Differential Expression Analysis on Septic Shock and Cardiogenic Shock

Shreyoshi Ghosh Chandana Prakash

Background

Shock can be defined as the clinical expression of circulatory failure that results in inadequate cellular oxygen utilization. The most common form of shock in ICUs is septic shock, caused by sepsis, which is a life threatening condition in response to tissue damage from an infection (Vincent JL, 2014). Septic shock (SS) happens when a patient with sepsis has a large drop in blood pressure and that results in severe organ problems and sometimes death, with a mortality rate of about 40% (Mayo Clinic). Cardiogenic shock (CS) occurs when the heart cannot pump enough blood for the body to survive and is usually caused by a severe heart attack. CS is typically deadly if not treated immediately, however when treated, there is only about a 50% survival rate (Mayo Clinic). While the molecular alterations in SS patients have been studied extensively, there is very limited information about the molecular factors that impact CS (Braga D et al., 2019). The goal for our project is to look at differentially expressed genes in both SS and CS patients in order to comprehend the molecular factors by understanding the GO terms and pathways involved. We will assess the transcriptome in the whole blood samples of the patients using RNA sequencing.

Data

The data we will be using comes from a paper which highlights aspects of transcriptomic response to cardiogenic and septic shock (Braga D et al., 2019; DOI: 10.1186/s13054-019-2670-8 GEO: GSE131411). The study was part of a multicenter prospective study called ShockOmics, in which patients were gathered from the ICUs of hospitals in Geneva, Switzerland and Brussels, Belgium. Clinical data with factors such as age and sex were collected along with blood samples at three time points. Blood was collected from each patient at T1 - within 16 hours of ICU admission, T2 - on day 2, and T3 - either on day 7 or before discharge from the ICU. This data set has two factors: shock and time. The shock factor has two levels: septic shock and cardiogenic shock, while the time factor has three levels: T1, T2 and T3. There is a biological replicate of the same patient giving samples at three different times (Braga D et al., 2019). The study has 21 SS patients and 11 CS patients, while we will subset 6

male patients with no mortality from both the septic and cardiogenic shock patients using the sex and mortality factors.

Methods

Raw reads are obtained as SRA data. We will run a quality check using FastQC to check sequence quality, GC content, PCR duplicates/artifacts, etc. Then we will trim the sequences using Trimmomatic to improve sequence mappability. The trimmed, high-quality reads will then be aligned against the human genome (GRCh38) using Salmon (Patro R et al., 2017). After obtaining the count matrix of mapped reads, we plan to use the edgeR package to study differentially expressed genes (DEG) over time in CS and SS patients separately with a paired-end analysis. To classify DEGs with similar patterns, we will perform Hierarchical clustering. We then plan on performing Gene Set Enrichment Analysis (GSEA) (Subramanin A et al., 2005) to identify molecular factors, and enriched biological processes and transcription factors by looking at the associated Gene Ontology (GO) terms. The flowchart for our process can be seen in Figure 1.

Obstacles

An obstacle we foresee is high variability between patients due to the fact that both septic and cardiogenic shock are heterogeneous conditions. This is due to the different etiology, or the various causes of the sepsis amongst the patients, causing difficulty in identifying phenotype subtypes. Along with this, we have a relatively small sample size. This causes difficulty in the detection of small yet important gene expression changes. The small groups of patients makes it challenging to identify different phenotypic groups (Braga D et al., 2019).

Figures

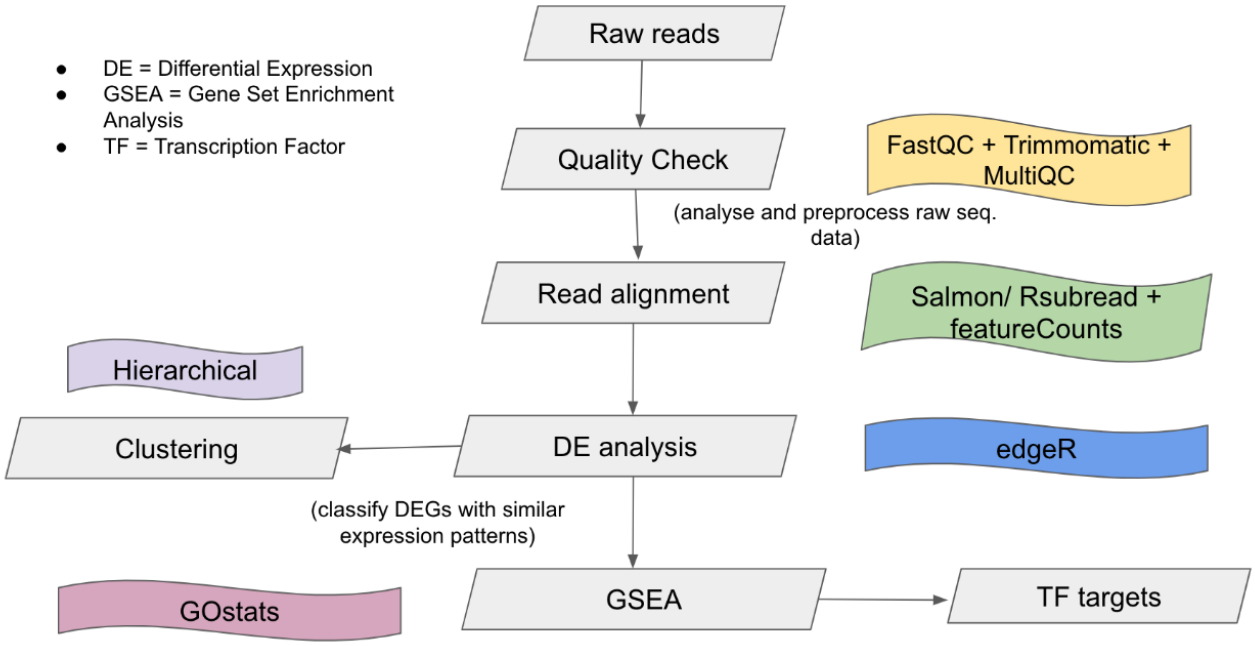

Figure 1: Workflow with tools

References

Braga D, Barcella M, Herpain A, Aletti F, Kistler EB, Bollen Pinto B, Bendjelid K, Barlassina C. A longitudinal study highlights shared aspects of the transcriptomic response to cardiogenic and Septic shock [Online]. *Critical care (London, England)* BioMed Central: 2019. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6921511/ [13 Apr. 2022].

Cardiogenic shock [Online]. *Mayo Clinic* Mayo Foundation for Medical Education and Research: 2021. https://www.mayoclinic.org/diseases-conditions/cardiogenic-shock/symptoms-causes/syc-20366739#:~:text=Cardiogenic%20shock%20is%20a%20life,Cardiogenic%20shock%20is%20rare. [13 Apr. 2022].

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nature Methods.

Sepsis [Online]. *Mayo Clinic* Mayo Foundation for Medical Education and Research: 2021.https://www.mayoclinic.org/diseases-conditions/sepsis/symptoms-causes/syc-20351214 [13 Apr. 2022].

Subramanian A, Tamayo P, Mootha V, Mesirov J. Gene set enrichment analysis: A knowledge-based approach for ... - PNAS [Online]. *PNAS*: 2005. https://www.pnas.org/content/pnas/102/43/15545.full.pdf?with-ds=yes [13 Apr. 2022].

Vincent JL. Circulatory shock [Online]. *American College of Cardiology*: 2014. https://www.acc.org/latest-in-cardiology/journal-scans/2014/01/17/13/38/circulatory-shock [13 Apr. 2022].