**CAPSTONE PROJECT- 1**

***"Pathogen-Specific Host Response in Infectious Keratitis: Insights from Transcriptional Profiling"***

**RASUTI - BT22GBT287**

**PRIYANSHU CHATTERJEE - BT22GBT374**

**Abstract:**

**Purpose:-** One of the most common corneal diseases is Keratitis. Numerous pathogenic infections have an impact on it. Our goal was to use RNA sequencing techniques to gain an understanding of the cellular and molecular mechanisms that underlie the host's response to viral keratitis. We aim to get an insight into the mechanisms that occur in human corneal tissue by examining the transcriptional profiles.

**Methods:-** After retrieving the RNA sequence data for healthy normal corneal samples (n=4), bacterial/fungal samples(n=4), viral samples(n=4) and keratoconus samples(n=4), we uploaded the data to the Galaxy Web Server and used a variety of tools to analyse the sequences in order to gain a better understanding.

**Result:-** We identified our primary DEGs using the provided data and we performed a Functional Enrichment Analysis on them to determine which genes were suppressed and which were highly expressed, which also helped us in understanding the mechanisms of the underlying DEGs.

**Conclusion:-** In conclusion, this analysis provides us valuable insights into the cellular and molecular mechanisms of the identified DEGs, enabling the selection of potential biomarkers for further investigation.

**Keywords:** RNA Sequencing, Galaxy Web Server, Infectious keratitis, viral keratitis, bacterial keratitis, fungal keratitis, Differential Expressed Genes, Functional Enrichment Analysis.

**Introduction:** Keratitis, a corneal illness that causes inflammation and is brought on by pathogenic bacteria, served as the basis for this investigation. Viral keratitis, bacterial/fungal keratitis, and keratoconus were the main subjects of our conversation. We primarily looked at the ways that various forms of keratitis impact one another. We performed a separate analysis for each type of keratitis and the results were as anticipated.

Since it was the sole platform used for all research experiments and analyses, the Galaxy web server was widely utilised.

**Related work:** We read a lot of different research papers to get a better idea of what we needed to do in order to conduct our research correctly.

* The first paper we took into consideration was titled as “**Transcriptional profiling specifies the pathogen-specific human host response to infectious keratitis”.** The primary focus of this publication was their investigation of the condition using high throughput RNA transcriptional profiling techniques. They collected a wide range of keratitis samples. They even collected some healthy corneal samples in order to distinguish between the bacterial/fungal, fungal, viral, and parasite infections. As possible diagnostic biomarkers, they found 561 genes specific to bacterial/fungal keratitis and 216 genes specific to viral keratitis. Additionally, they employed human corneal specimens from patients with infectious keratitis and keratoconus that were FFPE (Formalin-Fixed Paraffin-Embedded). It also displayed the DEGs analysis that they discovered during their investigation. This paper also provided us with a number of outcomes in the form of graphs that displayed the number of genes that were suppressed and those that were expressed. They took the shape of volcano plots and heat maps.
* The next article we cited was called **"RNA-Seq-Based Transcriptome Analysis of the Gene Expression Profiles Associated with Fungal Keratitis in Mice".** This gave us a better knowledge of how to carry out our RNA sequence analysis for this project. It largely discussed studying the transcriptional patterns of fungal keratitis and conducting RNA sequence analysis on it.
* This research paper, **"Comparison of RNA-seq and microarray-based models for clinical endpoint prediction"** discusses the key distinctions between microarray-based models and RNA-sequence analysis. Additionally, this research discusses which approach performs better over the long term and which approach produces better outcomes.
* The title of the final paper we cited was **"Transcriptome profiling: methods and applications".** The primary goal of this paper was to give readers a better grasp of transcriptional profiling techniques.

**Proposed Methodology:**

1. We retrieved RNA sequence data from NCBI using Sequence Read Archive (SRA) filter which included 4 healthy normal corneal samples, 4 bacterial/fungal samples, 4 viral samples and 4 keratoconus samples.
2. Then we ran a quality check on the following sequences using fastQC in Galaxy Web Server which gave us the quality check in the form of

graphs, tables etc.

1. Further, we aligned the RNA sequences and removed the unnecessary parts of the sequence using the TopHat tool.
2. We also realigned the above sequences using the HISAT2 tool in which the reference genome “hg38” played an important role.
3. After getting the HISAT2 results, we used a tool called featureCounts which helped us in measuring the Gene Expression of the sequences. This gave us the results in form of two files, namely -

* Counts file
* Summary file

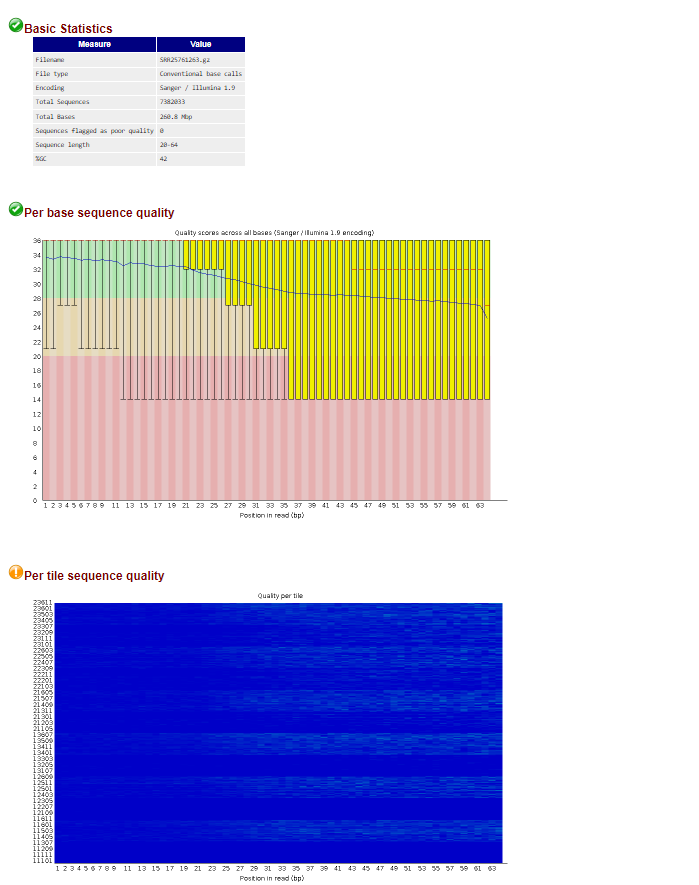
1. Further we used the edgeR tool to perform Differential Expression on the counts data which gave us a number of gene IDs. These IDs helped us in finding the Top 10 DEGs which we used further.
2. After finding the Top 10 DEGs, we uploaded them on the Enrichr website and got the results in the form of pathways and ontologies which helped us in understanding the mechanism of the pathways.

**Technology/Techniques/Methods:**

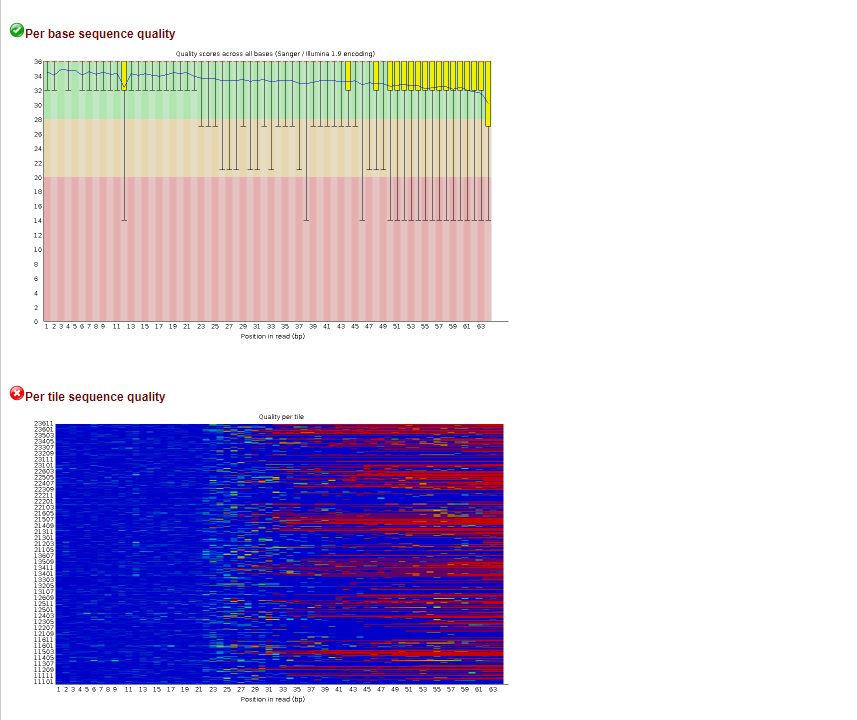
For our project, we used different methods, numerous tools in “**Galaxy Web Server”** and these tools gave us the result that we expected.

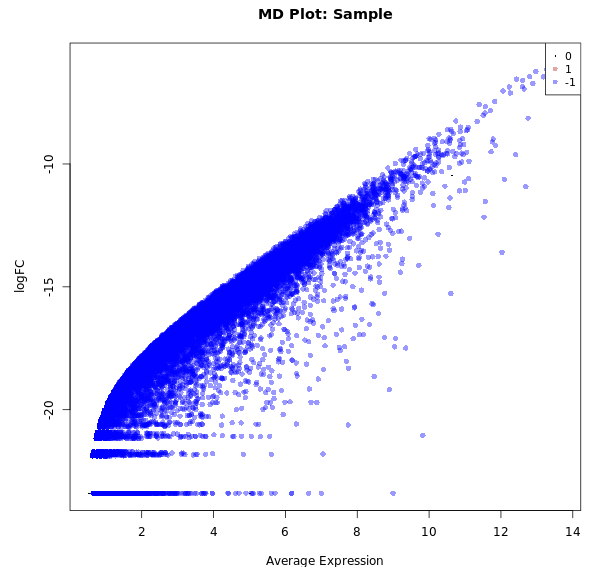
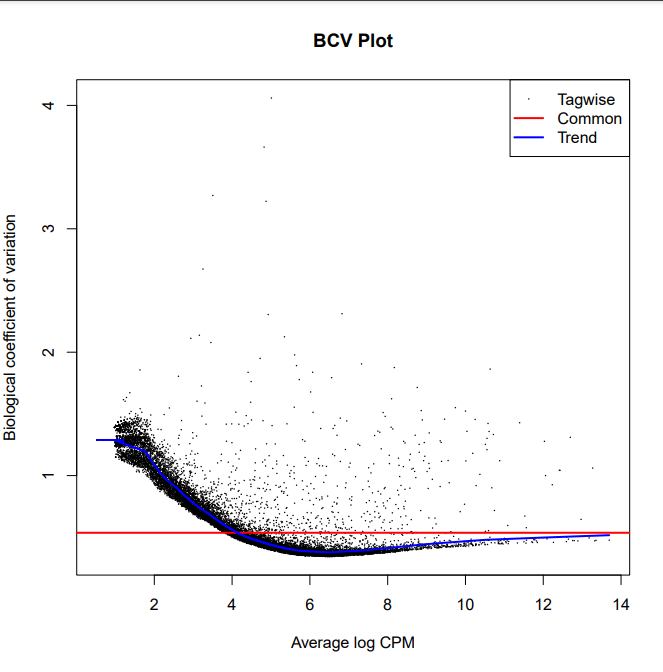
1. **RNA Sequencing:** A method for studying the transcriptome is RNA sequencing (RNA-Seq), which involves sequencing the RNA molecules in a sample. RNA is transformed into complementary DNA (cDNA) and sequenced in order to discover RNA variations, identify new transcripts, and analyse gene expression. Widely employed to investigate gene regulation and biological reactions, RNA-Seq offers in-depth insights into cellular processes.
2. **NCBI:** It provides access to biomedical and genomic information and facilitates research in these fields. In this we particularly use the SRA filter which is a powerful tool that helps users refine their search for high-throughput sequencing datasets. It allows filtering based on specific parameters such as organism, sequencing platform, study type, library preparation method, and data type (e.g., RNA-Seq).
3. **TopHat**: To obtain a more precise result, we utilised the Tophat tool to remove unnecessary portions of the sequence.
4. **HISAT2**: Following TopHat trimming, we utilised HISAT2 to realign our sequence, which was run against the 'hg38' genome file.
5. **featureCounts**: This tool was used to measure the gene expression of the sequence and for this we enabled an option called "featureCounts-built in".
6. **EdgeR**: This tool was used to do Differential Expression on the count data that we obtained from the featureCounts. To improve the results, we grouped the samples to obtain better results, which mostly provided us with tabular data with many gene IDs.
7. **Enrichr**: This web tool gives us the pathways and ontologies in the form of bargraphs.

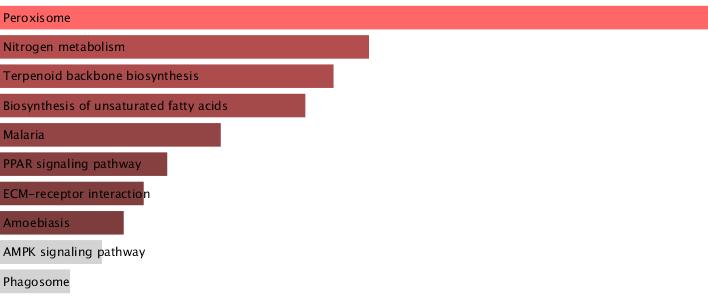
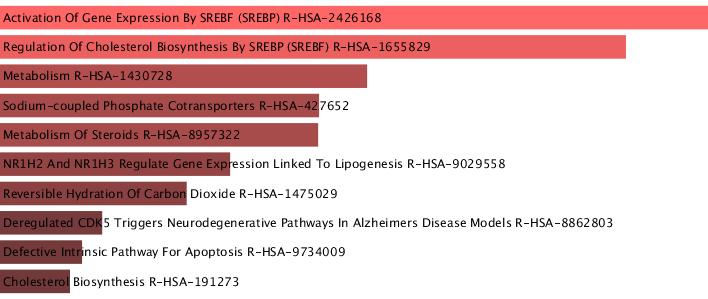
**Result and Analysis:**

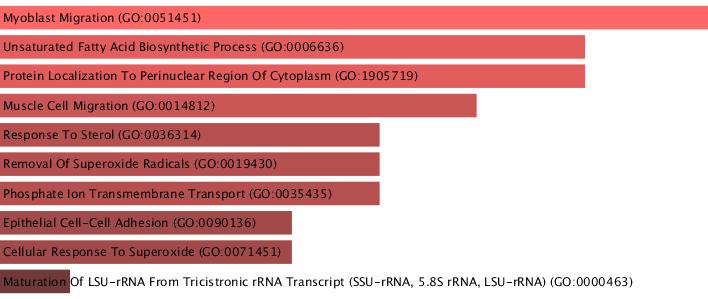
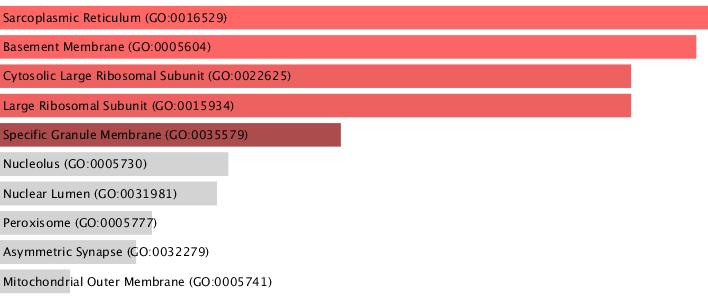


FastQC report of bacterial/fungal keratitis

  
FastQC report of TopHat Viral keratitis



BCV plot of EdgeR MD plot: Sample of EdgeR  
  
  
  
  
  
KEGG pathway  
  
  
Panther 2015  
  
  
Reactome Pathway

  
GO Biological Process  
  
  
GO Cellular Process  
  
**Conclusion and Future prospects:**

In conclusion, our study effectively analysed transcriptional patterns linked to keratitis and keratoconus using RNA sequencing methods and sophisticated bioinformatics tools. We gained valuable insights about the cellular and molecular mechanisms behind keratoconus, bacterial/fungal, viral and healthy corneal samples by analysing differentially expressed genes (DEGs). A greater comprehension of the host's reaction to pathogenic infections was made possible by the identification of important pathways and ontologies by functional enrichment analysis of the detected DEGs. The identification of possible biomarkers and therapeutic targets for additional study and clinical applications in corneal disorders is made possible by these findings.

In order to gain a deeper understanding of the pathways and ontologies involved in infectious keratitis, this project intends to broaden its focus by concentrating on the functional annotation and visualisation of DEGs. The development of a reusable analysis workflow for efficient biomarker identification and the identification of highly significant DEGs with great biological relevance will focus efforts on biomarker discovery. Potential therapeutic targets will also be found by examining enriched pathways and hub genes found using Enrichr and heatmaps. Further, our focus will be on precision immunotherapy techniques that target immune-cell-specific responses, such as T-cells in viral keratitis and neutrophils in bacterial/fungal keratitis, to enhance disease management and treatment results.

**References:**

Mycotic keratitis: epidemiology, diagnosis and management P. A. Thomas and J. Kaliamurthy https://pubmed.ncbi.nlm.nih.gov/23398543/

INSIGHTS INTO HOST RESPONSES AGAINST PATHOGENS FROM TRANSCRIPTIONAL PROFILING Richard G. Jenner\* and Richard A. Young\*‡§ https://www.researchgate.net/publication/7930281\_Jenner\_RG\_Young\_RA\_Insights\_into\_host\_responses\_against\_pathogens\_from\_transcriptional\_profiling\_Nat\_Rev\_Microbiol\_3\_281-294

Transcriptional profiling specifies the pathogen-specific human host response to infectious keratitis Thabo Lapp,;Thabo Lapp1,2\*†Paola Kammrath Betancor;Paola Kammrath Betancor1†Günther SchlunckGünther Schlunck1Claudia Auw-HdrichClaudia Auw-Hädrich1Philip MaierPhilip Maier1Clemens LangeClemens Lange2Thomas ReinhardThomas Reinhard1Julian Wolf,,\*Julian Wolf1,3,4\* https://www.frontiersin.org/journals/cellular-and-infection-microbiology/articles/10.3389/fcimb.2023.1285676/full

Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data Franck Rapaport1, Raya Khanin1, Yupu Liang1,Mono Pirun1,Azra Krek1,Paul Zumbo2,3, Christopher E Mason2,Nicholas D Socci1 and Doron Betel3,4\* https://genomebiology.biomedcentral.com/articles/10.1186/gb-2013-14-9-r95

Transcriptome Analysis of the Gene Expression Profiles Associated with Fungal Keratitis in Mice Based on RNA-Seq Qing Zhang,1 Jian Zhang,2 Mengting Gong,2 Ruolan Pan,2 Yanchang Liu,3 Liming Tao,1 Kan He https://pubmed.ncbi.nlm.nih.gov/32539135/

“Molecular Mechanisms Underlying the Corneal EndothelialPump” Joseph A. Bonanno https://pubmed.ncbi.nlm.nih.gov/21693119/

Transcriptional profiling to identify the key genes and pathways of pterygium Yihui Chen1, Haoyu Wang2, Yaping Jiang1, Xiaoyan Zhang3, Qingzhong Wang4 https://www.researchgate.net/publication/341133394\_Transcriptional\_profiling\_to\_identify\_the\_key\_genes\_and\_pathways\_of\_pterygiumr