The effects of 17α-ethinylestradiol on hepatic transcriptome of early-life stage and adult Japanese quail

Yeon Seon Jeon, Jin Kweon, Qian Li, Janice Ou

Introduction

17α-ethinylestradiol (EE2) is an endocrine disruptor widely found in the environment. It is toxic to many organisms such as fish[1-3] and birds[4]. Specifically, in Japanese quail (JQ), embryonic exposure to excessive EE2 causes reproductive organ abnormalities. It is suspected that EE2 exerts these cellular effects by activating the nuclear estrogen receptors (ERs) type alpha. Subsequently, the activated ER acts as a transcription factor to regulate targeted genes in the quails[4].

JQ has been frequently used as a research model in toxicity studies because of its relatively short sexual maturation period, minimum genetic variability and high sensitivity to environmental chemicals[5]. To be specific, when it comes to investigating the impacts of environmental chemicals on wildlife, domestic JQ is assumed to be more appropriate as a bridge to wild avian species than other widely used lab model organisms such as mice. However, there is only one standardized toxicity test protocol for adult and none for early life stage (ELS) JQ. In the case of other species, researchers have used ELS under the premise that it is the most susceptible stage to toxicants[6]. Currently, there is little research on the gene expression alterations on JQ, either ELS or adult, upon the exposure to EE2. Moreover, many avian RNA sequencing data has not been produced and analyzed in perspective of ecotoxicology[7] though it has huge potentials to be applied for predictive toxicology[8]. Therefore, by analyzing the first avian RNA sequencing data, our study would allow us to tackle questions about utilizing gene expression data as endpoints and substituting ELS for Adult in toxicity testing. The answers would have a meaning as a cornerstone for developing an alternative toxicity testing which is faster, cheaper, and more ethical.

Objectives & Hypothesis

We aim to determine the transcriptomic responses of ELS and adult JQ. We hypothesize that ELS would be more susceptible to EE2 than adult. If so, ELS would have more differentially expressed genes (DEGs), higher degree of dysregulation, and more dysregulated pathways than adult. According to the previous studies[4,9-10], we expect that genes such as apoVLDL, and

the ones responsible for sex organ development and sexual behaviors will be detected as DEGs in both adult and ELS JQs.

Materials & Methods

1. Data collection

In this study, we take advantages of RNA sequencing data produced via EcoToxChip project[11]. Five individuals per group (high, medium and control) were used to isolate RNAs from left-lobe livers. ELS and adult JQ cDNAs were sequenced by Illumina HiSeq 4000 PE100 and Illumina NovaSeq 6000 S4 PE100, respectively. The abundance of transcripts was quantified on Galaxy using Kallisto workflow. Details of exposure experiments for embryonic and adult JQs are described below.

ELS: EE2 was dissolved in dimethyl sulfoxide (DMSO) to be 50 mg/mL and the dosing solution was injected into the air cell of JQ eggs at two concentrations (High: 33.3 µg/g/egg; Medium: 3.33 µg/g egg). The control group was similarly injected with dimethyl sulfoxide (DMSO). Eggs were incubated for 9 days: the half of time necessary for hatching.

Adult stage: Following US EPA guideline OCSPP 850.2100, Six- to ten-week old adult JQs were dosed (High: 5 mg/kg; Medium: 0.5 mg/kg) and sampled 4 days after.

2. Data Analysis

To answer the research questions, DEGs were performed using EdgeR. Volcano plots and heatmaps with both transcript counts and fold changes were used to identify DEGs. DEGs were determined with adjusted *p*-value (or false discovery rate, FDR) of 0.05 and logarithmic fold change ($\log_2$FC) of 1 using EdgeR[12], which was recommended for a study with less than 12 replicates[13]. Enrichment analysis was used to find the associative pathways for the representative DEGs. For unsupervised machine learnings, both (sparse Independent) Principal Component Analysis ((sI)PCA) and hierarchical clustering algorithms (with different linkages) were applied to further investigate EE2 effects on both ELS' and adults' datasets.

Results

After obtaining the ELS and adult JQ countables, the duplicated rows were removed by summing up the transcript counts with the same row names. Both ELS and adult JQ datasets were then filtered by low count and low variance methods. Low variance filtration were perform with a cutoff of 20%, and low count filtration was conducted based on the low count of mean abundance of 10 transcripts. After filtration process, the remained gene numbers for ELS and adult were 14408 and 13472, respectively. Next, data normalization was performed with trimmed

mean of M-values (TMM). The normalization result of both datasets was visualized by boxplots and qqplots (Fig 1 a and b). We confirmed that the normalization performed well, and, in addition, it removed any significant outlier.

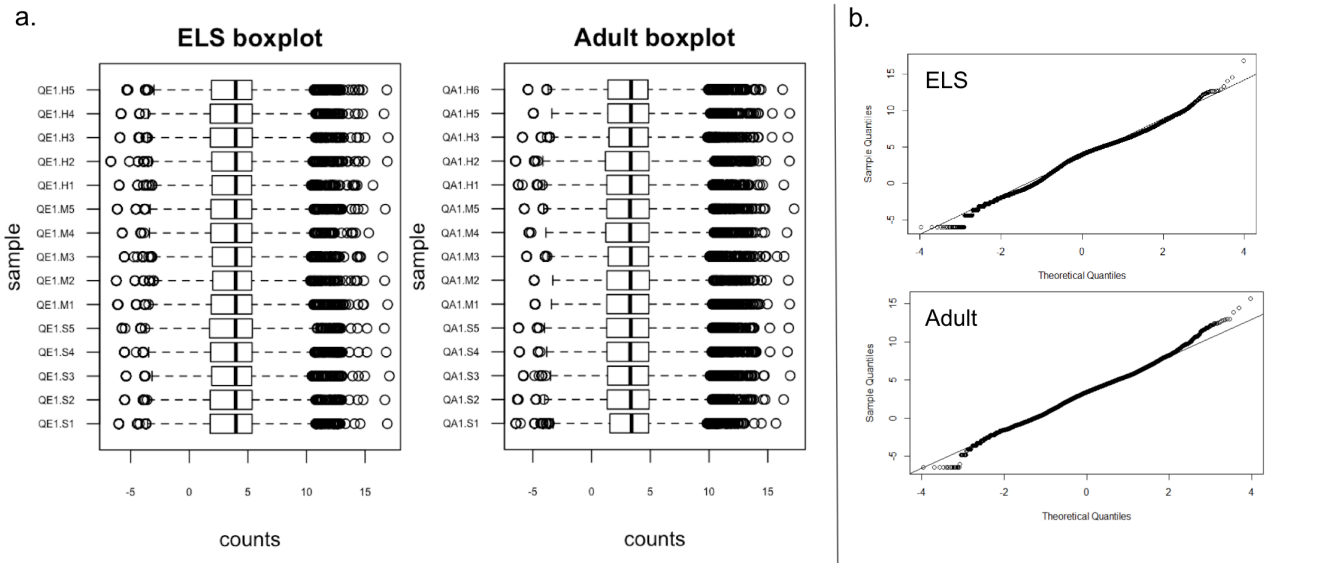

Figure 1. a. Boxplots of the normalized datasets across 15 samples in ELS and adult JQ. **b.** The QQ-plots show result from one of the 15 samples in each life stage. Similar QQ-plot patterns was shown in the other 14 samples in each life stage.

1. ELS JQ transcriptomic response upon exposure to EE2 treatment

Differential expression analysis was performed on the ELS data by comparing both medium and high doses of EE2 exposure with control. 246 genes were identified as DEGs. We plotted a heatmap with the normalized transcript counts across 15 samples categorized into control, EE2 medium and high groups (Fig 2a). Moreover, two volcano plots were drawn based on the $\log_2$FC of ± 1 and -$\log_{10}$P-values (P<0.05) of DEGs in control versus medium and control versus high (Fig 2b). The numbers of up and down regulated genes are 193 and 53, respectively.

We further performed enrichment analysis based on the list of 246 DEGs. A total number of 10 associated pathways were revealed from the enrichment analysis result (Table 1). These pathways are related to lipid metabolism, endocrine system, circulatory system, metabolism of cofactors and vitamins, as well as signaling transduction and molecules and interaction in the KEGG pathway categories.

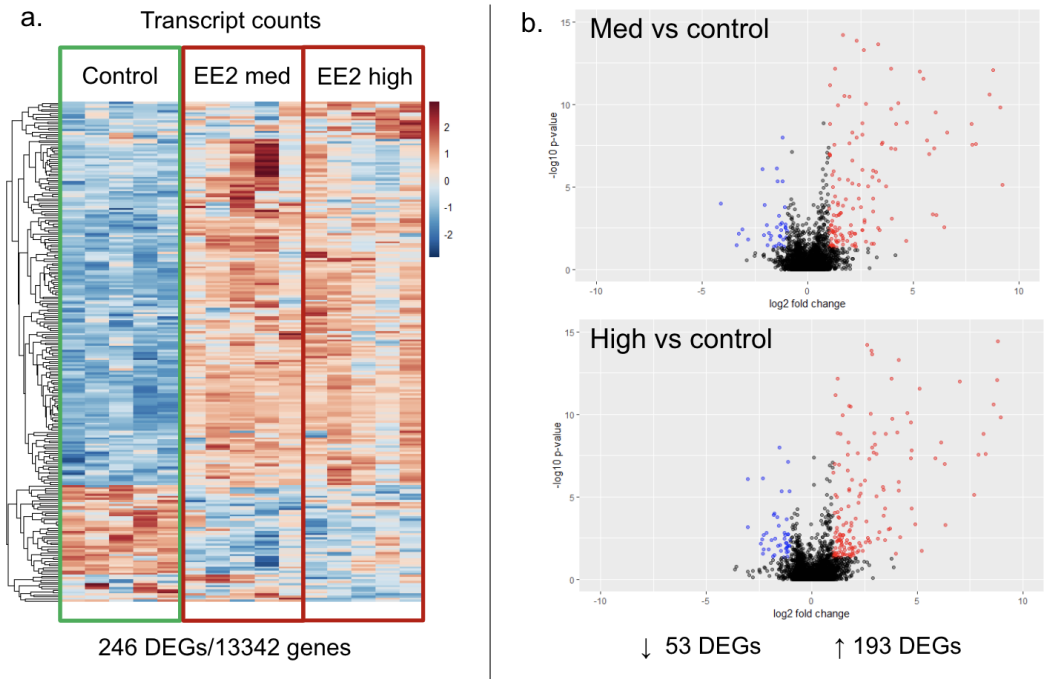

Figure 2. a. Heatmap shows the normalized transcript counts across 15 samples categorized in control, medium and high for the 246 DEGs in ELS stage JQ. **b.** DEGs log_2FC of ±1 versus $-log_{10}P$-values ($P<0.05$) from control v.s. medium and control v.s. High groups are displayed on the two volcano plots for ELS, respectively.

Enriched Pathway Names	Hits	Pval	AdjP
Primary bile acid biosynthesis	2/15	0.0247	0.865
Cytokine-cytokine receptor interaction	7/176	0.0253	0.865
Vascular smooth muscle contraction	5/109	0.0334	0.865
ErbB signaling pathway	4/79	0.0409	0.865
Steroid biosynthesis	2/20	0.0424	0.865
Progesterone-mediated oocyte maturation	4/80	0.0426	0.865
alpha-Linolenic acid metabolism	2/28	0.0776	1
PPAR signaling pathway	3/61	0.0793	1
Vitamin B6 metabolism	1/5	0.0802	1
Glycerophospholipid metabolism	4/102	0.0876	1

Table1. Enrichment analysis result of 10 pathways associated with the 246 DEGs in ELS JQ.

2. Adult JQ transcriptomic response upon exposure to EE2 treatment

Similarly, differential expression analysis was conducted on the adult stage JQ in response to EE2 exposure. It resulted in 128 DEGs. Next, we plotted a heatmap with the normalized transcript counts for the 128 adult DEGs in all the sample groups including control, medium and high doses (Fig 3a). Furthermore, volcano plots were drawn based on log_2FC of ± 1 on the x-axis and $-log_{10}P$-values ($P<0.05$) of DEGs on the y-axis in both control versus medium or high group. In total, there were 59 down regulated and 69 up regulated genes.

Using the identified 128 DEGs, we conducted enrichment analysis to pinpoint dysregulated pathways in JQ. In the end, we found 10 pathways shown in Table 2. These pathways are associated with amino acid metabolism, endocrine system and signaling transduction according to KEGG pathway map categories.

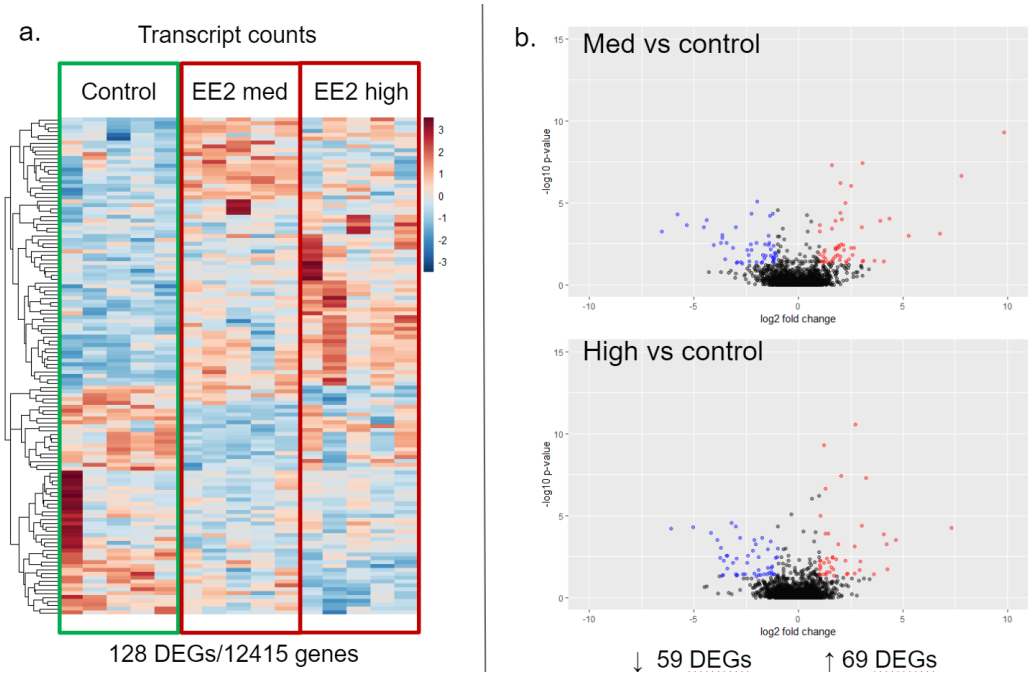

Figure 3. a. Heatmap shows the normalized transcript counts of 128 DEGs in adult stage JQ across 15 samples in the control, medium and high doses groups. **b.** Volcano plots indicate log_2FC of ± 1 versus -log_{10}P-values (P<0.05) of 128 DEGs by comparing control with medium and high groups in adult JQ.

Enriched Pathway Names	Hits	P.Value	AdjP
TGF-beta signaling pathway	3/83	0.0281	1
Tryptophan metabolism	2/37	0.035	1
Phenylalanine, tyrosine and tryptophan biosynthesis	1/5	0.0396	1
Insulin signaling pathway	3/124	0.0762	1
Glycerolipid metabolism	2/61	0.0855	1
PPAR signaling pathway	2/61	0.0855	1
Cardiac muscle contraction	2/64	0.0929	1
Thiamine metabolism	1/13	0.0998	1
Phenylalanine metabolism	1/14	0.107	1
ErbB signaling pathway	2/79	0.132	1

Table 2. Enrichment analysis result of 10 pathways associated with 128 DEGs in Adult JQ.

Discussion

1. The transcript counts of ELS Japanese hepatic genes show higher differences than the ones in adult

Comparing the K-means clustered heatmaps from the normalized counts (Fig 2a and 3a), we found that the differences between a control group and treated groups are higher in ELS than in adult. On the heatmap of ELS, ¾ of the DEGs are clearly up-regulated. However, the changes are not obvious in adult case. Although samples in high from the adult relatively showed stronger up-regulations compared to its medium, it still had smaller differences compared to ELS'. On the bottom ¼ of the heatamps on ELS, medium and high samples were down-regulated compared to the control, and again, adults showed the smaller difference. One of the reasons of different changes between ELS and adult is that adult quail's liver is fully developed, and thus, its metabolism of the chemical EEG works better.

To support our arguments from the heatmaps, we performed PCA, sIPCA and hierarchical clusters. Firstly, in PCA plots (Fig 4), as what heatmaps suggested, the control group is differentiated well from treatment groups in ELS. Unlike ELS, adults do not show the clear separation between control and treatment groups. Again, the sample 1 in control from the adults stayed away from all other samples in PCA.

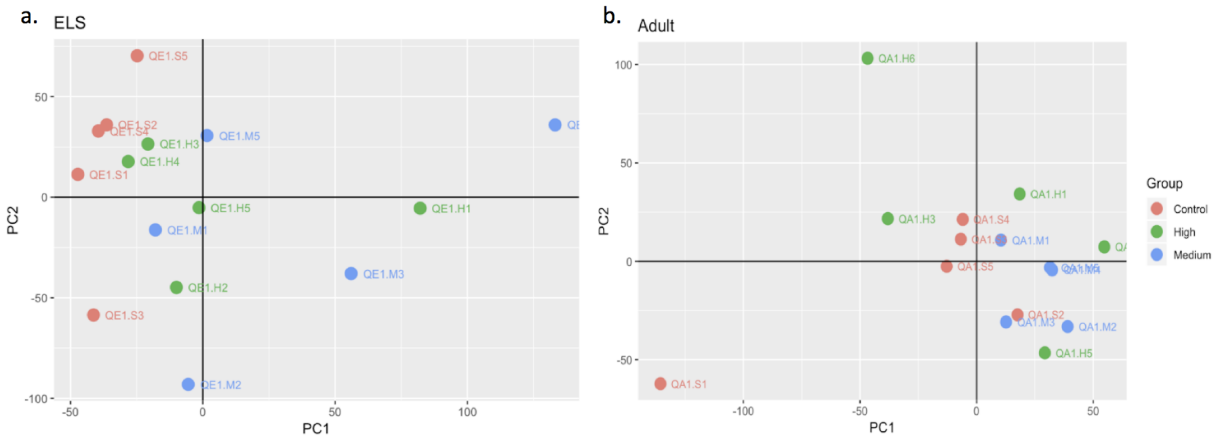

Figure 4. PCA plots. **a.** JQ ELS PC1 explains 18.2% and PC2 13.7% sample variance. **b.** JQ Adult PC1 represents 16.2% and PC2 11.6% sample variance.

Secondly, we performed sIPCA, since PCA has some drawbacks. PCA has the strong assumption that the gene expression follows a multivariate normal distribution, but it is not a case in our datasets. Furthermore, PCA decomposes the gene expressions based on maximization of its variance; however, our biological questions are not related to the highest variance. Again, in sIPCA (Fig 5), the control in ELS is well separated from medium and high groups; however, the adult's is not. The answers from previous analyses are strengthened by the hierarchical clusters (Fig 6). In conclusion, there are more numbers of DEGs (ELS has 246,

and Adults have 128) and higher log_2FC in ELS implying that ELS quails are more susceptible to EE2 than the adults.

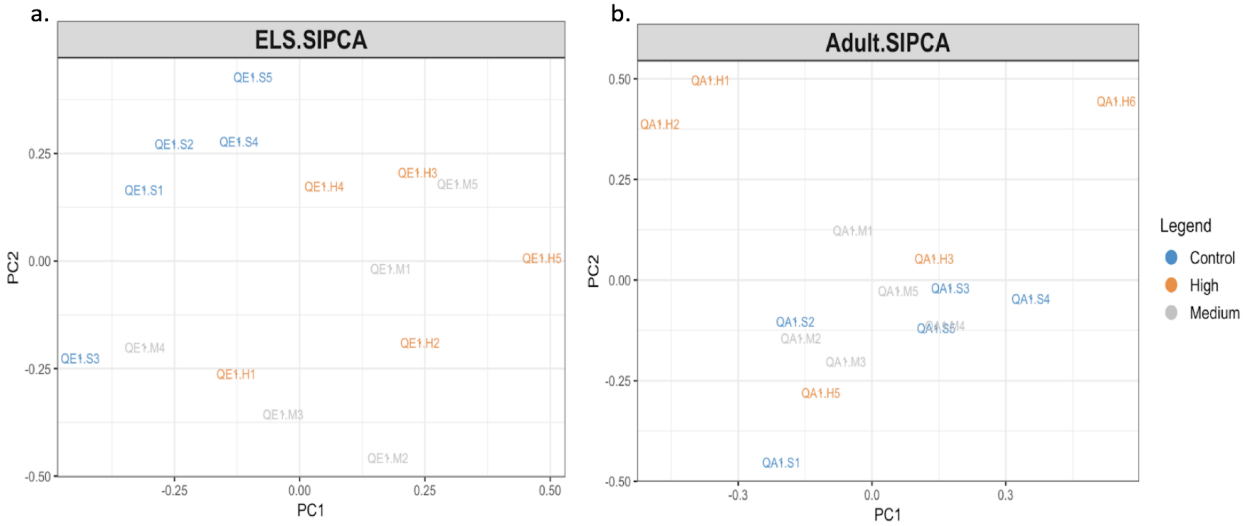

Figure 5. a. JQ ELS sIPCA with 25% explained variance. **b.** JQ Adult sIPCA with 24% explained variance.

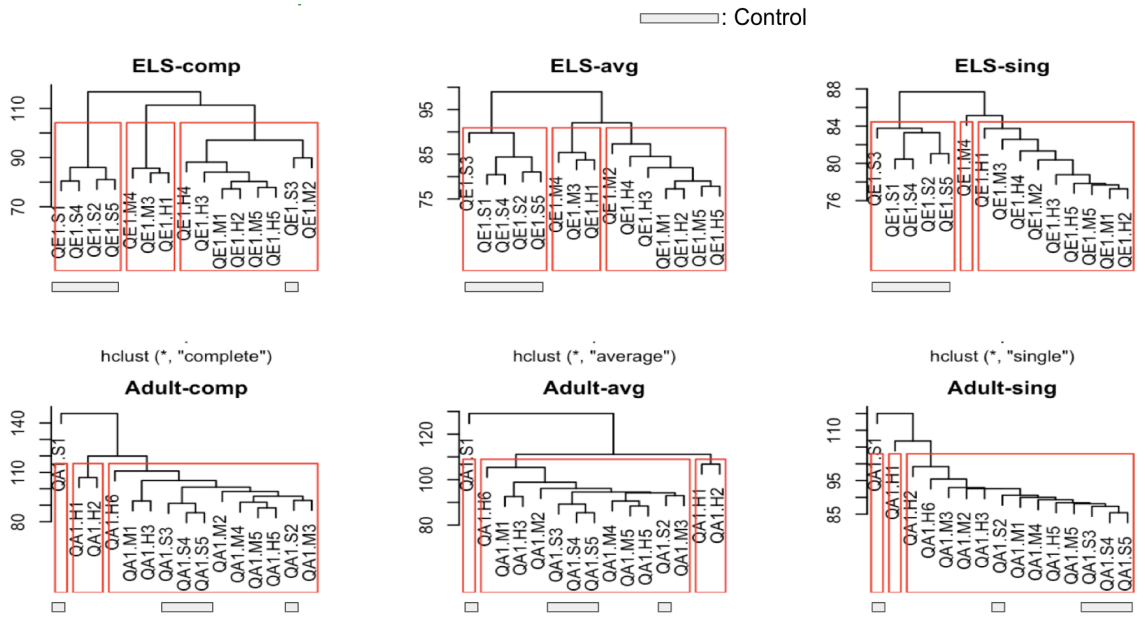

Figure 6. Hierarchical clusters with the linkages: complete, (first column) average, (second) and single (third)

2. Common DEGs in ELS and adult JQ

There were 17 DEGs shared by both ELS and adult JQ (Table 3). We further investigated their degrees and directions of FCs (Table 3). The degree of dysregulations presented, by absolute log_2FC, was higher in ELS than it was in adult, and the pattern was visually captured in Fig 7. We observed that six genes were dysregulated in the opposite direction. In contrast, seven

genes were dysregulated in the same direction in ELS and adult JQ. Specifically, all of them were upregulated in both medium and high doses of EE2 from both life stages.

17 Overlapping DEGs	ELS Medium vs Control	ELS High vs Control	Adult Medium vs Control	Adult High vs Control
FABP7	8.888	3.317	3.539	-0.229
NPTXR*	8.928	3.187	0.793	1.878
TCN2**	6.069	3.596	-5.393	-0.798
LOC107317492*	0.892	3.812	3.341	3.384
GPR50**	2.174	5.188	-0.153	-2.534
LOC107317574**	1.062	5.714	-2.676	-0.223
LOC107314278**	6.572	4.210	-5.203	-4.699
LOC107310248	3.495	7.039	-2.990	3.876
LOC107317565	6.138	2.286	-0.558	3.733
LOC107305801*	2.940	0.943	2.355	1.238
LOC107319873*	2.520	6.460	2.474	0.678
LOC107317352*	6.327	8.216	2.740	1.530
ALCAM	3.283	2.393	-2.561	1.188
UNC93A*	3.611	3.021	2.193	0.141
ZBTB16**	2.778	1.376	-0.473	-0.421
LOC107309939*	0.993	1.502	0.267	3.179
LOC107319648**	-3.055	-1.079	1.300	1.010

Table 3. 17 overlapping DEGs between ELS and adult JQ, and their log_2FC values. The values for up-regulation and down-regulation are colored in red and blue, relatively. * and ** denote DEGs dysregulated in the same and opposite direction, respectively.

Moreover, strong associations between the log_2FC values of ELS and adult were found (Table 4). Overall, ELS captured both larger numbers and higher degrees of dysregulated genes compared to the adult, indicating that ELS is more susceptible to EE2 than adult in gene expression levels.

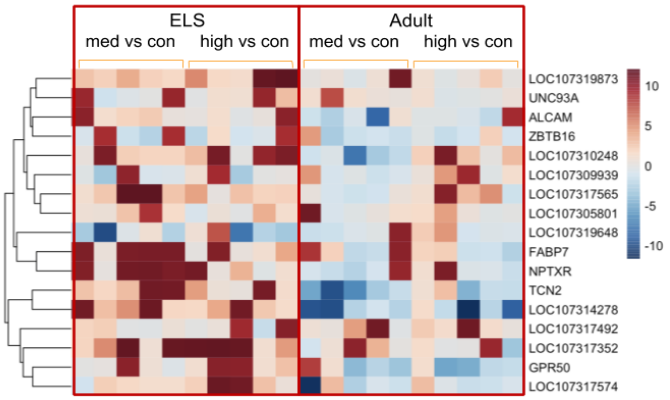

Figure 7. Heatmap of log_2FC values for common DEGs.

Parameters	Correlation
log_2FC (High vs Control)	0.855 *
log_2FC (Medium vs Control)	0.617 *
logCPM	0.615 *
LR	-0.237
p-value	-0.139
adjusted p-value	-0.134

Table 4. Correlation test result on ELS and adult JQ. Parameters of common DEGs are compared.

3. Common dysregulated pathways in ELS and adult JQ

Amongst total of 18 enriched pathways hit by ELS and adult JQ DEGs, only 2 pathways were in common: ErbB signalling pathway and PPAR signalling pathway (Table 1 and 2). Although the number of dysregulated pathways were the same, the number of genes' hits for the common enriched pathways were higher in ELS than in adult. Moreover, in ErbB signalling pathway, more upstream genes were present in ELS. Last but not least, in PPAR signalling pathway where ELS and adult had the same upstream DEGs, the common gene FABP7 has higher FC in ELS.

We further explored the KEGG categories of the 10 different dysregulated pathways in both life stages. Surprisingly, the KEGG categories mostly overlapped in ELS and adult, with the common ones: circulatory system, endocrine system, metabolism of cofactors and vitamins, lipid metabolism and signal transduction. The only exception was the amino acid metabolism dysregulation, which only presented in adult with 3 hits out of 10 pathways. However, in ELS, 4 out of 10 pathways were related to lipid metabolism. From this observation, we came up with a new hypothesis that ELS and adult counter the EE2 toxic effects with different molecular mechanisms.

4. Limitations and suggestions for future studies

First, we assumed that lab-raised JQ could biologically represent the wild JQ response. However, it might not be true due to the fact that domestic JQ is likely to have different genetic background compared to the wild one[14]. In addition, Bely et al.[15] suggested that endocrine system regulation is sensitive to environmental perturbations. Moreover, in the same study, variation in the diet between animal in the lab and in the wild could alter hormone levels in body, thus influencing gene expression level. Consequently, our result might not be able to capture the real biological effect of EE2 in terms of differential gene expression pattern in wild JQ. This gap

might be wider for other wild avian species due to genetic, anatomical, physiological and biochemical differences. So, we suggest to conduct RNA sequencing on wild JQ's or other wild birds' eggs from relatively less polluted area in similar experimental scheme and do comparative analyses with our data. It will deepen the understanding of difference between wild and lab organisms.

Second, there were limited information on the JQ genes' functions. For instance, only about 14% of DEGs for each of two life stages could be mapped to the KEGG pathways and many of the DEGs had no known functions online. More genes might be involved in the same pathways, but they had not been characterized yet. As a result, the adjusted p-value for our enrichment analysis was not significant. There are lots of space left for investigations to reveal JQ genes' functions and systematic roles.

Third, we assumed that the dosage (LD20) and timing (mid-incubation in ELS) of exposure would capture the gene expression in response to EE2. We tried to overcome lack of information on relationship between dose and gene expression in JQ by including one more dose (1/10 of LD20), but there is still needs to try more doses and conduct dose-response modelling. In addition, since gene expression pattern changes dynamically during embryonic development, other periods in ELS than mid-incubation should be explored.

Fourth, we cannot assure if all DEGs captured by RNA sequencing data are true DEGs. We want to validate the biologically relevant DEGs, for example the DEGs related to endocrine system regulation, with qPCR experiments. The qPCR would allow us to pinpoint the real differential gene expression effects and eliminate false positives.

Conclusion

Upon the exposure to EE2, the $\log_2$FC values of two life stages are correlated, and enriched pathways are related to the similar biological functions. Thus, gene dysregulation patterns of ELS and adult JQ are biologically similar. However, compared to adult, ELS shows higher number of DEGs, higher degree of dysregulation, higher number of DEGs mapped to KEGG pathways and higher number or FC of upstream genes in the common pathways. From these results, we concluded that ELS JQ is more susceptible to EE2 than adult JQ. It implies that ELS JQs can be a better ecotoxicological indicator than adult JQs.

References

1. Chen, Y., Li, M., Yuan, L., Xie, Y., Li, B., Xu, W., …, and Wang, R. (2017). Growth, blood health, antioxidant status and immune response in juvenile yellow catfish Pelteobagrus fulvidraco exposed to α-ethinylestradiol (EE2). *Fish & shellfish immunology*, *69*, 1-5.
2. Jobling, S., Nolan, M., Tyler, C. R., Brighty, G., and Sumpter, J. P. (1998). Widespread sexual disruption in wild fish. *Environmental science & technology*, *32*(17), 2498-2506.
3. Länge, R., Hutchinson, T. H., Croudace, C. P., Siegmund, F., Schweinfurth, H., Hampe, P., …, and Sumpter, J. P. (2001). Effects of the synthetic estrogen 17α- ethinylestradiol on the life- cycle of the fathead minnow (Pimephales promelas). *Environmental Toxicology and Chemistry*, *20*(6), 1216-1227.
4. Mattsson, A. and Brunström, B. (2017). Effects of selective and combined activation of estrogen receptor α and β on reproductive organ development and sexual behaviour in Japanese quail (Coturnix japonica). *PloS one*, *12*(7), e0180548.
5. Jaspers, V. L. (2015). Selecting the right bird model in experimental studies on endocrine disrupting chemicals. *Frontiers in Environmental Science, 3*, 35.
6. Mohammed, A. (2013). Why are early life stages of aquatic organisms more sensitive to toxicants than adults? *New insights into toxicity and drug testing*: InTech.
7. Jax, E., Wink, M., and Kraus, R. H. (2018). Avian transcriptomics: opportunities and challenges. *Journal of ornithology*, 1-31.
8. Alexander-Dann, B., Pruteanu, L. L., Oerton, E., Sharma, N., Berindan-Neagoe, I., Módos, D., and Bender, A. (2018). Developments in toxicogenomics: understanding and predicting compound-induced toxicity from gene expression data. *Molecular omics*.
9. Hanafy, A. M., Sasanami, T., and Mori, M. (2006). Effects of diethylstilbestrol and ethinylestradiol on gene transcription of very low- density apolipoprotein II in the liver of Japanese quail, Coturnix japonica. *Environmental toxicology and chemistry*, *25*(5), 1354-1359.
10. Halldin, K., Berg, C., Brandt, I., and Brunström, B. (1999). Sexual behavior in Japanese quail as a test end point for endocrine disruption: effects of in ovo exposure to ethinylestradiol and diethylstilbestrol. *Environmental Health Perspectives*, *107*(11), 861-866.
11. Basu, N., Crump, D., Head, J., Hickey, G., Hogan, N., Maguire, S., ... & Hecker, M. (2019). EcoToxChip: A next-generation toxicogenomics tool for chemical prioritization and environmental management. Environmental toxicology and chemistry, 38(2), 279.

12. Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics, 26*(1), 139-140.
13. Schurch, N. J., Schofield, P., Gierliński, M., Cole, C., Sherstnev, A., Singh, V., ... & Blaxter, M. (2016). How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use?. Rna, 22(6), 839-851.
14. Nunome, M., Nakano, M., Tadano, R., Kawahara-Miki, R., Kono, T., Takahashi, S., … Matsuda, Y. (2017). Genetic Divergence in Domestic Japanese Quail Inferred from Mitochondrial DNA D-Loop and Microsatellite Markers. PloS one, 12(1), e0169978.
15. Bely, A. E., & Weisblat, D. A. (2006). Lessons from leeches: A call for DNA barcoding in the lab. Evolution and Development.