Genome-wide analysis of genetic alterations in prostate cells in response to xenoestrogens

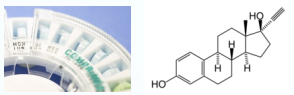
[musc-bioinformatics/]$

Matthew Huff[1,2], Sean M Courtney[1,3], E. Starr Hazard[1,4], Gary Hardiman[1,5]

[1]MUSC Bioinformatics, Center for Genomics Medicine, [2]MS in Biomedical Sciences Program, [3]Department of Pathology & Laboratory Medicine, [4]Library Science and Informatics, [5]Departments of Medicine & Public Health, Medical University of South Carolina (MUSC), Charleston SC

Background

- Many industrial chemicals are widely dispersed in the environment. In recent years these endocrine disruptors (ED) have been labeled 'contaminants of emerging concern' (CEC), as they have potential to cause adverse effects on wildlife and human health (1-5).
- Two important CECs are the synthetic estrogen 17α-ethynylestradiol (EE2) and bisphenol A (BPA) both of which are termed xenoestrogens (XEs) as they can bind the estrogen receptor (6) and disrupt estrogen physiology in mammals and other vertebrates.



17α-ethynylestradiol (EE2). EE2 is an orally bio-active estrogen used in almost all modern formulations of combined oral contraceptive pills. It is widely dispersed in municipal wastewater effluent discharges and active at very small amounts (pM levels) (3, 4)

Bisphenol A (BPA). Anthropogenic in origin, BPA serves as a chemical building block for the polycarbonate plastic and epoxy resins found in many consumer products. In 2006, 3.8 million tons of BPA were produced globally (5, 7-8).

- At present a fundamental gap of knowledge exists in understanding how exactly EE2 and BPA contribute to disease progression.
- In recent years the influence of XEs on oncogenes, specifically in relation to breast and prostate cancer has been the subject of considerable study. The long-term goal of our research is to better understand how physiological levels of key ED compounds act on the developmental mechanisms that integrate genetic and epigenetic interactions pertinent to prostate cancer.

Table 1. Amounts of BPA that have been reported from various global sites.

'Surface waters' refers to rivers and sources of drinking water. 'Non surface waters refer to landfill leachate, untreated and treated sewage influents and landfill plumes. The concentrations utilized in this proposal are 200 fold to 1000 fold less than the minimal concentrations detected in the aquatic environment.

sample	M amount	nM amount	µg/L
Surface Waters			
maximal	5.4E-02	5.4E+07	21
minimal	5.1E-06	5.1E+03	2.0E-03
Non Surface Waters			
maximal	4.4E+01	4.4E+10	1.7E+04
minimal	1.3E-03	1.3E+06	0.49
This Proposal			
BPA	5.1E-09	5	2.0E-06
BPA	2.5E-08	25	1.0E-05

- Bisphenol A was detected in 95% of the urine samples tested in the United States at concentrations ≥ 0.1mg/L (8). Urinary levels of BPA and its conjugates have been found in various populations in Southeast Asia (9-11).
- The largest BPA non-occupational exposure assessment reported urinary levels of 9.54 mg/L (8.91 mg/g creatinine) in a group of seventy-three adult Koreans (11).
- XEs act via multiple toxicity pathways to induce adverse health outcomes including the development of local/early disease in prostate cancer. The adverse outcome pathways (AOP) framework is a new strategy that organizes mechanistic and/or predictive relationships between initial chemical–biological interactions, pathways and networks, and adverse phenotypic outcomes.

Methodology

Cell culture

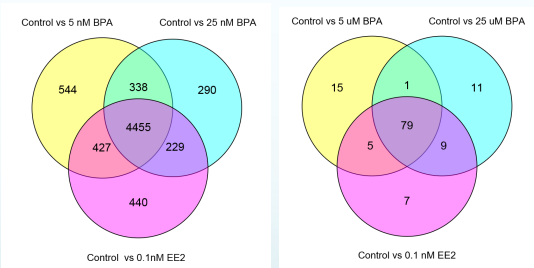
- Prostate Epithelial Cells (PrEC) obtained from Lonza, derived from a 23 year old male and were telomerase reverse transcriptase (TERT) expression negative. Cells were cultured in Clonetics™ PrEGM™ Prostate Epithelial Cell Growth Medium, in the absence or presence of environmentally relevant concentrations of BPA (low 5 and high 25 nM) and 17α-ethynylestradiol (EE2) (0.1 nM). RNA was extracted with TRIzol reagent (Invitrogen) and purified on a Qiagen RNeasy column. RNA integrity was verified using RNA 6000 Nano Assay chips run in Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA).

RNA sequencing (RNA-seq)

- High throughput sequencing (HTS) was performed using an Illumina HiSeq2000 with each sample sequenced to a minimum depth of ~50M reads.
- RNA sequencing data was analyzed using the OnRamp BioInformatics Genomics Research Platform RNAseq pipeline. OnRamp's GRP provides advanced genomics analysis with full data management to maintain analysis provenance across complex genomics analyses (RNAseq, DNAseq, etc).
- MUSC maintains an onsite 10-node, 240TB OnRamp Genomics Research Platform which utilizes hadoop software with automated data protection to seamlessly scale server and storage infrastructure.
- Gene Differential Expression (DE) analysis was carried out via Deseq2 and Tuxedo (Bowtie, Tophat, Cufflinks, HTseq) with OnRamp's RNAseq workflow.

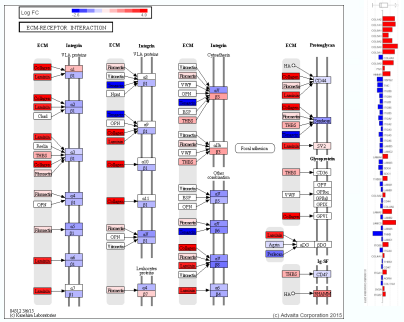
Results

- Transcript count data from DESeq2 analysis were sorted according to their adjusted p-value or q-value, which is the smallest false discovery rate (FDR) at which a transcript is called significant. FDR was calculated using the Benjamini-Hochberg multiple testing adjustment procedure.
- When we assessed differences in PrEC cells between control and EE2 exposed, this revealed that 5,806 transcripts were differentially expressed between the two groups. 5,357 DE transcripts were observed in 25 nM BPA exposed PrEC cells and 5,765 DE transcripts were observed in 5 nM BPA exposed PrEC cells ($q < 0.4$ for all comparisons).



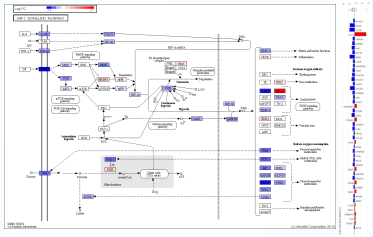
Venn diagram highlighting the overlap of differentially expressed transcripts by XE exposure in PrEC cells. Left panel: all significant genes $q < 0.4$ for all comparisons. Right panel: top 100 significant genes $q < 3.6^{E-19}$ for all comparisons.

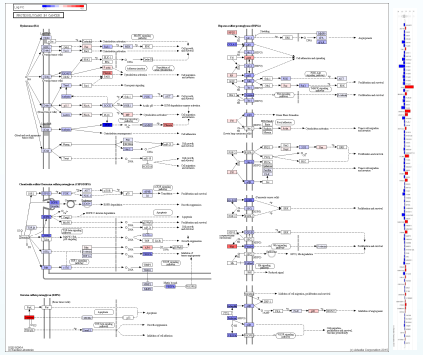
- Systems analysis was performed using Advaita iPathway Guide. A *Meta Analysis* was carried out where amongst the three experiments 5,664 transcripts were selected as differentially expressed ($q<0.4$) with an absolute log expression change (0.1). These data were analyzed in the context of the KEGG pathways and the Gene Ontology (GO) databases. In summary, 6 pathways were found to be significantly impacted,
  - ECM-receptor interaction,
  - Cytokine-cytokine receptor interaction,
  - Proteoglycans in cancer,
  - HIF-1 signaling pathway,
  - Regulation of actin cytoskeleton
  - Leukocyte transendothelial migration



Extracellular matrix (ECM) receptor interaction (KEGG: 04512): The ECM is a complex mixture of structural and functional macromolecules and serves an important role in tissue and organ morphogenesis and in the maintenance of cell and tissue structure and function. The pathway diagram is overlaid with the computed total perturbation of each gene. The highest negative perturbation is shown in dark blue, while the highest positive perturbation in dark red. **Gene measured expression bar plot**: Differentially expressed genes annotated to ECM receptor interaction are ranked based on the significance of their measured fold change. The plot displays top ranked DE genes. Upregulated genes are shown in red, downregulated genes are shown in blue.

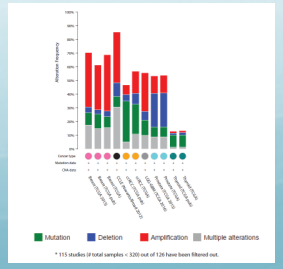
HIF1 signaling pathway (KEGG: 04066). Hypoxia inducible factor 1 (HIF1) is a transcription factor that functions as a master regulator of oxygen homeostasis. The pathway diagram is overlaid with the computed total perturbation of each gene. The highest negative perturbation is shown in dark blue, while the highest positive perturbation in dark red.
Gene measured expression bar plot: All the differentially expressed genes that are annotated to HIF1 signaling pathway are ranked based on the significance of their measured fold change. The plot is limited to the top ranked differentially expressed genes. Upregulated genes are shown in red, downregulated genes are shown in blue. The box plot on the left summarizes the distribution of all the differentially expressed genes that are annotated to this pathway. The box represents the 1st quantile, the median and the 3rd quantile, while the outliers are represented by circles.





Proteoglycans in cancer (KEGG: 05205): Many proteoglycans (PGs) in the tumor microenvironment have been shown to be key macromolecules that contribute to biology of various types of cancer including proliferation, adhesion, angiogenesis and metastasis, affecting tumor progress. The pathway diagram is overlaid with the computed total perturbation of each gene. The highest negative perturbation is shown in dark blue, while the highest positive perturbation in dark red. **Gene measured expression bar plot:** The plot displays the top ranked differentially expressed genes. Upregulated genes are shown in red, downregulated genes are shown in blue.

Cross-cancer alteration summary for the 79 shared transcripts. The cBioPortal for Cancer Genomics (12) was used to explore the 79 transcripts, shared across all three XE exposures $q < 3.6^{E-19}$ in the context of 'The Cancer Genome Atlas' (TCGA). The Y-Axis value is the alteration frequency for these 79 genes. A minimum number of 320 samples was required to add stringency to the analysis. The number of alterations, amplifications, deletions and mutations respectively are plotted. Alterations in these 79 transcripts have been implicated in Breast, Prostate, Thyroid and renal Cell Clear carcinomas, and Low-Grade Gliomas and Glioblastoma multiforme.



Conclusions

Low levels of BPA and EE2, comparable to those routinely detected in the environment, critically affect global transcriptomic responses in prostate epithelial cells. Systems level analyses indicated enrichment of biological pathways pertinent to cancer. Analysis of the highly significant 79 gene signature common to all exposures in TCGA data sets revealed alterations of these 79 transcripts in many cancers including prostate cancer.

Bibliography

Grun F *et al.* Endocrine-disrupting organotin compounds are potent inducers of adipogenesis in vertebrates. *Mol Endocrinol.* 2006; 20:2141-55.
Lam SH *et al.* Zebrafish whole-adult-organism chemogenomics for large-scale predictive and discovery chemical biology. *PLoS Genet.* 2006; 4:e1000121.
Vajda AM *et al.* Reproductive disruption in fish downstream from an estrogenic wastewater effluent. *Environ Sci Technol.* 2008; 42:3407-14.
Lange A *et al.* Sexual reprogramming and estrogenic sensitization in wild fish exposed to ethinylestradiol. *Environ Sci Technol.* 2009; 43:1219-25.
Vandenberg *et al.* Bisphenol-A and the great divide: a review of controversies in the field of endocrine disruption. *Endocr Rev.* 2009; 30:75-95.
Kuiper *et al.* Comparison of the ligand binding specificity and transcript tissue distribution of estrogen receptors alpha and beta. *Endocrinology.* 1997; 138:863-870.
Staples *et al.* Bisphenol A concentrations in receiving waters near US manufacturing and processing facilities, *Chemosphere* 2000; 405:521-525.
Calafat, Antonia M., et al. Urinary concentrations of bisphenol A and 4-nonylphenol in a human reference population. *Environmental health perspectives* 2005: 391-395.
Matsumoto, Akiko, *et al.* Bisphenol A levels in human urine. *Environmental Health Perspectives* 2003; 111.1:101.
Ouchi and Watanabe. Measurement of bisphenol A in human urine using liquid chromatography with multi-channel coulometric electrochemical detection. *Journal of Chromatography* 2002; *B* 780.2: 365-370.
Yang, Mihi, *et al.* Biological monitoring of bisphenol A in a Korean population. *Archives of environmental contamination and toxicology* 2003; 44.4: 0546-0551.
Gao *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* 2013; 6:269