Supplemental Materials for

Single-cell RNA sequencing reveals size-dependent effects of polystyrene microplastics on immune and secretory cell populations from zebrafish intestines

Weiqing Gu1#, Su Liu1,2#, Ling Chen1, Yuxuan Liu3, Cheng Gu1, Hong-qiang Ren1, Bing Wu1*

1 State Key Laboratory of Pollution Control and Resource Reuse, School of the Environment, Nanjing University, Nanjing, 210023, P.R. China

2 Department of Environmental Science, School of Engineering, China Pharmaceutical University, Nanjing, 211198, P.R. China

3 College of Environment, Hohai University, Nanjing 210098, P.R. China

# These authors contributed equally: Weiqing Gu, Su Liu.

* Corresponding author:
Email: bwu@nju.edu.cn
Tel & Fax: +86-25-89680720
Address: 163 Xianlin Avenue, Nanjing, 210023, P.R. China
Cytokine analysis method

The 2% intestinal tissue homogenate was prepared using a tissue homogenizer (Tissuelyser 48, Shanghai). The homogenate was centrifuged at 15000 rpm at 4 °C for 15 min to obtain the supernatant, which were diluted 10 times for the physiological tests. The intestinal levels of TNF-α and TLR2 were measured following the protocols of their fish-specific ELISA kits (Nanjing Jiancheng Bioengineering Research Institute, CHN). Standard curves of these detection kits were obtained to determine the amount of each intestine sample, with a satisfactory linearity of $R^2 > 0.99$. All samples and standards were operated in triplicate.

Single-cell suspension preparation

Intestinal segments were opened longitudinally and sliced into small fragments. The segments were transferred to 1 mL 0.04% BSA–PBS and gently shaken. After incubated in 0.5 mL of modified HBSS containing 0.3 U/mL of dispase (Yuanye, China) at 37 °C for 4 min with intermittent shaking, 50 μL of fetal bovine serum was added to inactivate the dispase. The cell suspension was then sequentially passed through 40 μm filter and pelleted 2.5 min at 300 g. After washed by 0.04% BSA–PBS, the cells were finally suspended in IESC media for the scRNA-seq\(^1\). The number of viable cells in the suspension cells was determined by an automatic cell counter (TC20, Bio-Rad, USA) to ensure that there are at least $5 \times 10^5$ cells in each sample and that the viability is above 80%.

Analysis of ScRNA-seq data

In Cell Ranger software, cellranger mkfastq was used to split data and check quality. Sample demultiplexing was performed based on the 8 bp sample index read to
generate FASTQs for the Read1 and Read2 paired-end reads, as well as the 16 bp GemCode barcode. The resulting FASTQ data were aligned to the reference genome using the comparison software STAR. Cellranger count was run on FASTQ data from each of the GEM wells individually. GemCode barcodes and UMIs were filtered. Cellranger aggr was used to aggregate outputs from multiple runs mentioned above of Cellranger count, normalizing those runs to the same sequencing depth and then recomputing the feature-barcode matrices and analysis on the combined data. Cellranger reanalyze pipeline was used to take feature-barcode matrices produced by cellranger aggr and rerun the dimensionality reduction, clustering, and gene expression algorithms using tunable parameter settings. PCA, t-distributed stochastic neighbor embedding (t-SNE) and $k$-means clustering were performed on the pooled matrix. GO biological processes, Reactome pathways, as well as KEGG pathways were enriched with the criteria that p-value < 0.01, a minimum count of 3, and an enrichment factor > 1.5. Cytoscape (version 3.7.1) plug-in ClueGO was used to decipher functionally grouped GO and pathway annotation networks. In this study we used ClueGO combined with CluePedia, another Cytoscape plug-in, to build networks of pathways based on the GO and KEGG databases, showing only pathways with p < 0.05. The online database Search Tool for the Retrieval of Interacting Genes (STRING) (https://string-db.org/) was used to construct the functional protein association networks. Cytoscape was utilized to visualize DEGs and pathways interaction networks.

**High-throughput sequencing of intestinal microbiota**

Total DNA in the intestinal contents was isolated with DNeasy Blood & Tissue Kit (QIAGEN, Germany). DNA concentration was measured using a NanoDrop spectrophotometer (NanoDrop One, Thermo, USA). After DNA quality analysis by
electrophoresis, the extracted DNA was amplified using primers (Forward primer: 5’-CCTAYGGGRBGCASCAG-3’; Reverse primer: 5’-GGACTACNNGGGTATCTAAT-3’) to target the V3-V4 regions of bacterial 16S rRNA gene. The obtained PCR products were purified by GeneJET Gel Extraction Kit (Thermofisher, USA) and applied to construct the sequencing library. The library was sequenced on an Ion S5™ XL platform.

Sequencing results were cut with low-quality fractions using Cutadapt (version 1.9.1), and the chimeric sequences were removed to obtain clean reads. Uparse (version 7.0.1001) was used for operational taxonomic units (OTUs) clustering at the 97% identity level. For each representative sequence, the Silva Database (https://www.arb-silva.de/) was used based on Mothur algorithm to annotate taxonomic information. Alpha diversity was determined by the Shannon index calculated in QIIME. The principal coordinates analysis (PCoA) was performed to get principal coordinates from complex, multidimensional data.
Figure S1. Single-cell RNA sequencing of zebrafish intestine. (A) Median number of genes and transcripts (UMI counts) detected per cell in a data mixture of 12000 cells. (B) t-SNE projection where cells that share similar transcriptome profiles are grouped by colors representing unsupervised clustering results. (C) Heat map analysis using known cell-specific marker gene expression profiles of cluster 1 (lymphocytes), cluster 2 (phagocytes), cluster 4 (enterocytes) and cluster 5 (secretory cells). (D) Violin plots depicting expression of genes Fabp10a and Spag6.
Figure S2. Feature plots on t-SNE showing expression of the marker genes of each cell type across these clusters. These plots show cells colored by the expression (log2(Exp + 1)) of markers of lymphocytes (Il7r and Lck), phagocytes (Cxc3r.2 and Ncf4), enterocytes (Fabp2), secretory cells (Anxa4), enteroendocrines (Neurod1), goblet cells (Agr2), T cells (Zap70), B cells (Blnk), neutrophils (Mpx), macrophages (Mpeg1.1), classically activated (M1) macrophages (Tnfa) or alternatively activated (M2) macrophages (Marco).
Figure S3. Venn diagram of the number of DEGs based on (A) all cell population-averaged measurement of each exposure group and (B) four intestinal cell populations.

Figure S4. Networks on altered GO biological processes and KEGG pathways obtained from ClueGO (Cytoscape plug-in) in enterocytes and secretory cells. The pattern in the dashed box indicates which of the displayed pathways are specific for a treated group. Blue and red mean the specific pathways for the 100 nm and 5 μm PS-MP groups, respectively.
Figure S5. Venn diagram of the number of DEGs in five kinds of immune cells

Figure S6. Networks on altered GO biological processes and KEGG pathways obtained from ClueGO in immune cells. The pattern in the dashed box indicates which of the displayed pathways are specific for a treated group. Blue, red and yellow mean specific pathways for the 100 nm, 5 μm, and 200 μm PS-MP groups, respectively.
Figure S7. Distinct network of DEGs in immune cells for 100 nm (A), 5 μm (B) and 200 μm (C) PS-MP groups. The M1 macrophages, M2 macrophages, T cells, and B cells were labeled in green, yellow, purple and pink, respectively.
Figure S8. Networks on cellular chemotaxis-related pathways obtained from ClueGO in immune cells. Colors in the annulus indicate the different treatment groups. Blue, red and yellow mean 100 nm, 5 μm, and 200 μm PS-MP groups, respectively.
References
