Title: *De novo* transcriptomic analysis predicts the effects of phenolic compounds in Ba River on the liver of female sharpbelly (*Hemiculter lucidus*)

Members:

Jiahua Guo[1], Jiezhang Mo[2], Jianglin Peng[1], Shan Liu[1], Mirella Kanerva[3], Hisato Iwata[3],

Affiliations

1. Shaanxi Key Laboratory of Earth Surface System and Environmental Carrying Capacity, College of Urban and Environmental Sciences, Northwest University, Xi'an 710127, China.

2. Department of Chemistry, City University of Hong Kong, Kowloon, Hong Kong, China

3. Center for Marine Environmental Studies, Ehime University, Bunkyo-cho 2-5, Matsuyama, Ehime prefecture, Japan. 790-8577.

1. Background

Ba River as one of the biggest urban river of Xi'an has been overburdened with the effluent discharged from 5th wastewater treatment plant (WWTP) in Xi'an, which mainly receives and treats the municipal and industrial sewage. Wang et al. (2018) have measured the concentrations and distribution of BPA, NP, and 4-t-OP in three sampling sites, including a mid-stream site (M) located near the outfall of WWTP, and two sites (up- (U) and down-stream (D)) situated approximately four kilometers away from the WWTP discharges. In regards to sites U, M, and D, the mean concentrations for BPA were 297, 1573.1, and 334.9 ng L-1, respectively; the exposure levels for NP were 125.2, 634.8, and 185.6 ng L-1; and concentrations for for 4-t-OP were 85.4, 126, and 56 ng L-1, respectively. In the U, M, and D sites, the mean concentrations in sharpbelly were 16.7, 26.4, 11.5 (ng g-1, dry muscle weight) for 4-t-OP, 15.5, 103.5, 30.8 ng g-1 for NP, and 24, 146.9, 28.7 ng g-1 for BPA, respectively.

2. Aims

The objectives of the present study were: (1) to assess the potential phenotypic changes induced by phenolic compounds containing WWTP effluent; (2) to decipher the molecular pathways dysregulated by phenolic compounds in WWTP effluent; and (3) to link the potential contaminants in the effluent with the dysregulated molecular pathways, predicting the potential adverse effects.

3. Procedure

3.1 Sampling sites and fish collection

Sharpbelly were collected from the Ba River that includes an outfall of the 5th municipal wastewater treatment plant (WWTP; Fig. 1). Three sampling sites were selected including a site located up-stream (U) from the WWTP outfal, a mid-stream site (M) located near the outfall of WWTP, and a site located down-stream (D) from the WWTP discharges. Here, sharpbelly collected from up-, mid- and down-stream of Ba River that were designed to be interpreted as control, high, and low treatment groups, respectively. While fifteen fish were

caught from each site by fishing, more than 90% were females. Thus, four adult female sharpbelly in each site were randomly selected for further transcriptomic analysis.

After capture, sharpbelly were immediately sacrificed following the animal handing procedures of Northwest University. Four biometric parameters including body weight (g), Fulton's condition factor (CF = fish weight/ fish length), gonado-somatic index (GSI = gonad weight/ body weight), and hepatosomatic index (HSI = liver weight/ body weigh) were employed to evaluate the phenotypic effects of WWTP effluent on all the captured sharpbelly. Liver tissues submerged in Trizol (Shanghai Yuanye Bio-Technology Co., Ltd, China) were subjected to snap freezing by liquid nitrogen and stored at -80 0C prior to RNA extraction.

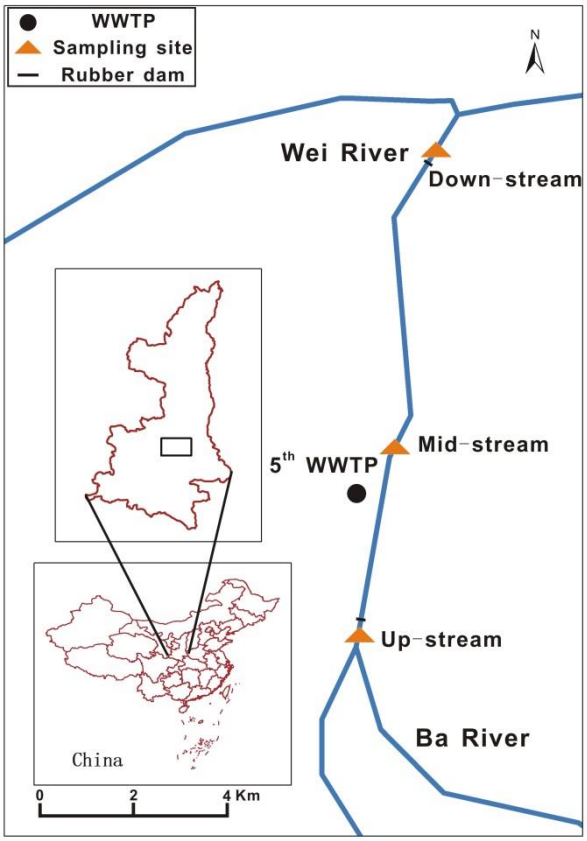

Fig. 1. Map of the study area and the sampled site location. The study design incorporated three sampling sites, including a site located up-stream from the WWTP outfall, a mid-stream site located near the outfall of WWTP, and a site located down-stream from the WWTP discharges.

3.2 RNA extraction, library preparation, and illumina sequencing

3.3 *De novo* assembly and functional annotation

3.4 Differentially expressed genes (DEGs) enrichment analysis

Methods in 3.2-3.4 followed the sophisticated bioinformatics steps so they are not repeated.

4. Results

4.1 Phenotypic assessment

Internal examination of these sharpbelly revealed that CF and body weight in the mid-stream fish were remarkably increased by approximate 20% in comparison with the up-stream (Fig. 2). Liver weights characterized by HSI in the sharpbelly dwelling in the mid- and down-stream were higher than that in the up-stream, but no significant difference was recorded

between sites (Fig. 2).

4.2 Transcriptomic analysis

4.2.1 Illumina sequencing and *de novo* assembly

RNA-seq data were obtained from 41.7 million to 52.9 million reads per tissue sample. The majority of reads were of high quality with Q20 values more than 96.52% across all the samples. After filtering the sequencing data containing a number of adaptors and low quality reads (< Q20), the clean data occupying > 92.38% were assembled *de novo* using the Trinity assembler. *De novo* assembly produced 261,296 unigenes with the mean length of 618.8 bp and a total length of 161.7 Mb.

4.2.2 Functional annotation

The unigenes were annotated using BLAST searches against NR, GO, KEGG, eggNOG, and Swissprot databases. The annotated RNA-seq genes were distributed in NR (25.18%), GO (7.09%), KEGG (12.66%), and Swissprot (17.41%). In the annotation to NR, sequencing data of sharpbelly showed high similarity to *Sinocyclocheilus rhinocerous* (18.85%), *Sinocyclocheilus anshuiensis* (15.86%), and *Sinocyclocheilus graham* (14.51%; Fig. 3A). Based on GO, 18538 unigenes were categorized into 3 main categories, including biological process (BP), molecular function (MF), and cellular components (CC), and 67 subcategories. From KEGG, 33,069 unigenes were assigned into 35 pathways that were primarily involved in the process of metabolism, genetic information processing, environmental information processing, cellular process, and organismal systems.

4.2.3 Differentially expressed genes and enrichment analyses

In total, we identified 160 (83 up- and 77 down-regulated), 15 (5 up- and 10 down-regulated), 162 (97 up- and 65 down-regulated) DEGs in U-M, U-D, and D-M groups of female sharpbelly, respectively (Fig. 3C, S3). In heatmap, DEG expression of sharpbelly living near WWTP outfall were clearly departed from three individuals collected from up- and down-stream sites, whereas the rest samples were indistinguishable from the mid group (Fig. 3B).

GO analysis was implemented to investigate the roles of DEGs in the cellular function. Given a set of DEGs, the GO analysis is able to figure out the CC, MF, and BP terms that were influenced by WWTP effluent in the female sharpbelly. Top 10 GO terms enriched in CC, MF, and BP in each group were summarized in Supplementary data. Contrast with U-M, and D-M groups, we merely identified 3 CC and 4 MF terms in U-D group, suggesting there was no significant difference in sharpbelly living in both sites. The functional KEGG pathway analysis determined 32 and 50 enriched pathways in U-M and D-M groups that may have been affected by WWTP effluent. Top 20 pathways enriched in each sites were illustrated in Tables 1 & 2. These pathways were mainly in relation to drug addiction and metabolism (e.g., amphetamine addiction, cocaine addiction, and drug metabolism – cytochrome P450 for citalopram), endocrine system (e.g., endocrine resistance, endocrine and other factor−regulated calcium reabsorption, thyroid hormone synthesis and signaling), and cellular process and lipid metabolism (e.g., mitogen-activated protein kinase (MAPK), adipocytokine signaling, hepatitis C, and the peroxisome proliferator-activated receptors (PPARs) pathway). Nearly no pathway was enriched in U-D group.

4.3 Predicted effects of phenolic compounds on sharpbelly

To sum up, our results provided an insight into the molecular mechanism of action of WWTP effluent incorporating phenolic compounds and predicted the potential adverse outcomes in female sharpbelly (Fig. 4).

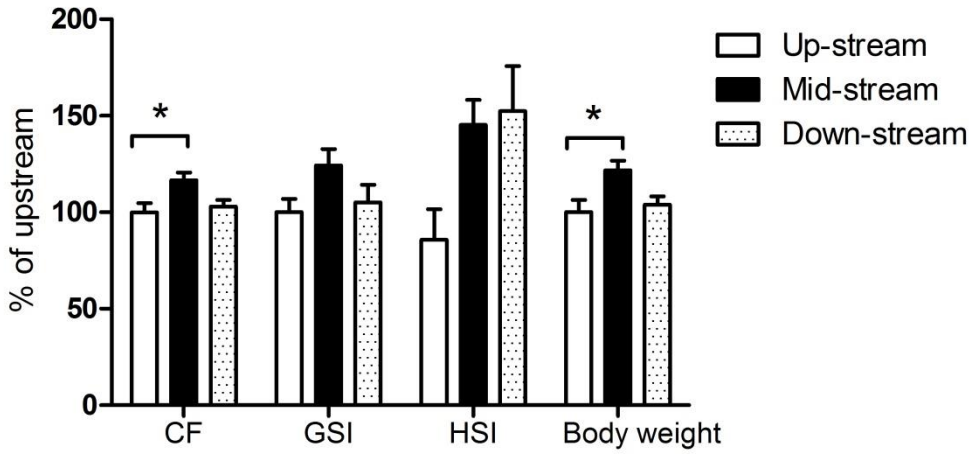

Fig. 2. Condition factor (CF), gonado-somatic index (GSI), hepatosomatic index (HSI) and body weight (BW) expressed as a percentage of values in sharpbelly collected from the mid- and down-streams relative to those in the up-stream. Data represent mean ± standard deviation (n = 15).

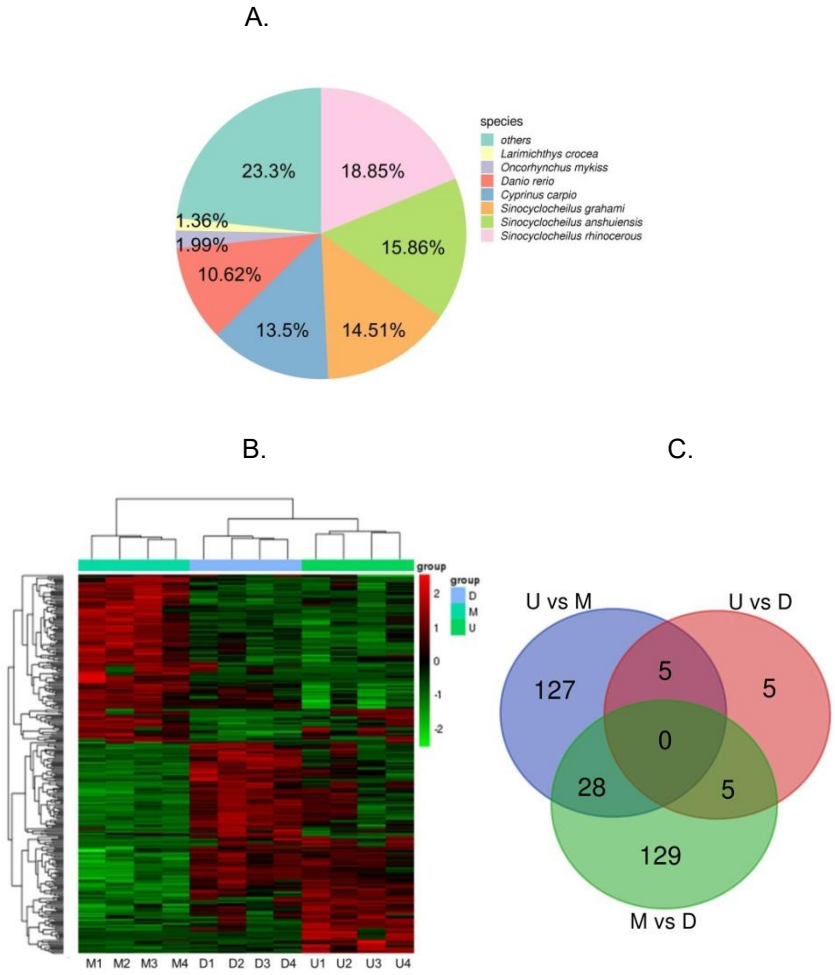

Fig. 3. Transcriptomic profiling of sharpbelly living down-stream from WWTP outfall. (A) Similarity of sharpbelly sequencing data to other fish species; (B) A heatmap of centered and scaled FPKM values of DEGs in sharpbelly from different sites; (C) Venn diagram of the number of DEGs in each group.

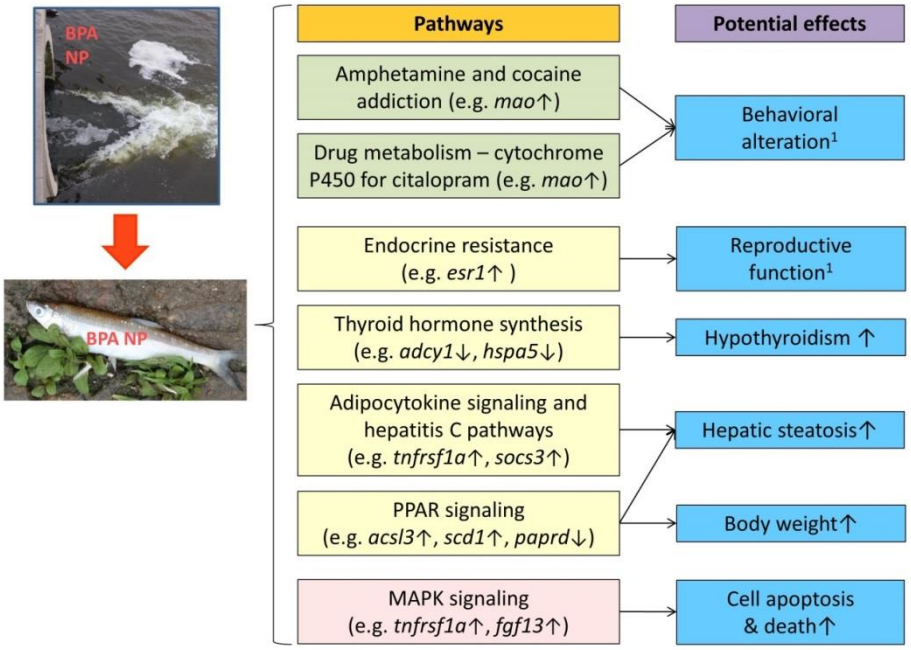

Fig. 4. Schematic diagram of the pathways activated in the female sharpbelly living in the mid-stream. *acsl3*: Long-chain-fatty-acid--CoA ligase 3; *adcy1*: adenylate cyclase 1; *esr1*: estrogen receptor 1; *fgf13*: Fibroblast growth factor 13; *hspa5*: heat shock protein family A (Hsp70) member 5; *mao*: monoamine oxidase; *ppard*: Peroxisome proliferator-activated receptor delta; *scd1*: stearoyl-CoA desaturase; *socs3*: suppressor of cytokine signaling 3; *tnfrsf1a*: tumor necrosis factor receptor superfamily member 1A. 1. Effects of WWTP effluent on the behavioral alteration and reproductive function were unclear.

Publication

This paper has been submitted to *Environmental pollution* and currently under peer-review.

Perspective In Future

As BPA and NP are typical endocrine disrupting chemicals, a further investigation into molecular changes in other organs (e.g. gonad) is required for a complete understanding of the adverse effects caused by phenolic compounds in the WWTP effluent.