Project 2 Presentation Script

Introduction

- Hello everyone. For my second project of this program, I worked with Dr. Larson on the generation of orthotopically functional salivary glands from embryonic stem cells.

Background

- This topic was derived from the scientific article shown here, which describes an experiment where the salivary glands of a mouse embryo were successfully regenerated after being induced with the Sox9 and Foxc1 transcription factors.
- My role was to see if the maturation of the induced salivary glands or iSGs was promoted by the addition of mesenchyme cells, a type of embryonic tissue that will later develop into lymphatic tissue as the mouse embryo matures.
- To answer this question, I performed differential gene analysis to find any significant genes within the iSGs that could possibly result in faster maturation. Then I used gene ontology enrichment analysis to understand the biological function of those genes and correlate that to better maturation.

Materials and Methods

- Since this project is based on an already performed experiment, the FASTQ files were already aligned, therefore, providing me with the ReadsPerGene files needed for my project.
- These files were then used for the differential gene expression analysis pipeline with the DESeq2 program in RStudio to create the visual figures shown in the following slides.
- Based on the results from this process, I performed gene ontology enrichment analysis using the online program Metascape, to get a better understanding of the roles of the statistically significant genes.

DGE Analysis

- Now, starting with differential gene expression or DGE analysis. Here, I searched for differentially expressed genes or DEGs within the experimental data.

PCA Plot

- This PCA plot shows the variance between the 3 samples of iSGs without the mesenchyme cells in orange and the 3 samples of iSGs with the mesenchyme cells in green.
- As you can see here, there is only a small amount of variance between the samples.
- And you can see that they do not cluster by condition, indicating that the conditions are more similar than they are different, and also that the samples of each condition have a bit of variance between them.

MA Plot

- Next, I generated a MA plot to show the overall trend of differential gene expression in regards to the relationship between the Log2Fold change and the mean normalized counts of the experimental data.
- The Log2Fold change describes the change in expression of a specific gene in the iSG samples compared to the iSG+mesenchyme samples, while the normalized counts is the result of the read counts of a specific gene divided by the total number of reads in that sample.
- As you can see here, there is a very small amount of genes that are differentially expressed, since there is only one gene that is significantly upregulated as indicated by the single blue point that I tried to zoom in on.
- You can tell that this gene is upregulated due to its position in the positive Log2Fold change region.

Volcano Plot

- This idea is further supported by the volcano plot shown here.
- This plot also displays the overall trend of differential gene expression but instead using the relationship between the Log2Fold change and the adjusted p-value.
- The cutoffs were set to a Log2Fold change greater than 1.5 and an adjusted p-value less than 0.05.
- As you can see zoomed in here, that the single Mir668 gene is significantly upregulated.

Normalized Plot Count

- However, this is a bit contradicted by the normalized count plot shown here.
- This plot displays the normalized counts for the most statistically significant gene based on adjusted p-value.
- This gene, Chst4, gives a low adjusted p-value in all samples except one iSG+mesenchyme sample, which does correlate with what was shown in the PCA plot. So it is possible that that PCA plot clustered by adjusted p-value.

Heatmap

- Lastly, I generated a heatmap to get a closer look at the genetic makeup of each of the samples, and to see if any more genes could be of interest in relation to my goals for this project.
- This heatmap displays the 10 most differentially expressed genes and their count abundance in comparison to other libraries.
- Expression levels are shown in a cold to warm color tone with down, or low, expression as cold, and up, or high, expression as warm.
- This plot supports how diverse the gene regulation is between the six samples, regardless of their condition.
- Some genes that were of interest to me were the upregulated genes in the iSG samples: Reg3g, Car3, Myo18b, and Atp2a1
- The upregulated genes in the iSG+mesenchyme samples were: Gm14744, H2-Q7, and Chst4 as seen in the normalized counts plot

GO Enrichment Analysis

- With these genes in mind, I then performed gene ontology enrichment analysis to learn their specific biological roles.
- To recap, we have a total of 8 genes of interest:
- Mir668 from the MA and Volcano plot
- Chst4 from the normalized counts plot and heatmap
- And Reg3g, Car3, Myo18b, Atp2a1, Gm14744, and H2-Q7 from the heatmap.

Metascape Gene Ontology Analysis Results

- I inputted these genes into Metascape to obtain the results shown here.
- To summarize:
  - Gm14744 did not come back with a result, so its function is most likely unknown
  - Chst4, Car3, and Reg3g play a role in metabolic processes
  - Myo18b and Atp2a1 are involved in cardiac development
  - And Mir668 is responsible for transcription and gene splicing
- All of which are important, but since my project is focused on mesenchyme cells which is related to the lymphatic system and immune response, H2-Q7 was of most interest to me, which promotes T-cell response to tumor cells.
- Chst4 and Car3 were also of interest because they are also involved somewhat in immune response, keeping of note that they both are mostly related to metabolic functions.

(Next Slide)

- Finally, I generated normalized counts plots for each of those 3 genes to see their regulation within each sample.
- As expected, H2-Q7 is up regulated in all 3 iSG+mesenchyme samples and down regulated in all iSG samples without mesenchyme cells.
- Chst4 is only up regulated in one of the iSG+mesenchyme samples, as seen previously.
- And Car3 is upregulated in two of the iSG samples, but down regulated in all other samples.
- All of this correlates with what was seen in the heatmap.

Conclusions

- In conclusion, the addition of mesenchyme cells to the induced salivary glands promoted immune health within each of the three samples.
- The H2-Q7 gene specifically associated with T-cell tumor targeted immune response, was up regulated significantly within all three samples.
- However, since the main genes of interest between the conditions did not produce any results that can be directly related to faster maturation, I cannot confidently assume that the addition of the mesenchyme cells promote better maturation.

Future Directions

- But this information does open room for further research into what cells or tissues directly develop from the mesenchyme cells. This way a maturation rate can be recorded.
- It would also be interesting to see this experiment done with samples not induced with transcription factors and on other organs instead of salivary glands. Increasing the amount of samples from just three of each condition could also help to increase accuracy of the results .
- Also, since we've seen how the addition of the mesenchyme cells contributes to immune health, it would also be interesting to see how the samples would react to exposure to factors that would test their immune health, such as infection or radiation that could induce tumor formation.

Acknowledgements

- And with that, I will like to thank Dr. Larson for supervising my research for this project, and also my group leaders Eric and Liv.
- I would also like to acknowledge all contributors to the scientific article that provided my research dataset, and the faculty and staff of this bioinformatics program for providing me this research opportunity.

Thank You!

- Thank you for listening.