone ace	427-51-0	Mouse	Repeat Dose	NOAEL	125
	Eugenol	97-53-0	Rat	Repeat Dose	NOAEL	57
	Fenofibrate	49562-28-9	Mouse	Developmental	LEL	11.7
	Flutamide	13311-84-7	Rat	Repeat Dose	NOAEL	10
	Genistein	446-72-0	Rat	Developmental	NEL	20
	Lovastatin	75330-75-5	Mouse	Repeat Dose	NOAEL	30
	PFHxA	307-24-4	Rat	Repeat Dose	NOAEL	200
	PFOA	335-67-1	Mouse	Reproductive	LOEL	0.02
	Prochloraz	67747-09-5	Dog	Repeat Dose	NOAEL	2.5
	Propiconazole	60207-90-1	Mouse	Repeat Dose	NOAEL	2.7
	Pyraclostrobin	175013-18-0	Rat	Repeat Dose	NOAEL	3.4
	Reserpine	50-55-5	Mouse	Repeat Dose	NOAEL	0.12
	Rotenone	83-79-4	Rat	Developmental	LOAEL	0.75
	Simazine	122-34-9	Mouse	Reproductive	NOEL	0.005
	Trifloxystrobin	141517-21-7	Rabbit	Developmental	NOAEL	10
	Troglitazone	97322-87-7	Mouse	Repeat Dose	NOAEL	1200

	Vinclozolin	50471-44-8	mouse	Reproductive	LOEL	1
	Zidovudine	30516-87-1	Mouse	Repeat Dose	LEL	100
REACH	2DD-Glucose	66-13-6	rat	Reproductive	NOAEL	20
	BPAF	1478-61-1	Rat	Repeat Dose	NOAEL	3.5
	BPS	80-09-01	Rat	Reproductive / Developmental	NOAEL	10
	Cyproconazole	94361-06-5	Rat	Reproductive	NOAEL	1
	Estradiol	50-28-2	Rabbit	Developmental	NOEL	0.0003
	Lactofen	77501-63-4	Rat	Reproductive	NOEL	2.5
	TGSA	41481-66-7	Rat	Repeat Dose	NOEL	15
	Thiram	137-26-8	Dog	Repeat Dose	NOEL	0.84
	Urea	57-13-6		Developmental	NOAEL	500
OECD	B[a]P	50-32-8		Repeat Dose	LOEL	0.05
	Fenpyroximate	111812-58-9		Reproductive	LOEL	8.45
	Imazalil	35554-44-0		Reproductive	NOAEL	5
	PFOS	1763-23-1		Reproductive	NOAEL	0.03
HC	Propyl gallate	121-79-9		Repeat Dose	NOAEL	135

Table B.3: Summary table of advantages, disadvantages, and results of evaluated approaches for derivation of transcriptomic points of departure.

	Result	Advantages	Disadvantages	Result	Advantages
The 5th percentile	The 5th percentile is a conservative value for tPOD derivation that targets the lowest and most responsive genes. Produces the highest potency ranking for select substances, even for chemicals identified as 'non-toxic' reference chemicals. Not considered as a reliable approach to derive tPODs using gene expression data.	Simplified calculation of estimate that uses the 5th percentile gene from the full distribution of genes with modeled BMCs. No additional models or calculations are required beyond BMD modeling software.	Subject to influence from the experimental design (e.g., the top dose/concentration). Subject to influence by the extent of transcriptional change (i.e., total number of genes with BMCs).	The 5th percentile	The 5th percentile is a conservative value for tPOD derivation that targets the lowest and most responsive genes. Produces the highest potency ranking for select substances, even for chemicals identified as 'non-toxic' reference chemicals. Not considered as a reliable approach to derive tPODs using gene expression data.
1st Mode ^a	The 1st mode uses those genes from the identified first mode based on the frequency of distribution of genes with BMCs. Considered the least conservative but is a consistent and simple approach to derive tPODs using distributions.	Consistently low (i.e., conservative) estimate across majority of listed chemicals. Corresponds with disrupted biomolecular targets and pathways, and provides an estimate of the initiation of molecular events Judson et al. (2016)	Requires sufficient biological activity (i.e., genes with BMCs) for identification and calculation of modes.	1st Mode ^a	The 1st mode uses those genes from the identified first mode based on the frequency of distribution of genes with BMCs. Considered the least conservative but is a consistent and simple approach to derive tPODs using distributions.

25th Ranked Gene ^a	The 25th ranked gene is set as a threshold that is identified from a relative potency ranking of the lowest to highest genes with BMCs. Considered a simple approach deriving tPODs using the BMC distribution that represents approximately 0.1 % of the genome.	Simplified calculation using the 25th ranked gene by potency based on the distribution of genes with modeled BMCs. No additional models or calculations are required beyond BMD modeling software.	Requires a minimal amount of biological activity (e.g., 25 concentration-responsive genes). Excludes select chemicals with insufficient biological activity (e.g., either non-liver toxic, or highly cytotoxic chemicals that do not meet a minimum required concentration-responsive genes).	25th Ranked Gene ^a	The 25th ranked gene is set as a threshold that is identified from a relative potency ranking of the lowest to highest genes with BMCs. Considered a simple approach deriving tPODs using the BMC distribution that represents approximately 0.1 % of the genome.
LCRD ^a	The LCRD is the value of the lowest of rank-ordered BMCs using consistent response groups of BMCs Crizer et al. (2021) . Represents the primary approach to derive conservative tPODs for most of the listed chemicals that were carried forward for comparison with apical PODs.	Uses a calculated derivation for tPOD derivation that represents a point of toxicological relevance that is considered a consistent response of all biological features. Predominantly the lowest and most conservative tPOD of all approaches based on the distribution of BMCs across all chemicals, and suggested as a promising candidate for deriving	Requires more complex modeling to obtain a relevant tPOD based on the literature.	LCRD ^a	The LCRD is the value of the lowest of rank-ordered BMCs using consistent response groups of BMCs Crizer et al. (2021) . Represents the primary approach to derive conservative tPODs for most of the listed chemicals that were carried forward for comparison with apical PODs.

		protective and conservative tPODs.			
Lowest Gene Set ^a	The lowest and most sensitive gene set is widely accepted as a means of obtaining a tPOD as outlined in a guidance document National Toxicology Program (2018) . Reliable approach in scenarios with sufficient biological activity to derive tPODs using the median value of the lowest and most potent (i.e., sensitive) gene set. tPODs primarily defined using annotations from the largest and most comprehensive database (e.g., GO).	Established and reliable approach to derive tPODs using dose-response models that reflect meaningful changes in biology while reducing the influence of background. Option to choose from multiple available databases to represent annotations used to define gene sets (e.g., GO, KEGG or REACTOME).	Requires more complex modeling and additional parameters and filtering to obtain a relevant tPOD based gene sets. Requires sufficient biological activity (i.e., genes with BMCs) to define subsets of genes related to gene sets and pathways.	Lowest Gene Set ^a	The lowest and most sensitive gene set is widely accepted as a means of obtaining a tPOD as outlined in a guidance document National Toxicology Program (2018) . Reliable approach in scenarios with sufficient biological activity to derive tPODs using the median value of the lowest and most potent (i.e., sensitive) gene set. tPODs primarily defined using annotations from the largest and most comprehensive database (e.g., GO).

^aApproaches that were considered to be alternatives to be applied in lieu of approaches using percentiles (e.g., 5th) to derive a transcriptomic point of departure.

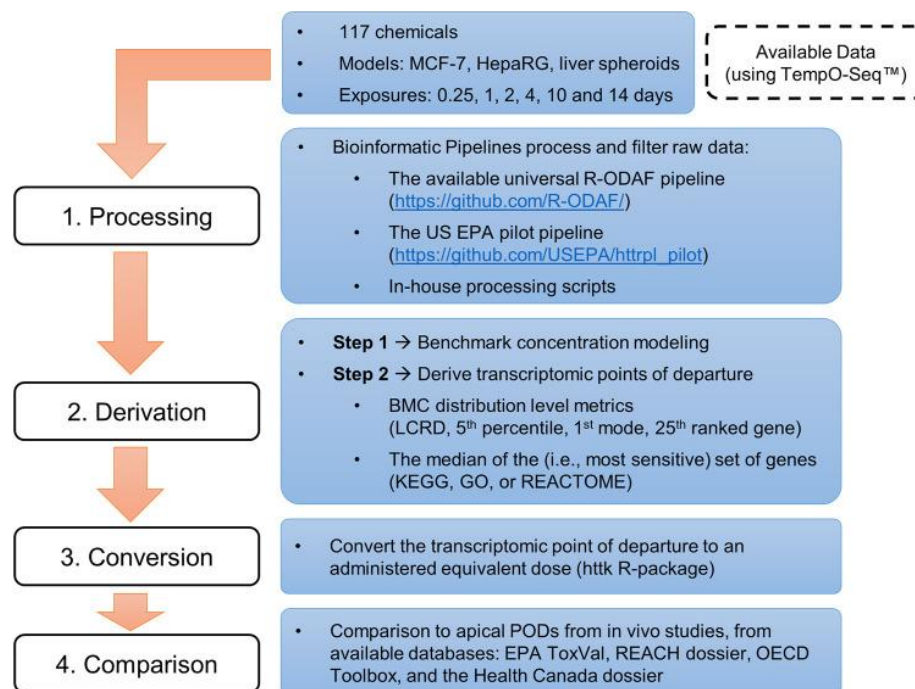


Figure B.1: Overview of the workflow used to compile available datasets and derive in vitro points of departure for comparison with in vivo apical points of departure from curated databases.

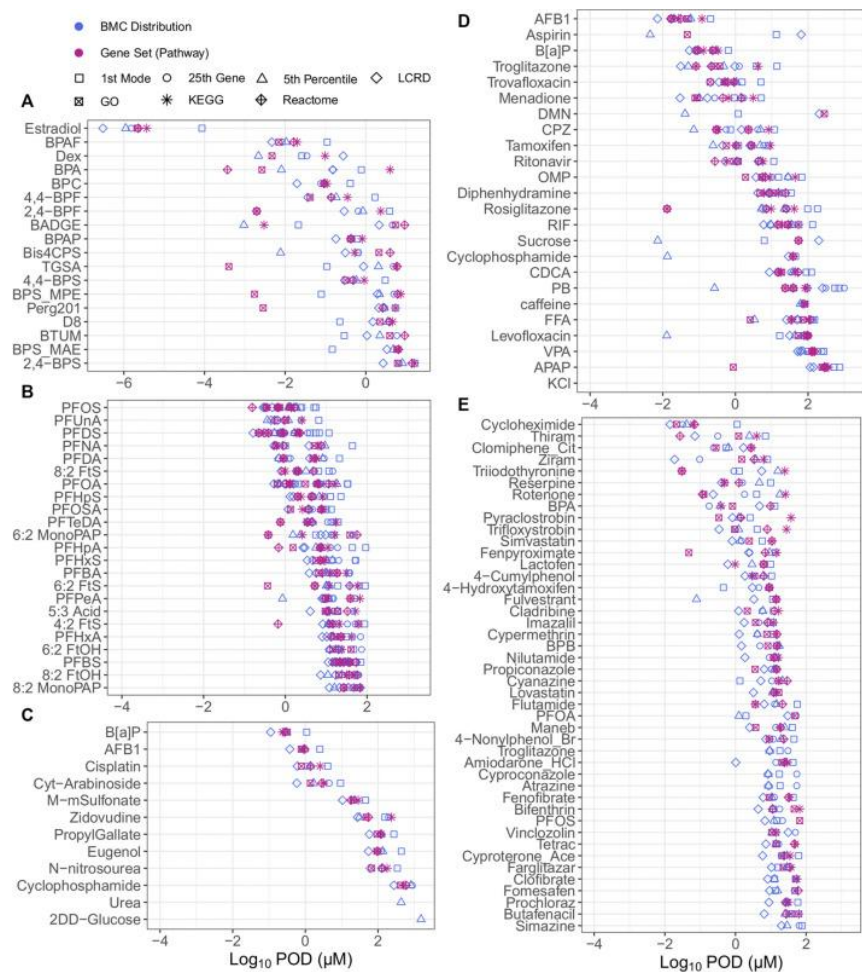


Figure B.2: All tPODs of chemicals from the bisphenol case study group from Health Canada data in human MCF7 cells (A), PFAS data from Reardon et al. and Rowan-Carroll et al. in human liver spheroids (B), Buick et al., from exposed human liver HepaRG cells (C), Ramaiahgari et al. from exposed human liver HepaRG cells (D) and Harrill et al. in human MCF7 cells (E).

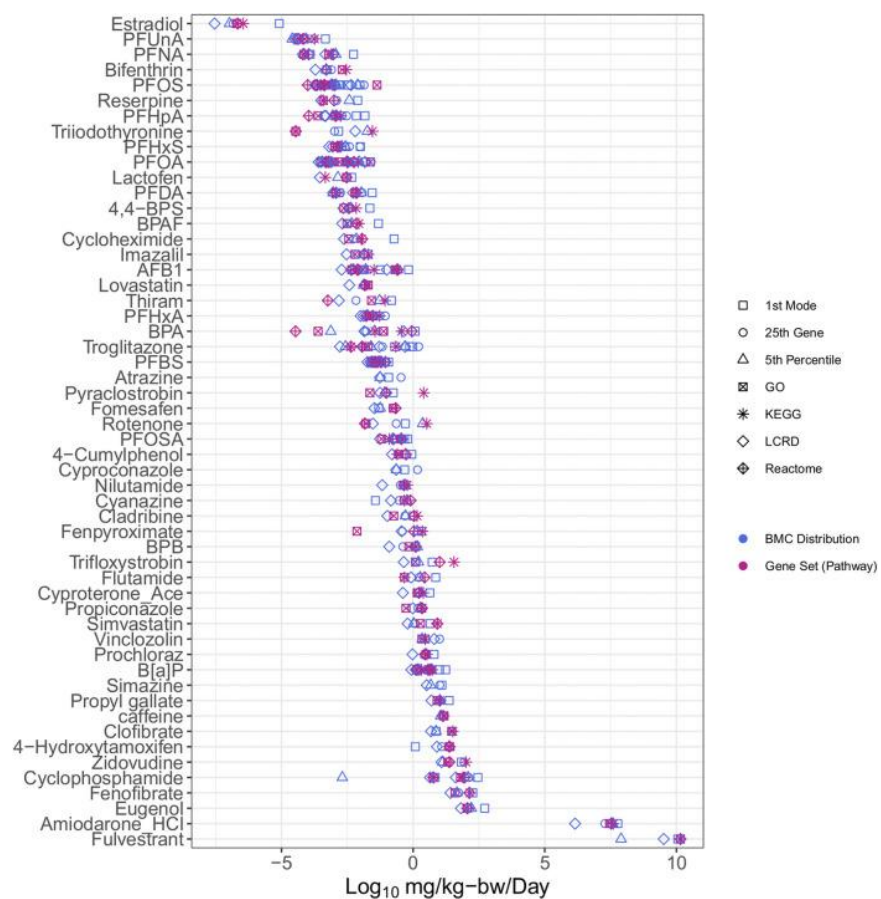


Figure B.3: All case study chemicals with available data for conversion using IVIVE ranked by median potency of AEDs.

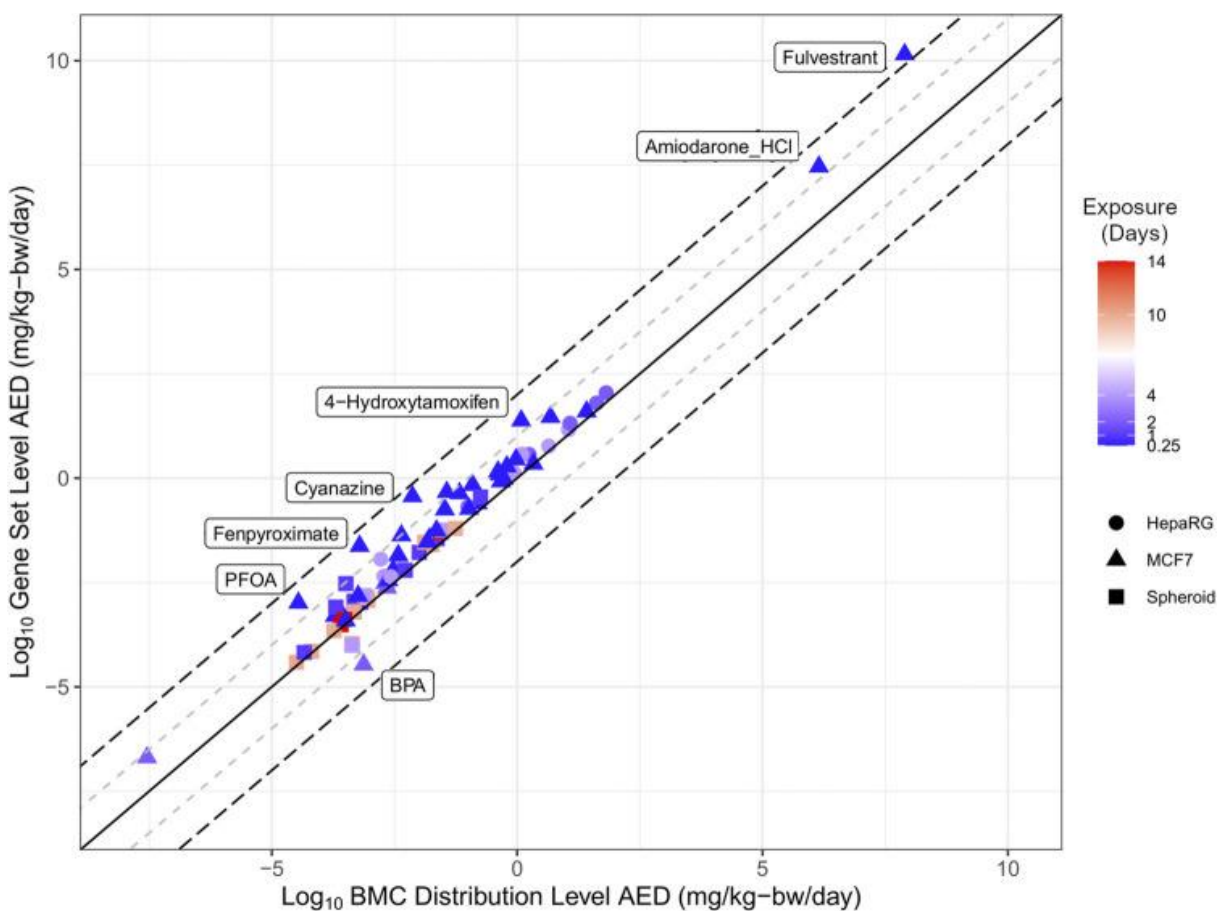


Figure B.4: Correlation plot of the relationship between the lowest derived AED from using the BMC distribution level and gene set level approaches. The solid line represents perfect agreement between data on the x-axis with data on the y-axis, with a range of one (+/-1 log₁₀, grey dashed line) or two (+/-2 log₁₀, black dashed line) orders of magnitude.

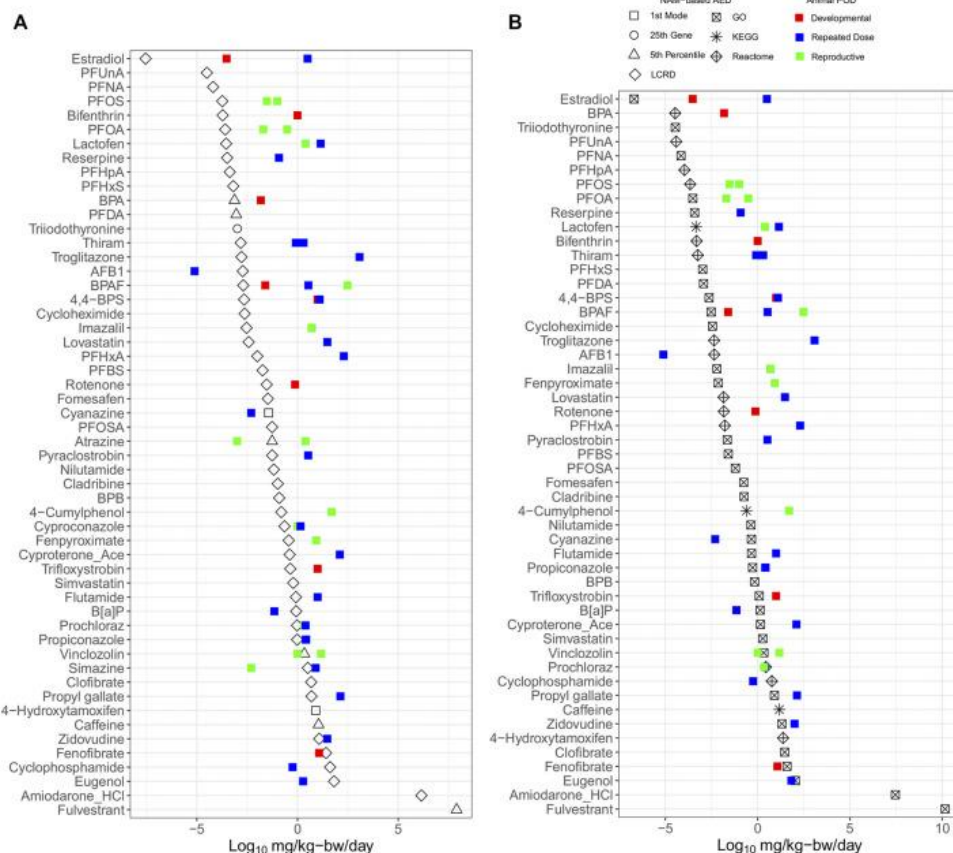


Figure B.5: Relative potency ranking of points of departure for each chemical selected from the lowest value derived using BMC distribution (A) and gene-set (B) level AEDs, compared to in vivo apical points of departure.

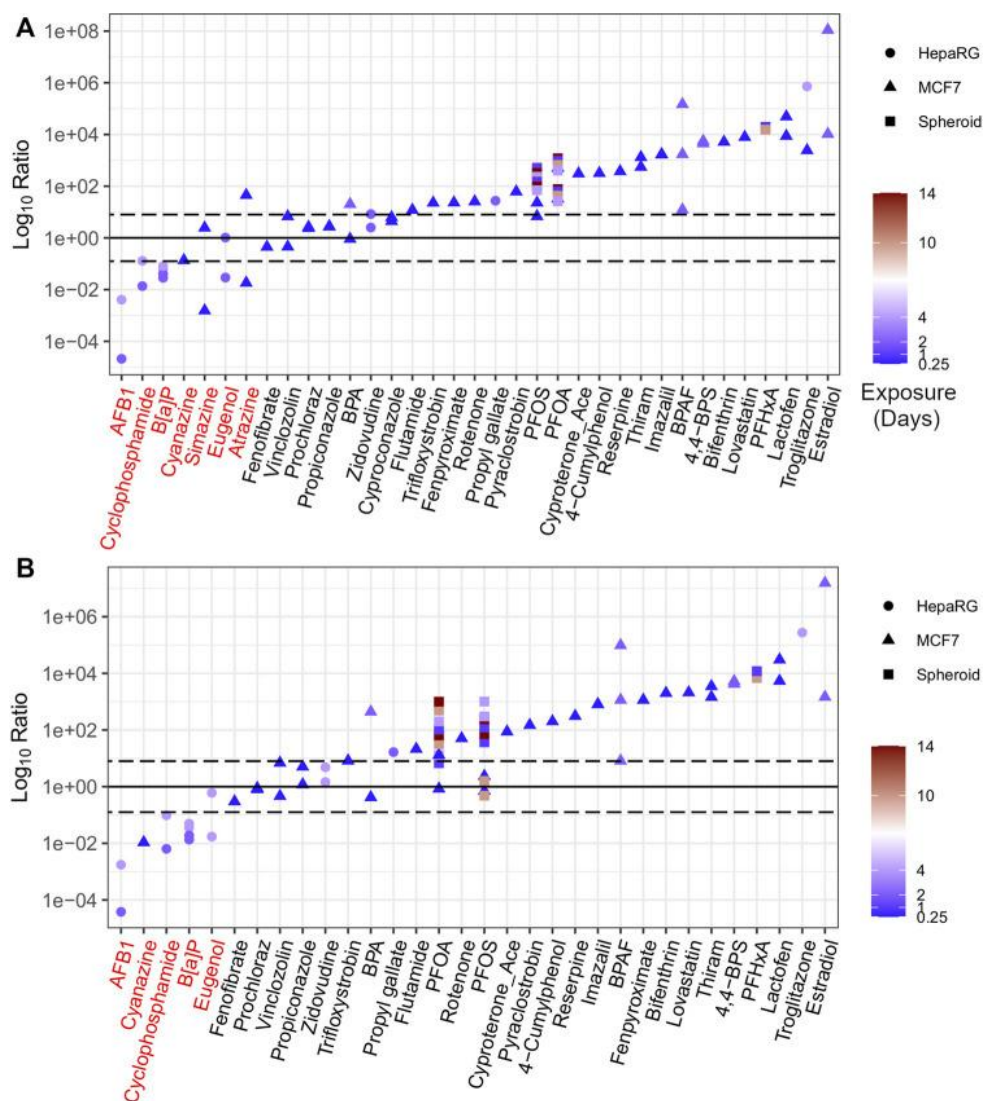


Figure B.6: The ratio of in vitro derived AEDs to apical PODs ranked using BMC distribution level (A) and gene set level (B) approaches. Chemicals highlighted in red represent outliers based on ratio less than 0.

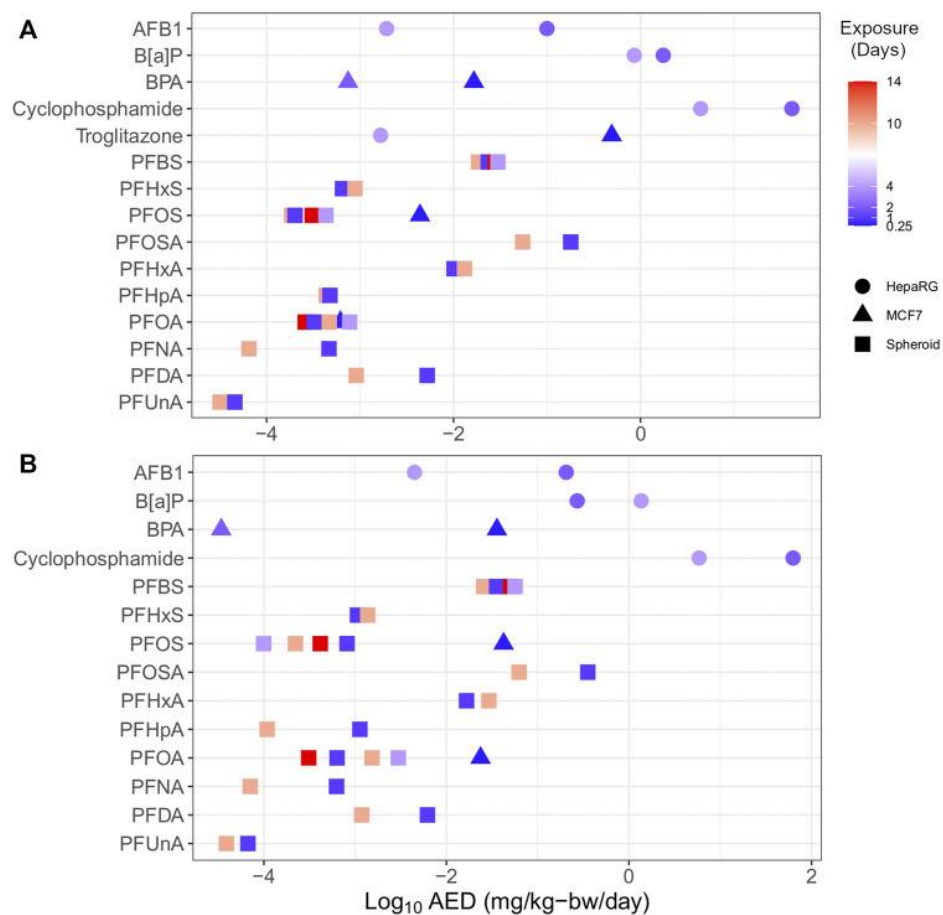


Figure B.7: Plot of substances with multiple endpoints demonstrating the influence of time (by colour; range 0.25–14 days) and model (shape) on transcriptomic AEDs derived from in vitro data using BMC distribution level (A) and gene set level (B) approaches.

Appendix C: Supplementary Material

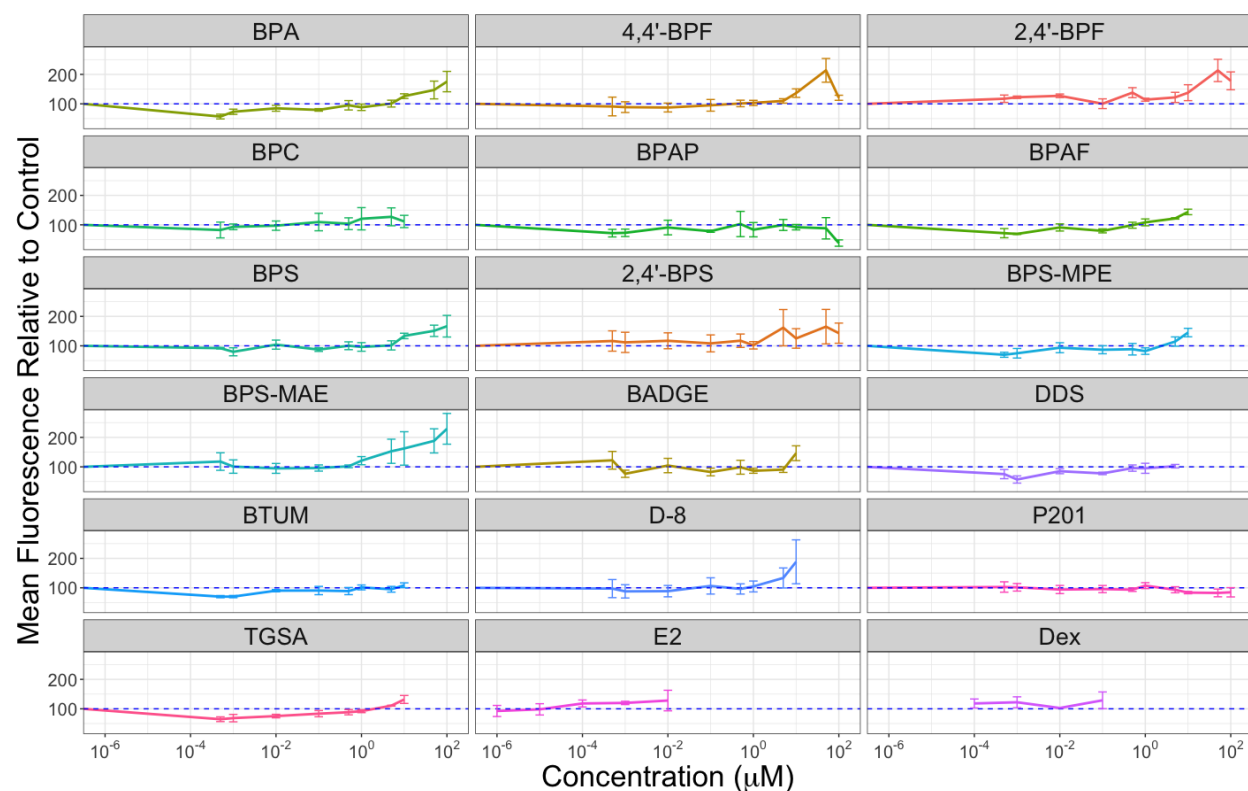
C.1 Chapter 2

Supplementary Table 2.1: Summary of stress response biomarkers of MCF-7 cells (n = 3 – 4) exposed to BPA alternative chemicals at a range of concentrations (0.0005 - 100 μ M) for 48 hours. Biomarkers were activated (orange) or inhibited (blue) based on thresholds for significance $-\log(p\text{-value}) \geq 4$ or ≤ -4 .

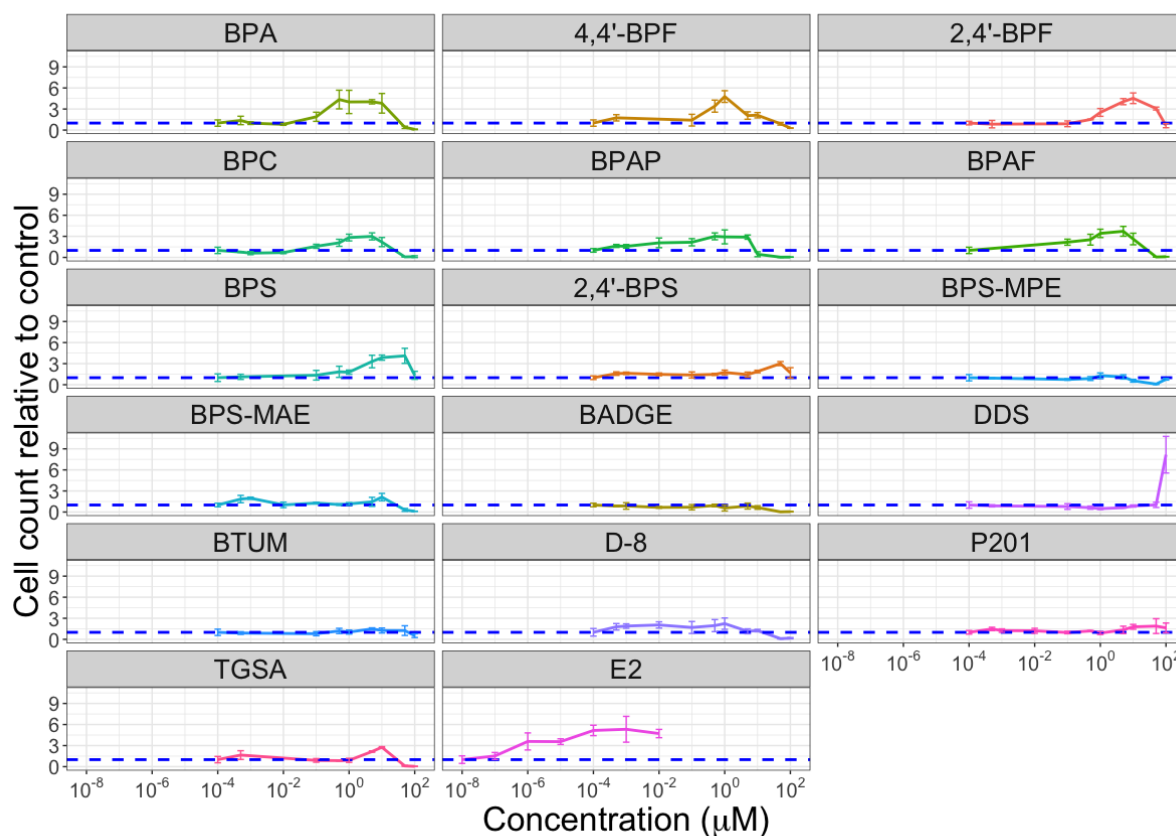
Chemical	HDACi	NFkB	HSF1	Nrf2	MTF1	TGx-DDI
2,4'-BPF		50		50, 100		
2,4'-BPS		100				
4,4'-BPF	5, 10, 50, 100			100		10, 50
BADGE					0.1, 0.5, 1	
BPA	5, 10, 50			100	1, 10 100	10
BPAF	1				1, 5	0.1
BPAP	5, 10, 50	100	50, 100	50, 100	50, 100	100
BPC	5, 10				5, 10	
BPS	5	100			0.001	
BPS-MAE				100	100	
BPS-MPE						
BTUM					5	
D-8						
DDS						
P201		100			0.01, 0.1, 1, 10, 100	
TGSA						

Supplementary Table 2.2: Summary of concentrations removed from analysis. Samples were removed due to technical error (TE), visible precipitation (green), activating two or more stress response biomarkers (red; see Supplementary File 1.3), significantly decreasing cell proliferation (yellow; see Supplementary Figure 2.2), or both activating biomarkers and decreasing cell proliferation (orange).

Chemical	0.0005 μM	0.001 μM	0.01 μM	0.1 μM	0.5 μM	1 μM	5 μM	10 μM	50 μM	100 μM
2,4'-BPF									2 stress BMs	↓proliferation
2,4'-BPS										
4,4'-BPF										↓proliferation
BADGE									precipitated	
BPA		TE	TE						↓proliferation	2 BMs & ↓prolif
BPAF									precipitated	
BPAP								↓proliferation	2 BMs & ↓prolif	
BPC									precipitated	
BPS										
BPS-MAE									↓proliferation	2 BMs & ↓prolif
BPS-MPE									precipitated	
BTUM									precipitated	
D-8									precipitated	
DDS								precipitated		
P201										2 stress BMs
TGSA									precipitated	



Supplementary Figure 2.1: Cell viability using a CellTiter-Blue Cell Viability Assay. MCF-7 cells ($n = 3 - 4$) were exposed to BPA alternative chemicals at a range of concentrations (0.0005 - 100 μM) for 48 hours and are compared to their respective DMSO (0.1%) control samples. First row of chemicals from left to right: BPA, 4,4'-BPF, 2,4'-BPF. Second row: BPC, BPAP, BPAF. Third row: BPS, 2,4'-BPS, BPS-MPE. Fourth Row: BPS-MAE, BADGE, DDS. Fifth row: BTUM, D-8, P201. Fifth row: TGSA, E2, Dex. Blue horizontal dotted line denotes 100% cell viability relative to controls. There were no significant decreases in viability relative to controls.



Supplementary Figure 2.2: Cell proliferation assay. MCF-7 cells ($n = 3 - 4$) were exposed to BPA alternative chemicals at a range of concentrations (0.0001 - 100 μM) for 7 days and are compared to their respective DMSO (0.1%) control samples. First row of chemicals from left to right: BPA, 4,4'-BPF, 2,4'-BPF. Second row: BPC, BPAP, BPAF. Third row: BPS, 2,4'-BPS, BPS-MPE. Fourth Row: BPS-MAE, BADGE, DDS. Fifth row: BTUM, D-8, P201. Fifth row: TGSA, E2. Blue horizontal dotted line denotes 100% cell proliferation relative to controls. The following chemicals increased proliferation relative to controls: BPA (0.0005, 0.1, 0.5, 1, 5, 10 μM), 4,4'-BPF (0.0005, 0.1, 0.5, 1, 5, 10 μM), 2,4'-BPF (0.5 – 50 μM), BPC (0.1 – 10 μM), BPAP (0.0005 – 5 μM), BPAF (0.1 – 10 μM), BPS (0.0005, 0.1, 0.5, 1, 5, 10, 50, 100 μM), 2,4'-BPS (0.0005 – 100 μM), BPS-MPE (1, 5 μM), BPS-MAE (0.0005 – 10 μM), DDS (50, 100 μM), BTUM (0.5 – 50 μM), D-8 (0.0005 – 10 μM), P201 (0.0005, 0.001, 0.01, 0.5, 5, 10, 50, 100), TGSA (0.0005, 0.5, 10 μM), E2 (0.0000001 – 0.01 μM). The following chemicals decreased cell proliferation below 50% relative to solvent control: BPA (50, 100 μM), 4,4'-BPF (100 μM), BPAP (10, 50, 100 μM), BPS-MAE (50, 100 μM), 2,4'-BPF (100 μM).

C.2 Chapter 3

Supplementary Table 3.1: Total number of replicates per condition included in the analysis.

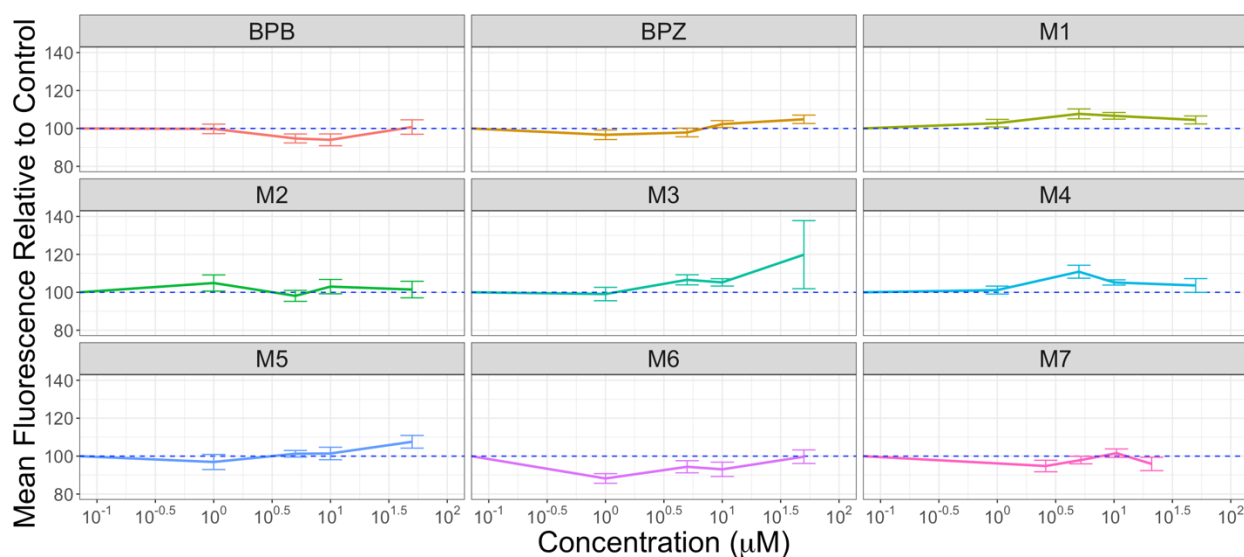
Chemical	Concentration (μM)													
	0.001	0.01	0.1	0.5	0.66	1	1.31	2.62	5	5.25	10	10.5	21	50
2,4'-BPF	3	3	4	4		4			3		4			4
2,4'-BPS	3	2	3	4		4			2		3			2
4,4'-BPF	3	3	3	3		3			3		2			2
BPA	4	4	4	4		3			3		4			3
BPAF	4	3	4	4		3			2		4			3
BPAP	4	4	4	4		4			4		4			4
BPB	3	4	3	4		3			4		3			4
BPC	4	4	3	4		4			4		4			3
BPS	4	4	4	4		4			4		3			4
BPZ	4	3	3	3		4			4		4			4
M1	3	4	4	4		3			4		4			4
M2	4	4	4	4		4			4		4			4
M3	4	4	4	4		4			4		4			3
M4	4	3	4	4		4			4		4			2
M5	4	4	4	4		4			3		3			3
M6	4	4	4	4		4			4		3			4
M7					4		4	3		3		4	4	
P201	3	3	3	4		3			3		3			3
TMBPF		3	2			3			4		3			3

Supplementary Table 3.2: Comparison of transcriptomic points of departure (tPODs) derived from whole transcriptome analysis of MCF-7 cells (n = 2 – 4 per concentration) exposed to seven mixtures of bisphenol A (BPA) alternatives and 12 individual chemicals at a range of concentrations (0.001 – 50 μ M) for 48 h.

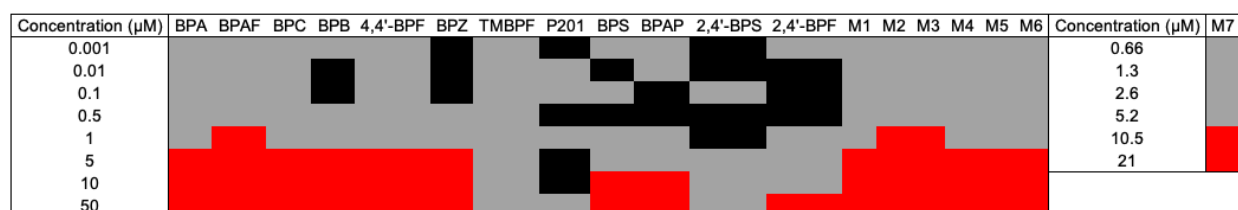
Chemical	Lowest Pathway 5th Percentile BMC	Lowest Pathway 5th Percentile BMCL - BMCU	Predicted Lowest Pathway 5th Percentile BMC	# Genes fitting ER α biomarker	ER α Median BMC	ER α Median BMCU - BMCL	Predicted ER α Median BMC	LCRD	LCRD BMCL - BMCU	Predicted LCRD
BPA	0.23	0.0052 - 0.49		21	2.73	2.03 - 4.15		0.18	0.091 - 0.32	
BPAF	0.12	0.014 - 0.23		20	0.58	0.19 - 1.29		0.11	0.055 - 0.17	
BPC	0.21	0.0043 - 0.48		23	2.90	1.55 - 5.39		0.21	0.11 - 0.34	
BPB	0.32	0.015 - 0.54		21	3.67	2.61 - 6.38		0.24	0.14 - 0.37	
4,4'-BPF	0.18	0.00089 - 0.91		20	3.48	2.40 - 6.16		0.41	0.12 - 0.77	
BPZ	0.51	0.0085 - 0.91		14	4.63	3.11 - 9.07		0.38	0.17 - 0.54	
TMBPF	0.07	0.00062 - 1.81		4				0.45	0.096 - 1.56	
P201	2.48	0.021 - 5.47		1				1.21	0.51 - 2.77	
BPS	0.98	0.040 - 3.48		12	22.27	15.7 - 38.0		0.93	0.49 - 2.90	
BPAP	0.99	0.024 - 2.59		2				0.96	0.50 - 2.69	
2,4'-BPS	1.02	0.014 - 3.18		0				0.93	0.49 - 2.23	
2,4'-BPF	1.31	0.020 - 3.28		3	36.8	22.2 - 106		1.64	0.53 - 2.76	

M1	0.18	0.0039 - 0.42	0.16	18	0.50	0.22 - 1.11	0.96	0.14	0.073 - 0.24	0.14
M2	0.20	0.017 - 0.48	0.17	20	0.81	0.24 - 2.24	1.24	0.17	0.092 - 0.28	0.15
M3	0.20	0.014 - 0.40	0.17	19	0.64	0.22 - 1.63	1.48	0.19	0.10 - 0.29	0.18
M4	0.28	0.011 - 0.56	0.21	23	1.94	1.16 - 4.69	1.84	0.22	0.12 - 0.36	0.22
M5	0.24	0.0065 - 0.52	0.24	24	3.50	2.43 - 6.01	2.18	0.23	0.12 - 0.38	0.25
M6	0.28	0.0021 - 0.66	0.33	17	4.09	2.82 - 7.32	3.23	0.33	0.16 - 0.51	0.34
M7	0.32	0.00029 - 1.05	0.32	17	3.95	2.56 - 5.59	5.73	0.37	0.15 - 0.69	0.58

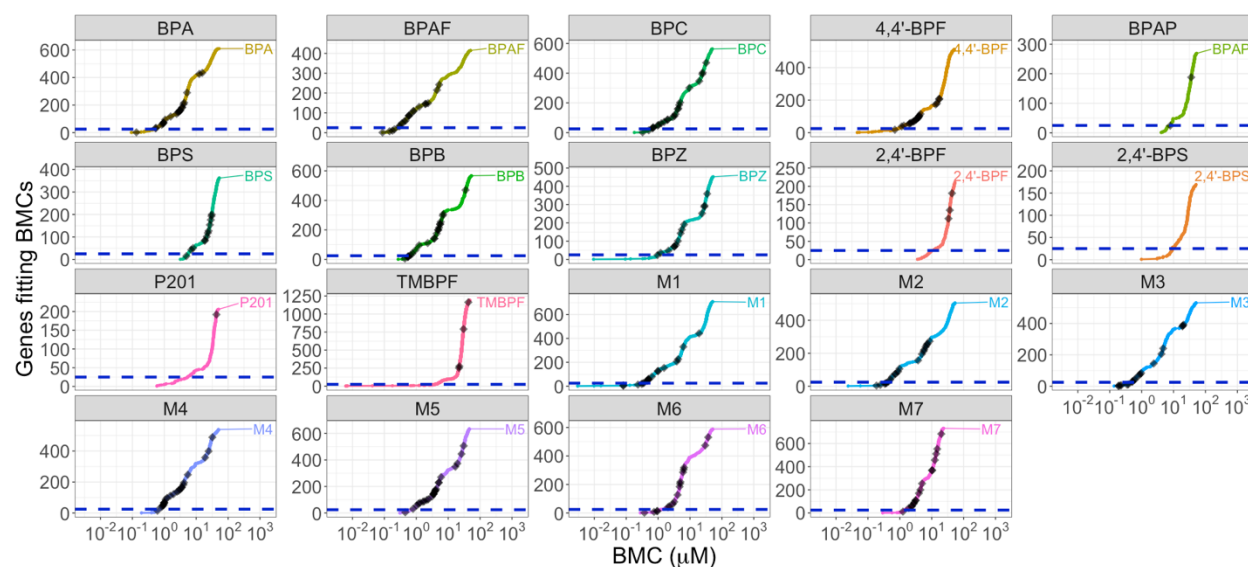
tPODs were derived by prefiltering data using the Williams trend test ($p < 0.05$) and an absolute fold-change filter of ≥ 1.5 . Data were also postfiltered with the following settings in BMDEpress v3: Best BMD/BMDL <20, Best BMDU/BMD <20, Best BMDU/BMDL <40, and Best fitPvalue > 0.1. tPODs (shown in μM) representing the 5th percentile gene BMC for the lowest pathway (at least 3 genes and 5% of pathway), the median gene benchmark concentration (BMC) for the estrogen receptor alpha (ER α) biomarker gene set, and the lowest consistent response dose are shown. The ER α tPOD is based on at least one concentration yielding an agonist or antagonist call and at least 3 genes producing BMCs. BPA is shown at the top and then most chemicals are shown in decreasing order of potency based on tPODs from the 25th gene BMC, followed by the mixtures.



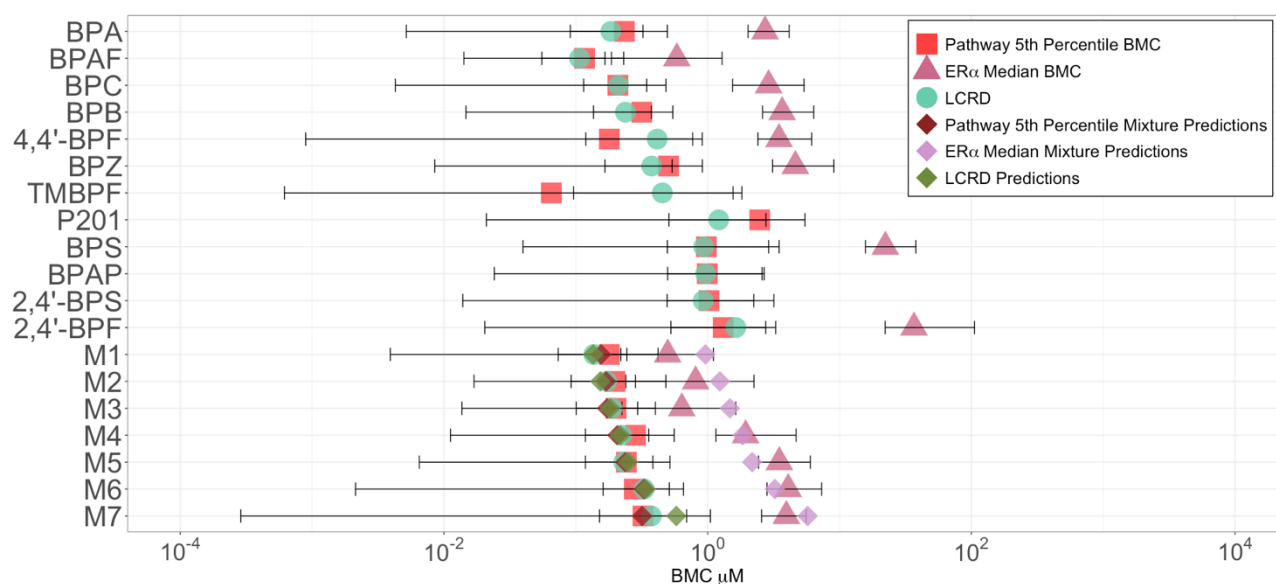
Supplementary Figure 3.1: Cell viability using a CellTiter-Blue Cell Viability Assay. MCF-7 cells ($n = 2 - 4$) were exposed to seven mixtures of bisphenol A (BPA) alternative chemicals and individual chemicals at a range of concentrations (0.001 – 50 μM) for 48 hours and are compared to their respective control (dimethyl sulfoxide; 0.1%) samples. First row of chemicals from left to right: BPB, BPZ, M1. Second row: M2, M3, M4. Third row: M5, M6, M7. Blue horizontal dotted line denotes 100% cell viability relative to controls. There were no significant declines in viability relative to controls.



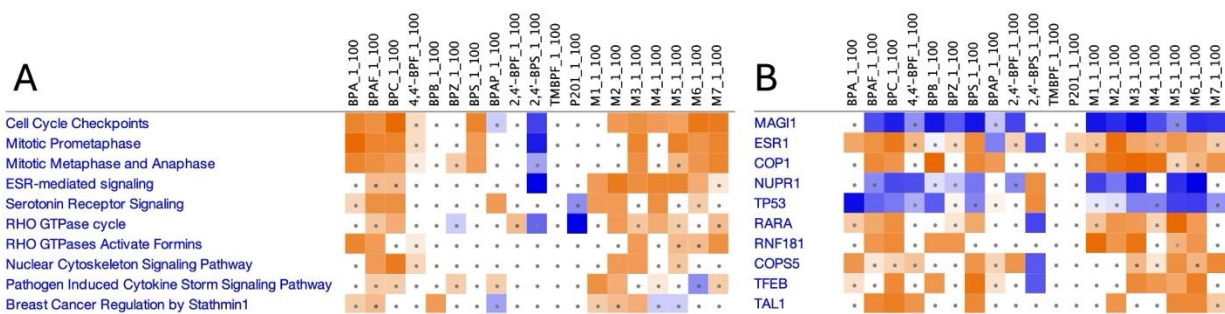
Supplementary Figure 3.2: Summary of estrogen receptor alpha transcriptomic biomarker activity of MCF-7 cells exposed to seven mixtures of bisphenol A (BPA) alternative chemicals and 12 individual chemicals at a range of concentrations (0.001 – 50 μM) for 48 hours. The biomarker was activated (red) or inhibited (black) based on thresholds for significance $-\log(\text{p-value}) \geq 4$ or ≤ -4 .



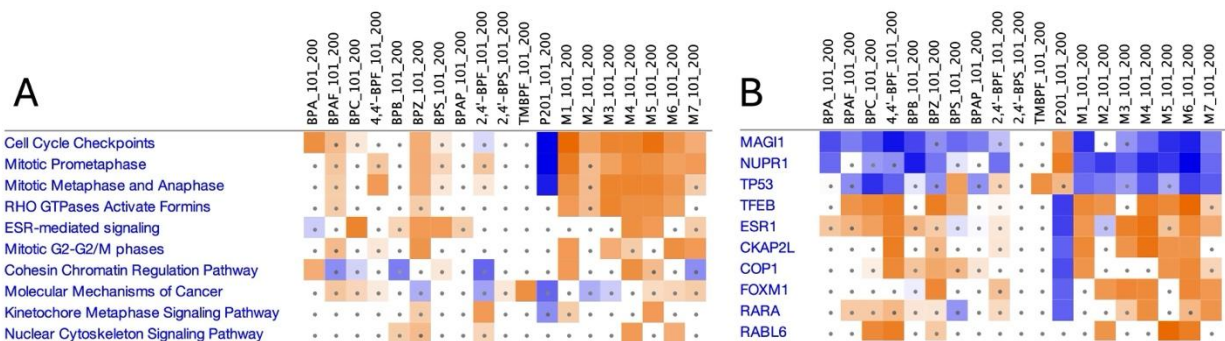
Supplementary Figure 3.3: Gene accumulation plot of MCF-7 cells ($n = 2 - 4$ per concentration) exposed to seven mixtures of bisphenol A (BPA) alternatives and 12 individual chemicals at a range of concentrations ($0.001 - 50 \mu\text{M}$) for 48 h. Genes fitting the 50 gene estrogen receptor alpha biomarker are overlaid (black diamonds). The blue dashed line indicates the 25th ranked gene. Data were prefiltered using the Williams trend test ($p < 0.05$) and an absolute fold-change filter of ≥ 1.5 , and post filtered with the following settings in BMDEExpress v3: Best BMD/BMDL < 20 , Best BMDU/BMD < 20 , Best BMDU/BMDL < 40 , and Best fitPvalue > 0.1 . First row of chemicals from left to right: BPA, BPAF, BPC, 4,4'-BPF, BPAP. Second row: BPS, BPB, BPZ, 2,4'-BPF, 2,4'-BPS. Third row: P201, TMBPF, M1, M2, M3. Fourth row: M4, M5, M6, M7.



Supplementary Figure 3.4: Comparison of transcriptomic points of departure (tPODs) derived from whole transcriptome analysis of MCF-7 cells ($n = 2 - 4$ per concentration) that were exposed to seven mixtures of bisphenol A (BPA) alternatives and 12 individual chemicals at a range of concentrations ($0.001 - 50 \mu\text{M}$) for 48 h. tPODs were derived by prefiltering data using the Williams trend test ($p < 0.05$) and an absolute fold-change filter of ≥ 1.5 . Data were also postfiltered with the following settings in BMDEExpress v3: Best BMD/BMDL < 20 , Best BMDU/BMD < 20 , Best BMDU/BMDL < 40 , and Best fitPvalue > 0.1 . tPODs (shown in μM) representing the lowest pathway 5th percentile benchmark concentration (BMC), median gene BMC for the estrogen receptor alpha biomarker gene set, and the lowest consistent response dose are shown. The ER α tPOD is based on at least one concentration yielding an agonist or antagonist call and at least 3 genes. BMCL and BMCU are used for the lower and upper bounds, respectively. The reference chemical BPA is shown at the top. Most chemicals are shown in decreasing order of potency based on tPODs from the 25th gene BMC, followed by the mixtures. Predictions based on additivity of individual components are shown as diamonds for the mixtures.

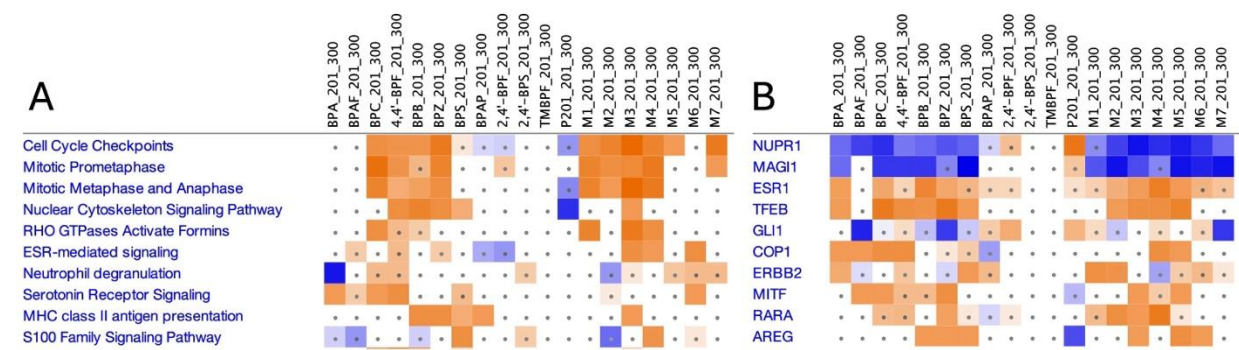


Supplementary Figure 3.5: Ingenuity Pathway Analysis of genes with rank order 1–100 (Bin 1) in MCF-7 cells ($n = 2-4$ per concentration). Cells were exposed to seven bisphenol A (BPA) mixtures and 12 individual chemicals (0.001–50 μM , 48 h). Genes were ranked from lowest to highest benchmark concentration. The list of genes (rank order 1–100, Williams trend test p-value, maximum fold-change) was imported into IPA. Orange indicates predicted activation, blue predicted inactivation (dots = Z-score < 2). Filters: Z-score ≥ 2.0 , $p < 0.05$. (A) Top 10 canonical pathways. (B) Top 10 upstream regulators.



Supplementary Figure 3.6: Ingenuity Pathway Analysis of genes with rank order 101–200 (Bin 2) in MCF-7 cells ($n = 2-4$ per concentration). Cells were exposed to seven bisphenol A (BPA) mixtures and 12 individual chemicals (0.001–50 μM , 48 h). Genes were ranked from lowest to highest benchmark concentration. The list of genes (rank order 101–200, Williams trend test p-value, maximum fold-change) was imported into IPA. Orange indicates predicted activation, blue predicted inactivation (dots = Z-score < 2). Filters: Z-score ≥ 2.0 , $p < 0.05$. (A) Top 10 canonical pathways. (B) Top 10 upstream

regulators.



Supplementary Figure 3.7: Ingenuity Pathway Analysis of genes with rank order 201–300 (Bin 3) in MCF-7 cells (n = 2–4 per concentration). Cells were exposed to seven bisphenol A (BPA) mixtures and 12 individual chemicals (0.001–50 μ M, 48 h). Genes were ranked from lowest to highest benchmark concentration. The list of genes (rank order 201–300, Williams trend test p-value, maximum fold-change) was imported into IPA. Orange indicates predicted activation, blue predicted inactivation (dots = Z-score < 2). Filters: Z-score $\geq$ 2.0, p < 0.05. (A) Top 10 canonical pathways. (B) Top 10 upstream regulators.

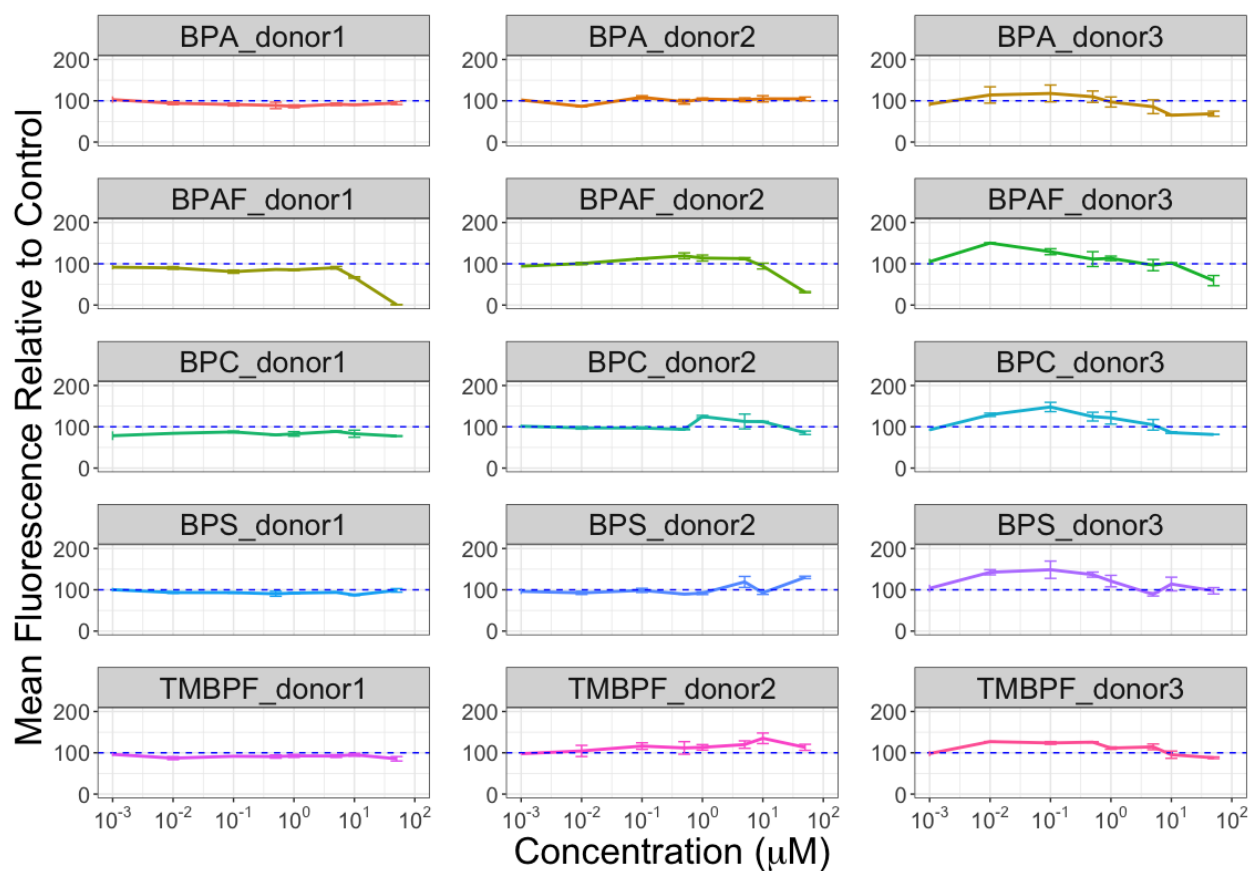
C.3 Chapter 4

Supplementary Table 4.1: Summary of biological donors

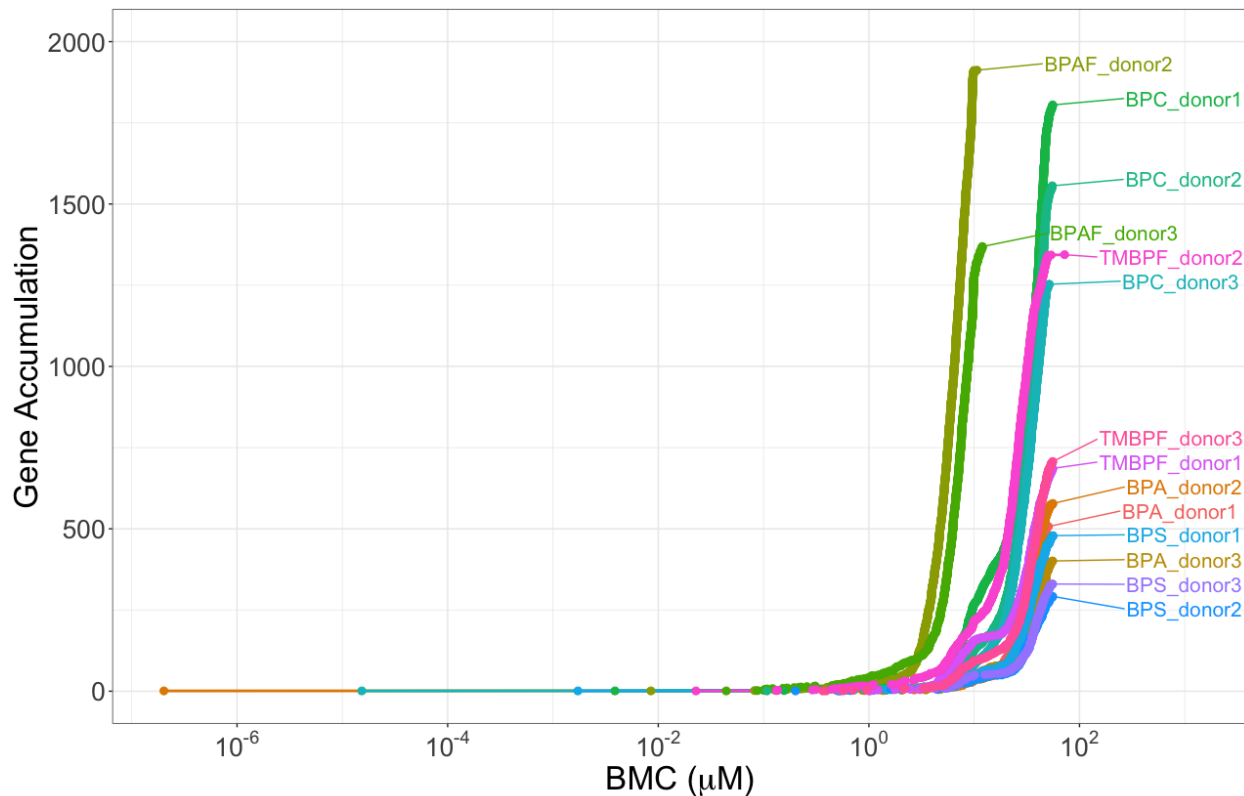
Donor	Height (in)	Weight (lb)	Age	Sex	Race	Smoking
1	65	155	29	F	Caucasian	N
2	65	155	29	F	Caucasian	N
3	65	145	24	F	Caucasian	N

Supplementary Table 4.2: Total number of replicates per condition included in the analysis

Chemical	Donor	0.001	0.01	0.1	0.5	1	5	10	50
BPAF	1	4	3	3	3	4	3		
	2	4	4	3	4	4	4		
	3	4	3	4	4	4	4	3	
TMBPF	1		4	4	4	4	3	3	3
	2		4	4	3	3	4	4	3
	3		4	4	4	3	4	4	4
BPC	1	4	4	4	4	4	4	2	4
	2	3	4	4	4	3	4	4	4
	3	3	4	4	4	4	3	4	3
BPS	1	4	3	4	3	3	4	4	3
	2	3	4	4	4	2	3	3	3
	3	2	3	3	3	4	3	4	4
BPA	1	4	4	4	4	4	4	3	2
	2	4	4	4	4	4	4	4	4
	3	4	4	4	4	3	4	3	4



Supplementary Figure 4.1: Cell viability using a CellTiter-Blue Cell Viability Assay. Human mammary epithelial cells ($n = 2 - 4$ per concentration) from three biological donors exposed to bisphenol A (BPA) and four alternatives at a range of concentrations ($0.001 - 50 \mu\text{M}$) for 48 h and are compared to their respective control (dimethyl sulfoxide; 0.1%) samples. First row of chemicals from left to right: BPA donor 1, donor 2, donor 3. Second row: BPAF donor 1, donor 2, donor 3. Third row: BPC donor 1, donor 2, donor 3. Fourth row: BPS donor 1, donor 2, donor 3. Fifth row: TMBPF donor 1, donor 2, donor 3. Blue horizontal dotted line denotes 100% cell viability relative to controls. BPAF significantly decreased cell viability at the highest concentration for all three donors. No other condition affected cell viability.



Supplementary Figure 4.2: Gene accumulation plot of human mammary epithelial cells ($n = 2 - 4$ per concentration) from three biological donors exposed to bisphenol A (BPA) and four alternatives at a range of concentrations ($0.001 - 50 \mu\text{M}$) for 48 h. Data were prefiltered using the Williams trend test ($p < 0.05$) and an absolute fold-change filter of ≥ 1.5 , and post filtered with the following settings in BMDExpress v3: Best BMD/BMDL < 20 , Best BMDU/BMDL < 40 , and Best fitPvalue ≥ 0.1 .