Project 08

Meng Han, Sharmim Sultana, Shalmali Hemant Kadam, Feng Chen

The goal of this assignment is to use RNA-seq data from BMAT and MM cells under conditions such as co-culture and mono-culture, and figure out how gene expression is changed when the two cell grow together. When two cells put together, there is chance they communicate with each other by change their expression, so the idea of this project is to check if there is any transcriptional shift under co-culture and what biology pathways.

The matrix is inside the RData file, which is design = replicate + co_cultured, which means we will first remove the replicate difference by DESeq2 then looks at the co_cultured effect. It will not be idea to remove the replicate from the analysis since this might create more noisy.

The parameters I used for the analysis is padj < 0.1 and |log2FC| >= 1, the reasons I set those threshold is because padj that are less then 0.05 will be too sensitive to this data, it will return very little information, therefore, is more reasonable to use 0.1, as for |log2FC| gives genes at least two fold change, and I want to only keep genes that really change a lot, therefore I set it to greater or equal to 1.

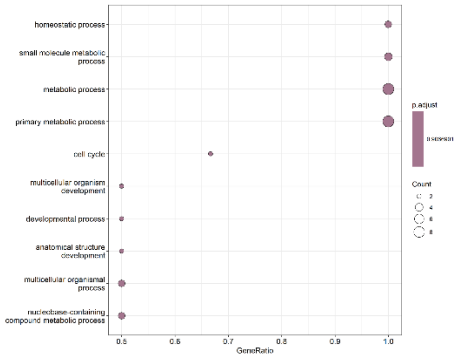
Based on the differential results, BMAT doesn't change much under co-cultured conditions because it only have small numbers of differential expressed genes, most of its genes are metabolic genes. However, for MM cell have a lot of differential genes and many of them are immune related genes. When looking at the shrinkage result, it lowers the high fold change values from low count genes, by that the fold become more stable and make much more sense. For the pathway enrichment analysis, BMAT's pathways are mainly metabolic related and not significant, which is expected, as for MM cell it have strong enrichment for immune activation pathways such as cytokine production. Through this analysis, we can kind of understand how each cell different from each other, BMAT barely respond under co-cultured but MM respond significantly, and this can be support by the DEGreport plots shows in the below, the pattern are consistent across the replicates, which means the differential expressions differences are real not by the random chance.

Overall, BMAT did not shows strong differential expression under co-cultured conditions, and enrichment analysis did not give significant pathways, means co-cultured did not affect BMAT cells. For MM, it shows very strong change under co-cultured conditions, many of its gene are up-regulated especially ones to related to immune system, which be support by GSEA plots, ridge plots, and dot plots.

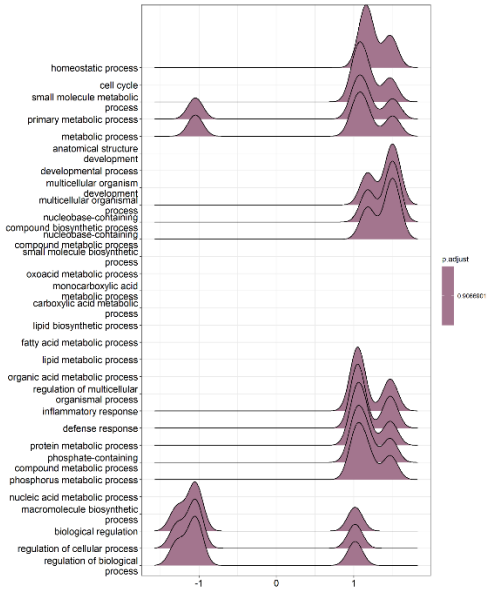
One big challenge of this assignment is figure out how to run the analysis without install the clusterProfiler package in the server, which I eventually did in my local Rstudio where I can install the clusterProfiler.

Through this project me and my group learned how to apply the enrichment analysis which we think could be really helpful for the future. We also learned the importance of shrinkage step, which a step we never heard about before, but it is essential for remove the high, unstable fold change from low-count genes, this is something we will be apply for future DESeq2 analysis.

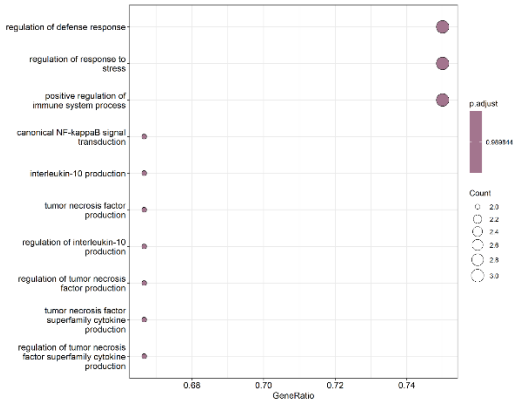
Graphs:



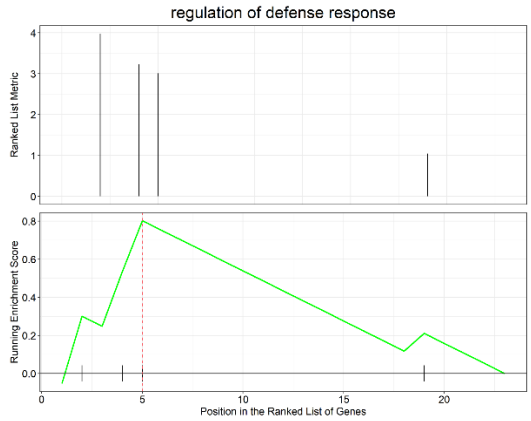
Pathways like metabolic and homeostatic process are upregulated, but none of them are significant due to the high p.adjust value. The GeneRatio are big, this mainly is because BMAT have small numbers of differential expressed genes, so genes are sized up make the ratio look large.



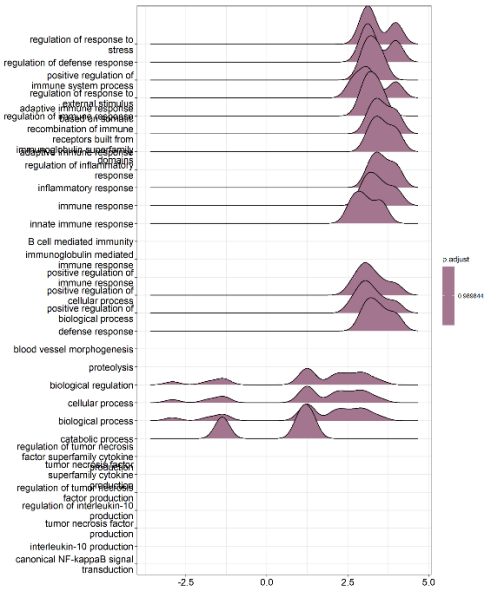
This ridgeplot shows the distribution of the enrichment score for each pathways in the BMAT. Most of the curves are flat, meaning low enrichment score, BMAT does not have strong pathway level shifts under the co-cultured conditions. Even few metabolic and lipid related pathways shows slight positive enrichment score, but those are not significant since the p.adjust value is pretty high.



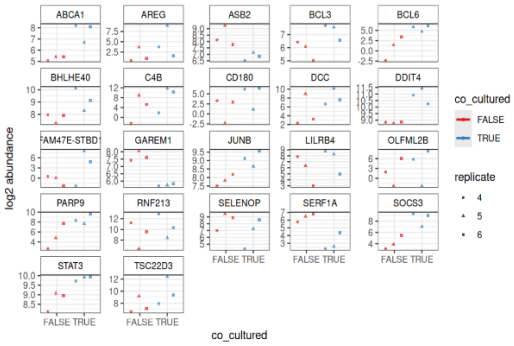
This dot plot shows the top enriched GO terms for MM cell under the co-cultured condition, compared to what BMAT have, MM shows a much clearer immune related signal. Even though the p.adjust value still high, the pattern suggests MM cell react more to co-cultured conditions.



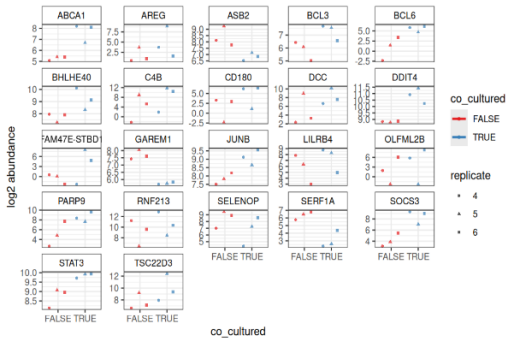
This plot shows how regulation of defense response pathway behaves across the gene list in MM. The green line represents the enrichment score, and it peaks pretty early, which means most of the genes in this pathway are highly ranked in the positive direction. The black ticks at the bottom show where the pathway genes show up in the list, and most of them are on the left side, so they are all pushed upward in co-culture.



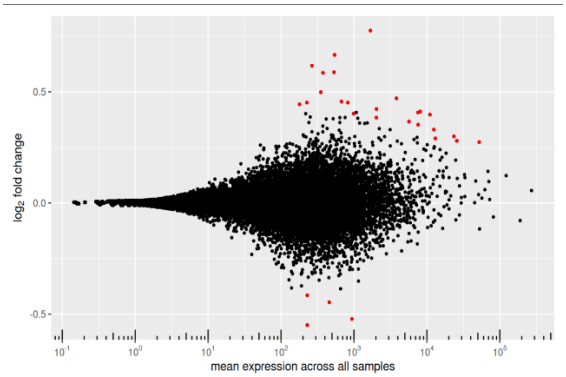
This ridge plot shows the distribution of gene scores of immune pathways in MM cell. Most of the peaks are toward the right side, means that genes in the pathways are more likely to have positive fold changes under co-cultured condition. Many of categories are immune related, means MM cell activate immune related genes when they are co-cultured.



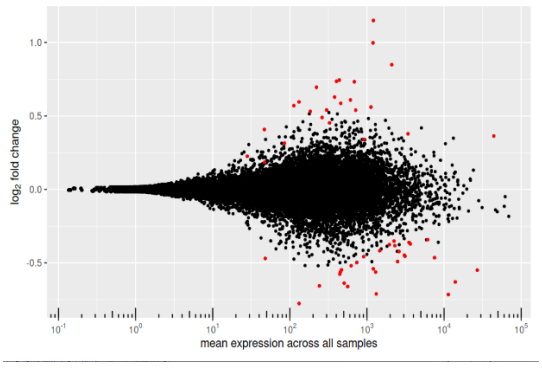
This plot is the expression pattern of top MM differential expression genes across the co-cultured vs. mono-cultured genes. Each panel is one gene, and the red points are the mono-culture while the blue points are the co-culture. One thing I notice is that most of these genes shift upward in the blue group, meaning they get upregulated when MM cells are co-cultured.



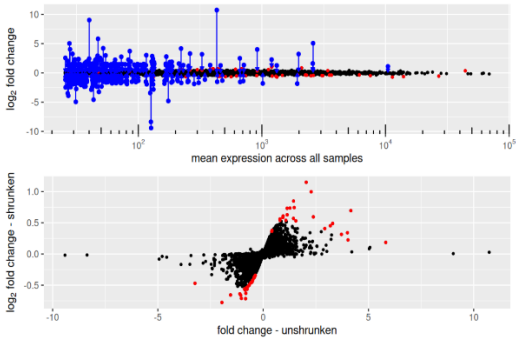
This figure shows how the MM differentially expressed genes change between mono-culture and co-culture. Each panel is one gene, and the three shapes inside each panel are the three replicates. A lot of these genes shift upward under co-culture, meaning they get higher expression when MM cells are grown with the other cell type.



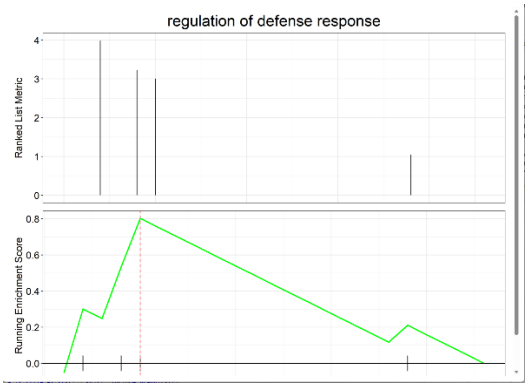
This plot is the MA plot for the BMAT differential expression results. Where each dot represent one gene, the x-axis is the mean expression level and y-axis is the log2 fold change between two conditions. Most genes are black and centered around zero, which means they don't really change that much. The red dots are the significant DE genes, and they mostly appear on the upper side, so these are the genes that get up-regulated under the co-culture condition.



This plot is the MA plot for the MM samples. Same as the BMAT one, each dot is a gene, x-axis is the average expression, and y-axis is the log2 fold change between co-culture and mono-culture. Most genes are around zero meaning no big change, and the red ones are the DE genes that passed the filter. For MM, a larger spread both above and below zero compared to BMAT, meaning MM cells have more genes responding in both directions.



This figure shows the shrinkage results from DESeq2. The top one is the unshrunken log2 fold change, which there are lot of high (blue points) values especially for low expression genes. The red ones are the significant genes. the bottom one is the shrunken log2FC compared to the original unshrunken values. After shrinkage, all the extreme values get pulled closer to zero, which makes the fold changes more stable and more realistic, especially for those low-count genes.



This graph is the GSEA running score plot for the pathway “regulation of defense response” in the MM cells. The black tick marks at the bottom show where the genes from this pathway appear in the ranked gene list. The green line is the running enrichment score, and it peaks around the left side, meaning most of the important pathway genes are ranked near the top and are up-regulated in the co-culture condition. The red dashed line marks the peak.