

Transcriptomic Profiling Identifies Differential Gene Expression Signatures Between Inflammatory And Noninflammatory Aortic Aneurysms

Submitted in Partial Fulfilment of the Requirement for the Award of
Degree of MSc. Computational Biology in NGS Data Analytics

by

NIGHITHA T. N.

Reg No: 97223606004

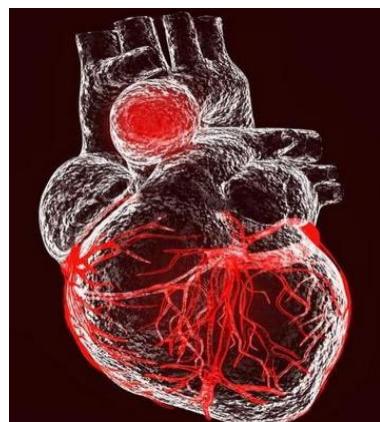
MSc. Computational Biology 2023-2025 Batch

July 2025

Under the guidance of

Dr. Anuraj Nayarisseri, Principal Scientist and Director,

Eminent Biosciences Private Limited, Indore, Madhya Pradesh



Department of Computational Biology & Bioinformatics

University of Kerala, Thiruvananthapuram, Kerala, India – 695581

Ref: EMBS1930631/25

Date: 16 July 2025

CERTIFICATE

This is to certify that this report entitled "**Transcriptomic Profiling Identifies Differential Gene Expression Signatures Between Inflammatory and Noninflammatory Aortic Aneurysms**" is a bonafide record of the project work done by **Nighitha T. N.** under my supervision and guidance, during the period (February 27, 2025 to June 30, 2025) towards the partial fulfillment of the requirement for the award of the Degree of MSc. In Computational Biology in NGS Data Analytics of the University of Kerala.



Dr. Anuraj Nayarisseri
Principal Scientist & Director,
Eminent Biosciences.





DEPARTMENT OF COMPUTATIONAL BIOLOGY & BIOINFORMATICS
UNIVERSITY OF KERALA

Re-accredited by NAAC with 'A++' Grade
Email: dcb@keralauniversity.ac.in



Prof. Dr. A. Helen, Head (i/c)
Email: helenabios@keralauniversity.ac.in
Mob: 9447399759

Kariavattom North Campus,
Thiruvananthapuram,
Kerala, India- 695581

17/07/2025

CERTIFICATE

This is to certify that this report entitled "**Transcriptomic Profiling Identifies Differential Gene Expression Signatures Between Inflammatory and Noninflammatory Aortic Aneurysms**" is a *bona fide* record of the project work done by **Ms. Nighitha T. N.**, (Reg. No. 97223606004) under my supervision and guidance, during the period 27th February to 30th June, 2025, towards the partial fulfilment of the requirement for the award of the Degree of MSc Computational Biology with specialization in NGS Data Analytics of the University of Kerala.

Prof. Dr. A. Helen
Head of the Department

Dr. Sunitha P.
Internal guide

DECLARATION

I hereby declare that the dissertation entitled "Transcriptomic Profiling Identifies Differential Gene Expression Signatures Between Inflammatory and Noninflammatory Aortic Aneurysms", submitted to the University of Kerala in partial fulfilment of the requirements for the award of the Degree of MSc. Computational Biology in NGS Data Analytics is a comprehensive record of the work I carried out under the supervision of Dr. Anuraj Nayarisseri, Principal Scientist and Director, Eminent Biosciences Private Limited.

I further declare that this thesis has not been submitted, either in part or in full, for the award of any degree, diploma, associateship, fellowship, or any other similar title of any other university or institution.

Nighitha T.N.

Date:

Place: Thiruvananthapuram

ACKNOWLEDGMENT

My appreciation to everyone who supported me and believed that I could complete this journey.

First and foremost, I would like to express my heartfelt gratitude to Dr. Anuraj Nayarisseri, my external guide, for his invaluable guidance, unwavering support, and generous encouragement throughout the course of my dissertation. It has been a privilege to work under his mentorship, and I am sincerely thankful for the opportunity.

I am equally grateful to my internal mentors, Dr. Sunitha P., Assistant Professor, and Ms. Arya V.V., for their continuous support, timely suggestions, and academic advice which significantly enriched the quality of my work. My heartfelt appreciation also goes to Dr. Shanitha A., Assistant Professor and Class Coordinator, for her constant guidance, encouragement and motivating discussions throughout my academic journey. I would also like to extend my sincere thanks to Professor Helen Antony, Head of the Department of Computational Biology and Bioinformatics, University of Kerala, and the previous professors for their immense support and for providing a conducive academic environment during the course of my study.

A special thanks to my classmates, lab mates, and a few seniors whose companionship and collaboration made this academic journey enjoyable and memorable. I would also like to acknowledge the non-teaching staff and library personnel for their assistance and support throughout my time at the institution.

Finally, I am deeply indebted to my parents for their unwavering love, prayers, and motivation. Their trust in my abilities and moral support kept me going even during the toughest moments of my academic life.

To all who have walked with me on this path—thank you

Abstract

Aortic Aneurysms represent a life-threatening vascular condition characterised by abnormal dilation of the aortic wall, which can occur with or without inflammation. While both inflammatory and noninflammatory aneurysms share overlapping clinical features, their molecular mechanisms remain poorly understood. This study aims to delineate transcriptomic differences between inflammatory and noninflammatory aortic aneurysms through comparative RNA sequencing analysis, thereby identifying distinct gene expression signatures associated with each subtype. Utilising datasets from SRA, I performed RNA-seq analysis using the Tuxedo II pipeline to identify differentially expressed genes, followed by functional enrichment analysis. Whole-transcriptome RNAseq data from surgically resected aneurysmal tissues (10 inflammatory, 10 noninflammatory) were processed using standard pipelines including HISAT2 aligner, featureCounts and Ballgown for differential gene expression analysis, followed by functional annotation and pathway enrichment using DAVID. Identified a panel of differentially expressed genes (DEGs) distinguishing inflammatory from noninflammatory aneurysms. I identified a total of 147 differentially expressed genes (DEGs) between the two groups, with 126 genes upregulated and 21 downregulated in inflammatory aneurysms compared to noninflammatory ones. Functional enrichment revealed upregulation of immune response and cytokine signaling pathways in inflammatory aneurysms, while noninflammatory aneurysms were enriched for genes involved in vascular structural integrity and metabolic processes. Enrichment analyses revealed activation of NF-κB signaling, cytokine–cytokine receptor interaction, and leukocyte migration pathways in inflammatory aneurysms, whereas noninflammatory samples showed enrichment in oxidative phosphorylation and metabolic pathways. These findings suggest that inflammatory and noninflammatory aortic aneurysms are driven by distinct transcriptomic programs, providing potential molecular targets for subtype specific diagnostics and therapeutic strategies. This transcriptomic distinction could aid in improving diagnostic precision and informing subtype specific therapeutic approaches. Future research should focus on validating these biomarkers in larger cohorts and exploring their utility in targeted therapies and early non-invasive diagnosis of aneurysm subtypes.

Contents

Chapter 1-

Introduction.....	1
1.1 Opening Remarks.....	1
1.2 Scope and objective of the work.....	3
1.3 Thesis organization.....	4

Chapter 2-

Background of the work.....	7
2.1 Definition of the problem.....	7
2.2 Basic biology related to the problem.....	8
2.2.1 Aortic Structure and Function.....	8
2.2.2 Aortic Aneurysm Pathophysiology.....	8
2.2.3 Genetic and Molecular Contributors.....	8
2.3 Computational Approaches/Resources Relevant to the Problem.....	8
2.3.1 Sequence Read Archive (SRA).....	9
2.3.1.1 Data Retrieval.....	9
2.3.2 Total RNA Sequencing.....	10
2.3.2.1 Steps involved.....	10
2.3.2.2 RNA-seq Data Analysis Workflow.....	10
2.3.3 Differential Expression Analysis.....	11
2.3.3.1 R packages for DGE Analysis.....	12
2.3.4 Annotation Tools.....	12
2.3.4.1 DAVID.....	12
2.3.5 Public Databases and Resources.....	13
2.4 National and International Status/Impact of the Problem.....	14
2.4.1 Global Health Perspective.....	14
2.4.2 National Status (India).....	14

2.5 Challenges and Future Prospects.....	14
2.5.1 Challenges in Clinical and Molecular Diagnosis.....	14
2.5.2 Future Prospects.....	15
Chapter 3-	
Review of Literature.....	17
Chapter 4-	
Materials and methodology.....	23
4.1.1 Wet Lab Procedures.....	24
4.1.2 Data Selection.....	24
4.1.2.1 SRP432690.....	25
4.2 RNA-Seq Data Analysis.....	27
4.2.1 Quality control and Preprocessing.....	27
4.2.2 Alignment to the Reference Genome using HISAT2.....	28
4.2.3 Transcript assembly and quantification.....	29
4.3 Differential expression analysis.....	30
4.3.1 Ballgown.....	30
4.3.2 Libraries used.....	31
4.4 Functional Annotation and Enrichment Analysis.....	31
4.4 Diagrammatic Workflow.....	32
Chapter 5-	
Results and Discussion.....	33
5.1 Quality Control of RNA-Seq Data.....	33
5.2 Alignment Statistics.....	34
5.3 Transcript Assembly and Quantification.....	34
5.4 Differential Gene Expression Analysis.....	34
5.5 Plots of differential expression analysis.....	35
5.5.1 Boxplot of Raw Normalized Data ($\log_2(FPKM+1)$).....	35

5.5.2 Boxplot of Filtered Normalized Data ($\log_2(\text{FPKM}+1)$).....	36
5.5.3 Expression Boxplot of the 20th Gene.....	37
5.5.4 Differential Expression Category Bar Plot.....	38
5.5.5 Heatmap of All Differentially Expressed Genes (DEGs).....	39
5.5.6 Heatmap of Top 50 Differentially Expressed Genes.....	40
5.5.7 Volcano Plot of Differentially Expressed Transcripts (DETs).....	41
5.5.8 Volcano Plot of Differentially Expressed Genes (DEGs).....	42
5.6 Functional Enrichment Analysis of DEGs using DAVID.....	43
5.7 Biological Interpretation of differentially expressed genes.....	52
5.7.1 Upregulated genes in inflammatory Aneurysms.....	54
5.7.2 Downregulated genes and suppressed protective mechanisms.....	54
5.8 Conclusion.....	54
Chapter 6-	
Conclusion & Future direction.....	56
References.....	59
Antiplagiarism Statement.....	63

List of figures

Figure 1.1 : Normal aorta and aneurysm.....	11
Figure 2.1 : Homepage of SRA.....	17
Figure 2.2 : RNA-Seq Workflow.....	19
Figure 2.3 : DAVID page for annotation.....	21
Figure 4.1 : RNA-Seq Workflow.....	22
Figure 4.2 : Data set.....	28
Figure 4.3 : List of downloaded samples.....	29
Figure 4.4 : Samples downloading using ‘prefetch’ command.....	30
Figure 4.5 : Splitting files using ‘fastq-dump’	33
Figure 4.6 : Code for FastQC.....	35
Figure 4.7 : Code for Trimming.....	36
Figure 4.8 : FastQC Report.....	36
Figure 4.9 : Code for Alignment.....	37
Figure 4.10 : SAM to BAM.....	37
Figure 4.11 : GTF from BAM.....	38
Figure 4.12 : Code for merging GTF files.....	38
Figure 4.13 : Diagrammatic Workflow.....	39
Figure 5.1 : Boxplot of $\log_2(\text{FPKM}+1)$ values for raw normalized data across all samples.....	39
Figure 5.2 : Boxplot of $\log_2(\text{FPKM}+1)$ values after transcript filtering and normalization.....	40
Figure 5.3 : Expression levels of the 20th gene across control and treatment samples.....	41
Figure 5.4 : Bar plot showing counts of upregulated, downregulated, and non-significant genes.....	42
Figure 5.5 : Heatmap displaying expression patterns of all DEGs ($\log_2 \text{FPKM} +1$, scaled).....	42
Figure 5.6 : Heatmap of the top 50 most significant DEGs ($\log_2 \text{FPKM} +1$,	

scaled).....	44
Figure 5.7 : Volcano plot showing log2 fold change and p-values for DETs, highlighting key transcripts.....	45
Figure 5.8 : Volcano plot showing log2 fold change and p-values for DEGs, highlighting key genes.....	46
Figure 5.9 : Top 30 Enriched KEGG Pathways.....	50
Figure 5.10 : Functional annotation chart (KEGG pathway enrichment).....	51

List of Tables

Table 5.2 : Highly Significant Pathways ($\text{Log}_{10}(p) > 4.5$).....	53
Table 5.3 : List of the 21 Downregulated genes.....	54
Table 5.4 : List of the 126 Upregulated genes.....	57
Table 5.5 : Selected DEGs and their biological interpretation.....	59

Abbreviations

DEGs - Differentially expressed genes

AA - Aortic aneurysm

TAA - Thoracic Aortic Aneurysm

AAA - Abdominal Aortic Aneurysm

IAAs - Inflammatory aortic aneurysms

NIAAs - Noninflammatory aortic aneurysms

KEGG - Kyoto Encyclopedia of Genes and Genomes

GO - Gene Ontology

ECM - Extracellular matrix

VSMCs - Vascular clean muscle cells

SRA - Sequence Read Archive

ncRNA - non-coding RNA

TPM - Transcripts Per Million

FPKM - Fragments Per Kilobase of transcript per Million mapped reads

TMM - Trimmed Mean of M-values

FDR - False discovery rates

CT - Computed tomography

MRI - Magnetic resonance imaging

MMPs - Matrix metalloproteinases

WGCNA - Weighted gene co-expression network evaluation

VSMCs - Vascular smooth muscle cells

AA - Arachidonic acid

cDNA - Complementary DNA

BAM - Binary Alignment/Map

SAM - Sequence Alignment/Map

BP - Biological Processes

MF - Molecular Functions

CC - Cellular Components

Chapter 1

Introduction

1.1 Opening Remarks

The aorta is the largest artery in the human body, carrying oxygenated blood from the left ventricle of the heart and delivering it to the rest of the body through a network of smaller arteries. It travels from the heart through the chest (thoracic aorta) and into the abdomen (abdominal aorta), where it branches into arteries that supply the lower body. An aortic aneurysm is a localised, abnormal bulge or dilation in the wall of the aorta due to weakening of the vessel wall. If no medical action is taken, it may enlarge and rupture, causing life-threatening internal bleeding. There are two main types: Thoracic Aortic Aneurysm (TAA), which occurs in the chest portion of the aorta and Abdominal Aortic Aneurysm (AAA), which occurs in the abdominal portion (*Hiratzka, L. F, et al., 2010*).

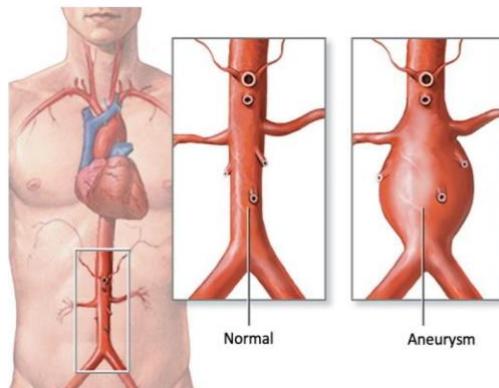


Figure 1.1: Normal aorta and aneurysm

(Source: <https://medlineplus.gov/ency/article/000162.htm>)

Aortic aneurysms are an important and potentially deadly cardiovascular condition characterised by the peculiar and permanent dilation of the aortic wall, which could lead to life-threatening complications, including dissection or rupture. The aorta, being the principal artery liable for offering oxygenated blood to the systemic circulation, plays a critical role in cardiovascular integrity; for this reason, any structural compromise poses massive medical dangers. Aortic aneurysms are typically categorised based on their anatomical area (thoracic or abdominal) and their underlying pathophysiological

mechanisms. Among those, abdominal aortic aneurysms (AAAs) are extra generic and are especially associated with aging, male gender, smoking, and atherosclerosis, at the same time as thoracic aortic aneurysms (TAAs) regularly have genetic predispositions such as Marfan syndrome or Loeys-Dietz syndrome (*Golledge, J., & Norman, P. E.*, 2011).

A vital but frequently underappreciated component of aneurysm pathology is the inflammatory nature of the lesion. Inflammatory aortic aneurysms (IAAs) are characterised using dense perivascular infiltration of immune cells, increased systemic inflammatory markers, and a thickened aneurysmal wall. In contrast, noninflammatory aortic aneurysms (NIAAs) lack such immune infiltration and might expand because of biomechanical pressure, vascular ageing, or hereditary defects in connective tissue. Despite overlapping scientific displays, those aneurysm subtypes vary fundamentally in their biological drivers, disease progression, and response to treatment. However, distinguishing between them remains difficult with the use of conventional imaging and histopathological techniques on my own (*Koo, T. K., et al.*, 2016).

In modern years, transcriptomic profiling the usage of RNA sequencing (RNA-seq) has emerged as a powerful tool to observe the molecular landscape of complex illnesses. By allowing the complete evaluation of gene expression styles, transcriptomics offers insights into the underlying cell techniques, signalling pathways, and regulatory networks that differentiate disease subtypes. In the context of aortic aneurysms, transcriptomic analysis holds the capability to discover key differences amongst inflammatory and noninflammatory forms, identify diagnostic biomarkers, and reveal novel therapeutic targets (*Biros, E., et al.*, 2015; *Pyo, R., et al.*, 2000).

The present study aims to fill this knowledge gap by performing differential gene expression analysis among inflammatory and noninflammatory aortic aneurysms using next-generation RNA sequencing. Through this method, I am looking to symbolise the distinct transcriptomic signatures that drive aneurysm pathogenesis in each subtype and contribute to more accurate molecular classification, better clinical management, and the development of targeted therapies in vascular medicine (*Wang, Z., Gerstein, M., & Snyder, M.*, 2009).

1.2 Scope and objective of the work

Aortic aneurysms constitute a severe and progressive vascular situation marked by means of the weakening and peculiar dilation of the arterial wall. While each thoracic and

abdominal aortic aneurysms make appreciable contributions to cardiovascular morbidity and mortality worldwide, their underlying etiologies remain complex and multifactorial. Among the various classifications, inflammatory and noninflammatory aneurysms are the main pathological subtypes that fluctuate not simplest in histological presentation but also in their medical route and potential response to treatment. Inflammatory aortic aneurysms are characterised by intense immune cell infiltration, adventitial thickening, and elevated inflammatory markers, while noninflammatory aneurysms progress with minimum immune involvement and are generally linked to factors together as ageing, mechanical stress, and genetic predisposition. However, distinguishing among these subtypes at the molecular level remains a challenge in clinical practice (Johnston, K. W., Rutherford, R. B., 1991).

The scope of this research is to address the gap in knowledge of the molecular distinctions between inflammatory and noninflammatory aortic aneurysms by leveraging high-throughput transcriptomic technology. With the development of next-generation sequencing (NGS), transcriptomic profiling has become an effective tool to comprehensively examine gene expression patterns across specific biological states. This has a look at the utility of RNA sequencing (RNA-seq) to aneurysmal tissue samples, aiming to capture the genome-wide transcriptional panorama that differentiates the two aneurysm subtypes. The work encompasses computational workflows involving examining alignment, quantification, differential gene expression analysis, and functional enrichment using established bioinformatics tools (Kasashima, S., Zen, Y., Kawashima, A., 2018).

The core objective of this observation is to discover differentially expressed genes (DEGs) that might be uniquely or appreciably altered among inflammatory and noninflammatory aortic aneurysms. By doing so, we aim to uncover awesome molecular signatures related to immune-mediated vascular remodelling in inflammatory aneurysms, as well as pathways related to structural degeneration in noninflammatory cases. Further, the study seeks to functionally annotate those DEGs to determine their biological relevance, particularly in processes of immune cell activation, cytokine signalling, extracellular matrix degradation, and smooth muscle cell dysfunction, all of which are essential in the development and progression of aneurysmal disease (Golledge, J., & Norman, P. E., 2011).

In addition to figuring out DEGs, this work additionally aims to map those genes onto known signalling pathways the usage of assets like Gene Ontology (GO) phrases, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, and disease enrichment databases. Through this analysis, I am seeking to gain deeper insight into how inflammation alters the vascular transcriptome in comparison to noninflammatory mechanisms. The findings from this study are expected to contribute towards the refinement of aneurysm class systems by way of incorporating molecular records, which can in the end help in advanced diagnosis, better risk stratification, and more personalised therapeutic approaches for patients with aortic aneurysms (Conesa, A., Madrigal, P., 2016).

1.3 Thesis organization

The thesis is organised into 6 chapters, preceded by several introductory pages. The introductory pages consist of the front page, certificate page, declaration, acknowledgement, abstract, table of contents, tables and figures. The main part of the thesis discusses the Background, methodology, results and discussion of the topic. The references section lists the sources that the author used in the thesis. The contents of the thesis are as follows.

A: Introduction

B: Background of the work

D: Materials and Methodology

E: Result and discussion

F: Conclusion and Future Direction

A: Introduction

The introductory chapter sets the stage for the entire study's work by providing the context and significance of studying aortic aneurysms. It begins by describing the anatomy and function of the aorta, mechanisms leading to aneurysm formation, and the role of inflammation in vascular remodelling, epidemiological burden and scientific demanding situations related to aortic aneurysms, emphasising their silent development and danger of rupture. The chapter then introduces the concept of aneurysm heterogeneity, in particular, the difference between inflammatory and noninflammatory

subtypes and the constraints of current diagnostic tools as it should be differentiating these forms. It further introduces transcriptomic profiling as a promising molecular method to research the underlying differences among these subtypes. The chapter concludes with a clear articulation of the research troubles, the rationale behind the study, the targets, and a brief outline of the methodology followed.

B: Background of the work

This chapter provides an in-depth discussion of the biological and clinical foundation necessary to understand the study. Special emphasis is placed on the pathological and histological features of inflammatory and noninflammatory aneurysms, supported by clinical markers and imaging findings. Furthermore, the chapter introduces the concept of transcriptomics, RNA sequencing technology, and its relevance in vascular disease research. It also discusses how gene expression profiles can provide insight into cellular pathways and molecular mechanisms contributing to disease phenotypes.

C: Review of Literature

This chapter critically evaluates existing scientific literature relevant to the study. It evaluates previous research on the molecular and genetic aspects of aortic aneurysms, highlighting recognised genes and pathways implicated in aneurysm development and progression. It additionally covers beyond applications of RNA-seq and other omics technologies in cardiovascular studies, and how these strategies have contributed to identifying biomarkers and therapeutic targets. Gaps in the current knowledge, especially with recognition of comparative transcriptomic analysis among inflammatory and noninflammatory aneurysms, are recognised to justify the novelty and necessity of the present study of inflammatory and noninflammatory aneurysms, supported by clinical markers and imaging findings. Furthermore, the chapter introduces the concept of transcriptomics, RNA sequencing technology, and its relevance in vascular disease research. It also discusses how gene expression profiles can provide insight into cellular pathways and molecular mechanisms contributing to disease phenotypes.

D: Methodology / Materials and Methods

This chapter comprehensively describes the methodological framework used in the study, detailing each of the experimental and computational additives. It outlines the processes involved in wet lab like RNA extraction, library preparation, sequencing and then it details

the dry lab part from preprocessing to analysis. The sequencing technique, including platform info and examine parameters, is truly defined. The bioinformatics pipeline for data analysis is presented step-by-step, which includes read alignment (HISAT2), expression quantification, normalisation, differential gene expression analysis (Ball Gown), and statistical thresholds used for gene selection. Functional annotation and pathway enrichment analyses achieved using tools such as DAVID, KEGG, and Gene Ontology databases are also discussed. The chapter ensures reproducibility through variations of software programs, gear and databases, wherever applicable.

E: Results and Discussion

This chapter presents the core findings of the study, which include an in-depth account of differentially expressed genes (DEGs) between inflammatory and noninflammatory aortic aneurysms. Results are visualised via appropriate graphical representations such as heatmaps, volcano plots, box plots, etc. The biological relevance of the top DEGs is explored via pathway and gene ontology analyses, revealing insights into the molecular mechanisms underlying every aneurysm subtype. The discussion integrates these findings with the existing literature, highlighting similarities and variations. Novel insights are emphasised, and possible elements for observed gene expression patterns are provided. The implications of these findings for disease diagnosis, risk stratification, and future therapeutic interventions are discussed critically. Limitations of the study, such as sample size, absence of functional validation, or lack of clinical follow-up data, are acknowledged.

F: Conclusion and Future Direction

The final chapter offers a concise summary of the studies' effects, reiterating the key findings and their relevance to understanding the molecular basis of inflammatory versus noninflammatory aortic aneurysms. It displays how transcriptomic profiling has effectively exposed notable gene expression signatures that can function as potential biomarkers or therapeutic targets. The chapter also highlights how the study contributes to the broader field of vascular transcriptomics and precision medicine. Finally, it outlines future directions, suggesting validation studies in large cohorts, integration of multiomics data (e.g., proteomics, Sc-RNA-seq), and experimental work to confirm the functional roles of identified genes in aneurysm biology. The need for translating these findings into clinical diagnostic or prognostic tools is emphasised as a prolonged-time period intention.

Chapter 2

Background of the work

2.1 Definition of the Problem

Aortic aneurysms represent a significant cardiovascular health issue, marked by a localised, irreversible dilation of the aortic wall that predisposes people to catastrophic rupture if left undetected. These aneurysms are classified primarily based on their anatomical region (thoracic or abdominal) and their underlying pathology, which includes inflammatory and noninflammatory types. The accurate differentiation among these subtypes is clinically crucial because it affects prognosis, management strategies, and patient outcomes (Johnston et al., 1991). Inflammatory aortic aneurysms (IAAs), characterised with the aid of using immune cellular infiltration, adventitial fibrosis, and multiplied systemic inflammatory markers, frequently exhibit different progression patterns and therapeutic responses compared to noninflammatory aortic aneurysms (NIAAs), which can be commonly associated with structural weakening of the vessel wall without significant immune involvement (Matsumura et al., 1999).

Current diagnostic tools, including imaging and serological markers, are regularly insufficient to differentiate between those cases at an early or molecular level. This ambiguity poses challenges in affected person stratification and targeted intervention. Therefore, there may be an urgent need to locate unique molecular signatures that could distinguish IAAs from NIAAs, which can be carried out through transcriptomic profiling. RNA-seq technology offers a promising avenue to find out genome-scale gene expression differences in affected tissues, potentially revealing novel diagnostic markers, recovery desires, and insights into disease mechanisms (van de Luijtgaarden et al., 2015).

This study is situated within this context, aiming to investigate and compare the transcriptomic landscapes of inflammatory and noninflammatory aortic aneurysms to define differentially expressed genes (DEGs) and associated pathways, thereby enhancing molecular understanding and informing clinical decision-making (Wang, Gerstein & Snyder, 2009).

2.2 Basic Biology Related to the Problem

2.2.1 Aortic Structure and Function

The aorta, the most important artery within the human body, originates from the left ventricle of the heart and distributes oxygenated blood to the systemic circulation. Structurally, the aorta consists of three layers: the tunica intima (internal layer), tunica media (middle muscular layer), and tunica adventitia (outer connective tissue layer). Integrity of the aortic wall is essential for its function as a pressure conduit, and it's far maintained via extracellular matrix (ECM) components like elastin and collagen, as well as vascular smooth muscle cells (VSMCs) (Golledge, J., & Norman, P. E., 2011).

2.2.2 Aortic Aneurysm Pathophysiology

Aneurysm formation involves progressive weakening and dilation of the aortic wall. Inflammatory mechanisms, proteolytic degradation of ECM components, oxidative stress, apoptosis of VSMCs, and hemodynamic strain contribute to aneurysm pathogenesis. In IAAs, there may be marked immune cell infiltration, especially of macrophages, T lymphocytes, and plasma cells, which results in adventitial thickening and fibrosis. Cytokines, together with IL-6, TNF- α , and IFN- γ , play pivotal roles in mediating inflammation and matrix degradation. In contrast, NIAAs typically result from mechanical and genetic factors, with a reduced inflammatory component but enhanced ECM degradation (Bianconi, E., et al., 2019).

2.2.3 Genetic and Molecular Contributors

Numerous genes have been implicated in aortic aneurysm formation, including MMPs (matrix metalloproteinases), TIMPs (tissue inhibitors of MMPs), TGF- β signalling genes, and genes involved in vascular remodelling. The differential expression of these genes, in particular those related to inflammation, apoptosis, and fibrosis, offers a molecular basis for distinguishing among IAAs and NIAAs. Deep information of those biological processes and gene regulatory networks is essential for effective classification and intervention (He, R., Guo, D. C., Sun, 2008).

2.3 Computational Approaches/Resources Relevant to the Problem

The field of transcriptomics, powered through next-generation sequencing (NGS), has revolutionised the capability to analyse gene expression at an unprecedented scale and

resolution. In this study, computational techniques play a central role in identifying transcriptomic differences among IAA and NIAA tissue samples.

2.3.1 Sequence Read Archive (SRA)

The Sequence Read Archive (SRA) is a publicly accessible database maintained by the NCBI that stores raw next-generation sequencing (NGS) data from various studies worldwide. It contains data from diverse experiments such as RNA-Seq, DNA-Seq, metagenomics, and more, allowing researchers to reuse and analyse sequencing data for secondary research. Users can access, download, and process this data using tools like the SRA Toolkit or through the NCBI web interface.

2.3.1.1 Data Retrieval

SRA data retrieval involves accessing raw next-generation sequencing data, such as RNASeq reads, from the Sequence Read Archive (SRA). This data is essential for analyses like transcript quantification, variant calling, and differential expression studies. Researchers can locate relevant datasets using SRA accession numbers (e.g., SRR, SRX, or PRJNA) via the NCBI SRA Run Selector or search tools. Once identified, raw data can be downloaded using the SRA Toolkit through commands like prefetch for retrieval and fastq-dump to convert .sra files into FASTQ format. These FASTQ files can then be processed through bioinformatics pipelines to generate gene-level or transcript-level count data for further investigation.



Figure 2.1: Homepage of SRA

(source: <https://www.ncbi.nlm.nih.gov/sra>)

2.3.2 Total RNA Sequencing

Total RNA sequencing (Total RNA-Seq) is a subset of RNA sequencing (RNA-Seq) that aims to capture a full snapshot of all RNA molecules found in a biological sample, including mRNA, non-coding RNA (ncRNA), and fragmented RNA species. Unlike mRNA sequencing (mRNA-Seq), which focuses largely on protein-coding transcripts, Total RNASeq gives a more comprehensive picture of the transcriptome, revealing information about diverse RNA populations within cells or tissues.

2.3.2.1 Steps involved

RNA Extraction: Total RNA can be extracted from biological samples, which include biofluids, cells or tissues.

Library preparation: It involves converting RNA to cDNA using reverse transcriptase, fragmentation of cDNA and ligating adapters for sequencing.

Sequencing: To generate millions of reads from cDNA, high-throughput sequencing like Illumina is performed. These reads capture both coding and non-coding RNA molecules.

2.3.2.2 RNA-seq Data Analysis Workflow

The RNA-seq pipeline includes a couple of steps:

- **Quality Control (QC):** Tools like FastQC and Trim Galore make certain high-quality reads using trimming adapters and filtering low-nice sequences.
- **Read Alignment:** Spliced aligners like HISAT2 map reads to the reference genome (GRCh38).
- **Transcript Assembly /quantification:** Generate counts. feature Counts, stringtie, etc, are used.
- **Normalization:** Methods that include TPM, FPKM, normalisation account for sequencing depth and gene period biases.
- **Differential expression analysis:** It is a method used to identify differentially expressed genes from the count data. R packages like Ballgown, EdgeR, etc, are used.

- **Pathway and Functional Enrichment:** Tools like DAVID pick out overrepresented biological techniques, pathways (KEGG), and molecular capabilities (Gene Ontology classes).

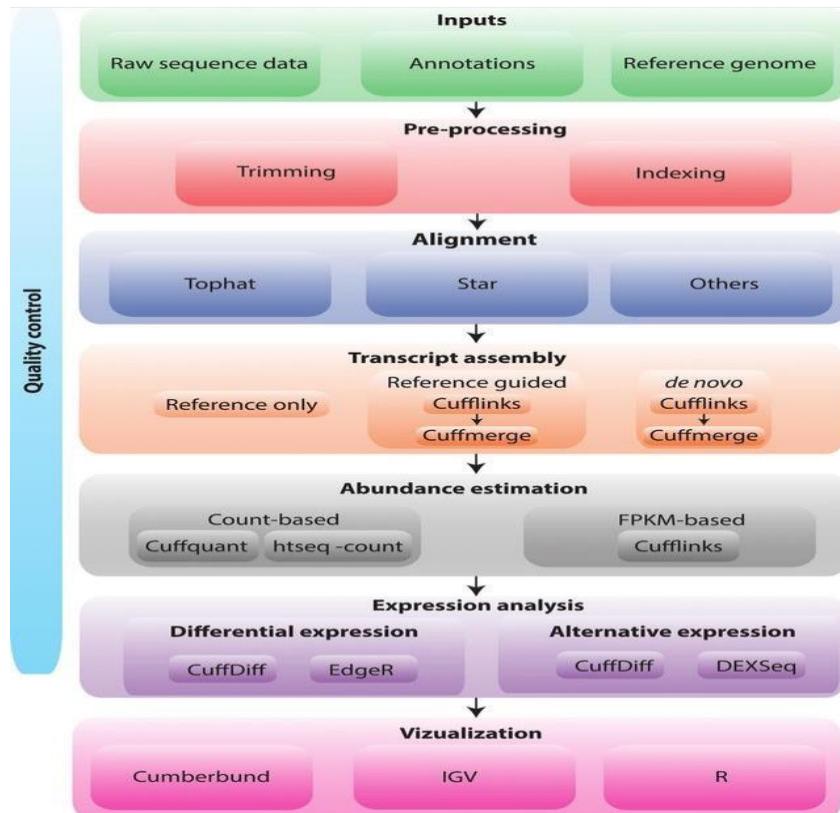


Figure 2.2: RNA-Seq Workflow

(Source:<https://www.researchgate.net/figure/RNA-seq-analysis-flow-chart-Anexample->
[RNA-seq-analysis-workflow-is-depicted-for-a fig4 280870031](https://www.researchgate.net/figure/RNA-seq-analysis-workflow-is-depicted-for-a-f))

2.3.3 Differential Expression Analysis

Differential expression analysis is a part of the RNaseq method where genes are differentially expressed when compared to normal or control, or treated conditions vs disease. Genes can be downregulated or upregulated. The main role is to identify relevant genes. Different R packages like Ballgown are available.

2.3.3.1 R packages for DGE Analysis

R is a programming language for statistical computing and graphics. It is widely used for data analysis, visualization, and modelling. It offers many packages for effectively analysing large data. For DGE analysis, R provides packages like EdgeR, Ballgown, etc.

EdgeR: EdgeR is an extensively used R package from the Bioconductor undertaking that performs differential gene expression (DGE) analysis using raw read records from RNA-Seq experiments. It models study counts the use of the negative binomial distribution, which is properly appropriate to account for organic variability and overdispersion in sequencing records. EdgeR is especially effective for experiments with small numbers of replicates, and it consists of strong normalization methods, inclusive of Trimmed Mean of M-values (TMM). The bundle helps with easy and complex experimental designs and allows customers to conduct statistical exams through actual checks or generalised linear models (GLMs). Its flexibility and statistical rigour make it a popular choice for gene-level expression evaluation.

Ballgown: In contrast, Ballgown is an R package specifically designed for transcript-level differential expression analysis, mainly when working with transcript assemblies generated by using tools like StringTie. Ballgown generally uses transcript abundance estimates together with FPKM or TPM, supplied in .ctab files. It employs linear modelling strategies to pick out differentially expressed genes or transcripts and gives visualisation tools to discover transcript systems and expression levels. Ballgown is especially useful when the research objective is to take a look at alternative splicing or isoform-level expression changes, making it an ideal tool for studies that require a detailed view of transcriptome dynamics beyond gene-level summaries.

2.3.4 Annotation Tools

2.3.4.1 DAVID

DAVID (Database for Annotation, Visualisation and Integrated Discovery) is a widely used bioinformatics resource designed for useful annotation and enrichment analysis of large gene or protein lists derived from high-throughput experiments, together with RNA-Seq, microarrays, or proteomics. This tool enables researchers to interpret the biological significance of genes by identifying enriched pathways, diseases, and molecular processes. DAVID integrates information from several public databases, including KEGG,

Reactome, UniProt, InterPro, and the Gene Ontology (GO), allowing comprehensive annotation in terms of biological strategies, cell additives, molecular functions, protein domain names, and biochemical pathways. It supports practical clustering, reduces redundancy in GO terms, and provides statistically enriched biological themes through metrics such as p-values and Benjamini-corrected false discovery rates (FDR). The platform offers detailed reports that help identify key pathways or gene groups involved in specific biological conditions. Overall, DAVID plays a critical role in enabling researchers to extract meaningful biological insights from complex genomic and transcriptomic data, thus bridging raw gene lists with functional interpretation.

Figure 2.3: DAVID page for annotation

(Source: <https://davidbioinformatics.nih.gov/summary.jsp>)

2.3.5 Public Databases and Resources

This study may additionally make use of:

- Ensembl, NCBI Gene, UCSC Genome Browser: For genomic annotations.
- SRA: For getting access to transcriptomic datasets.
- STRING, BioGRID: For gene interaction networks.

- Reactome, KEGG, Gene Ontology Consortium: For pathway-based total evaluation.

These computational procedures allow a structure-level research of ailment biology, going beyond individual genes to understand broader cellular mechanisms.

2.4 National and International Status/Impact of the Problem

2.4.1 Global Health Perspective

Globally, aortic aneurysms make a considerable contribution to cardiovascular mortality, in particular in ageing populations. Abdominal aortic aneurysm (AAA) is one of the top 20 leading causes of death in advanced nations. The majority of aneurysm-related deaths occur due to sudden rupture, which carries a mortality rate exceeding 80% if not treated emergently. Despite improvements in imaging and surgical techniques, the incidence of aneurysms continues to rise, partly due to lifestyle factors and increased longevity.

Inflammatory aortic aneurysms, although less common than their noninflammatory counterparts, are associated with higher surgical complexity and increased perioperative morbidity. Their differentiation from other retroperitoneal inflammatory diseases is often challenging (Howard, D. P., Banerjee, A., 2015).

2.4.2 National Status (India)

In India, the load of aortic aneurysms is underreported, largely because of a lack of widespread screening programs and limited access to advanced imaging in rural areas. However, hospital-based studies and case reports indicate a rising trend in aneurysm detection, especially with improved access to CT angiography. The scarcity of molecular diagnostic tools and the absence of population-based molecular profiling studies in the Indian context further underscore the importance of transcriptomics in improving aneurysm diagnosis and classification in India (Venkatesh, P.K., & Jayaraman, 2018).

2.5 Challenges and Future Prospects

2.5.1 Challenges in Clinical and Molecular Diagnosis

Despite advances in vascular imaging and surgical interventions, numerous challenges persist within the clinical and molecular diagnosis of aortic aneurysms, mainly when trying to differentiate among inflammatory and noninflammatory subtypes. One of the major limitations is the absence of reliable molecular markers that can accurately classify

aneurysm types at a preclinical stage. Current diagnostic approaches are predominantly reliant on imaging modalities, together with computed tomography (CT) and magnetic resonance imaging (MRI), together with histopathological evaluation post-surgery. While these techniques are fundamental, they often fail to capture subtle molecular and cellular differences between aneurysm subtypes, particularly in the early stages of disease development.

Another critical mission is the limited accessibility of aneurysmal tissue samples for molecular analysis. Aneurysm resections are generally performed only when the aneurysm reaches a critical size or presents symptoms of imminent rupture, which restricts possibilities to study early molecular events. This problem extensively hampers the ability to conduct large-scale transcriptomic research, as obtaining representative tissue samples before surgical intervention remains difficult.

Moreover, the inherent biological heterogeneity of aortic aneurysms poses a significant obstacle to both scientific control and research. Inter-individual variability in genetic predisposition, immune reaction, and disease progression complicates the translation of transcriptomic records. Even within a single aneurysm, spatial variability in cellular composition and gene expression can influence findings (Kuivaniemi, H., Ryer, 2015).

2.5.2 Future Prospects

Despite those challenges, the field holds gigantic promise, with numerous interesting future directions arising from ongoing advances in molecular biology, bioinformatics, and systems medicinal drugs. One of the most immediate prospects is the improvement of blood-based, noninvasive biomarkers derived from differentially expressed genes recognized through transcriptomic profiling. Such biomarkers could aid in the early detection and subtype differentiation of aortic aneurysms, decreasing reliance on invasive tissue sampling or late-stage imaging diagnoses.

Another promising road lies in the exploration of targeted therapeutic strategies. Deeper information of the gene expression signatures particular to inflammatory aneurysm in particular those concerning immune pathways, extracellular matrix degradation, or fibrotic remodelling, could guide the development of drugs that specifically inhibit pathological techniques in this subtype. Furthermore, genes and pathways discovered through RNA-seq analysis can also be candidates for drug repurposing, offering a faster route to clinical application.

Looking ahead, the integration of transcriptomic information with other omics platforms, such as proteomics, metabolomics, and single-cell RNA sequencing, will provide a more complete picture of aneurysm biology. This systems-level method can discover dynamic interactions between genes, proteins, metabolites, and cell phenotypes, thereby improving our understanding of disease progression and response to therapy.

As precision medicine continues to evolve, molecular types of aortic aneurysms can transform current management paradigms. Stratifying patients based on molecular subtype could enable personalized monitoring schedules, risk prediction models, and individualized treatment regimens. Such a method might enhance medical results even as minimising unnecessary interventions.

Finally, the incorporation of artificial intelligence and machine learning techniques gives an effective toolset for extracting clinically relevant styles from complicated transcriptomic and imaging datasets. Predictive models skilled on large, multidimensional datasets could aid early prognosis, subtype type, and rupture threat estimation with high accuracy. These computational frameworks may ultimately become integral to the clinical decision-making process, bridging the gap between bench research and bedside application (Nienaber, C. A., & Clough, R. E., 2015), (Buerkle, M. A., Minc, A., 2022).

Chapter 3

Review of Literature

Aortic aneurysms, specifically those associated with inflammatory procedures, represent a crucial and life-threatening cardiovascular condition that contributes drastically to international morbidity and mortality. A developing frame of literature has explored the transcriptomic landscape of aortic aneurysms, highlighting the roles of immune mobile infiltration, extracellular matrix reworking, cytokine signalling, and vascular mobile disorder. Many studies have been conducted and continue to be carried out on aortic aneurysms to understand the risk factors and biology behind them. A number of research papers are used in this study to get a better idea of the disease and confirm the results obtained in this study.

The study titled "*Identifying Nexilin as a Central Gene in Neutrophil-Driven Abdominal Aortic Aneurysm Pathogenesis*" provides important insights into the inflammatory mechanisms underlying abdominal aortic aneurysm (AAA) improvement. The authors are interested in the characteristics of neutrophil infiltration in aneurysmal tissues and use differential gene expression (DEG) assessment in combination with Weighted Gene Co-expression Network Analysis (WGCNA) to identify key molecular drivers. One of the number one findings turned into the identification of Nexilin (NEXN) as a hub gene that is appreciably dysregulated in AAA samples. The study demonstrates that NEXN is closely associated with neutrophil-related immune pathways and can contribute to extracellular matrix destabilisation and vascular wall contamination. By validating their transcriptomic findings with functional enrichment and community-based techniques, the authors improve the case for NEXN as a potential biomarker or therapeutic target in AAA. This study offers a strong parallel to the modern research, which moreover explores inflammation-associated transcriptional adjustments in aortic aneurysms and emphasises the role of immune-mediated gene signatures in distinguishing among inflammatory and non-inflammatory disease phenotypes.

The paper by Liu et al. (2020), titled "*Identification of Key Genes and Pathways in Abdominal Aortic Aneurysm by Integrated Bioinformatics Analysis*", gives a comprehensive systems biology technique to decipher the molecular mechanisms underlying abdominal aortic aneurysm (AAA). The study integrates multiple GEO microarray datasets to pick out always dysregulated across independent AAA cohorts. Through differential gene

expression evaluation and pathway enrichment, the authors reveal that inflammation and extracellular matrix (ECM) transformation are central topics in AAA pathogenesis. Key upregulated genes included matrix metalloproteinases (MMPs), STAT1, and components of the complement cascade, highlighting immune activation and ECM degradation. Additionally, CANX (calnexin), a gene involved in protein folding and endoplasmic reticulum stress, was recognised as a significant node in the regulatory network. The study emphasises that immune cellular infiltration, cytokine signalling, and proteolytic enzyme activity collectively contribute to vascular wall weakening. These findings align closely with the present research, which investigates transcriptomic signatures associated with inflammatory versus non-inflammatory aortic aneurysms, particularly focusing on ECM regulation and immune-driven gene expression patterns.

The study titled "*Whole genome expression profiling reveals a significant role for immune function in human abdominal aortic aneurysms*" (BMC Genomics, 2007) represents one of the earliest transcriptomic investigations aimed at elucidating the molecular basis of abdominal aortic aneurysm (AAA) formation. By performing complete-genome expression profiling on aneurysmal tissues compared to non-aneurysmal controls, the authors identified a suggested immune signature in AAA pathophysiology. Notably, genes associated with numerous immune cell types, including macrophages, natural killer (NK) cells, T lymphocytes, and B cells, were significantly upregulated in AAA tissues. This upsurge in immune-related gene expression highlighted the primary function of chronic irritation and immune cell infiltration in vascular wall degradation and aneurysm enlargement. The study's findings are immediately applicable to the present thesis, which also makes a speciality of transcriptomic differences among inflammatory and noninflammatory aortic aneurysms. The 2007 study supports the view that immune mechanisms are a vital driving force of aneurysmal disease progression, thereby validating the exploration of inflammatory DEGs as a biologically significant technique in figuring out biomarkers and therapeutic goals in aneurysm subtypes.

The study "*Transcriptional profiling and network analysis of the murine angiotensin II-induced abdominal aortic aneurysm*" by using Spin et al. (2011) employed a time-series transcriptomic approach to characterise gene expression dynamics in a properly mounted Angiotensin II (AngII)-precipitated abdominal aortic aneurysm (AAA) model using ApoE^{-/-} mice. This study stands out by no longer most effectively profiling the global gene

expression modifications throughout aneurysm development but also mapping the temporal styles of inflammatory gene activation and vascular remodelling. Key upregulated pathways covered cytokine signalling, leukocyte recruitment, and extracellular matrix degradation, hallmarks of inflammatory AAA progression. Importantly, the study used network-based total analysis to become aware of hub genes within co-expression modules, presenting insight into transcriptional regulators and potential therapeutic targets. This research gives a robust murine model parallel to human inflammatory AAA and aligns closely with the targets of the existing thesis, which focuses on identifying differentially expressed genes (DEGs) associated with inflammation in human aortic aneurysms. The technique and findings, in addition, support the relevance of transcriptomic profiling for uncovering inflammation-driven molecular signatures in aneurysmal pathogenesis.

The paper "*Involvement of Myeloid Cells and Noncoding RNA in Abdominal Aortic Aneurysm Disease*" (Antioxidants & Redox Signalling, 2020) investigates the immunopathological mechanisms driving abdominal aortic aneurysm (AAA) formation, with particular emphasis on the roles of myeloid cells and noncoding RNAs such as microRNAs. The study demonstrates how immune mediators like TNF- α and IFN- γ , along with CD4 $^+$ T-cell infiltration, make contributions to inflammatory degeneration of the aortic wall. Importantly, the authors become aware of identifying miR-155 as a key regulator linking immune signalling to vascular pathology, highlighting the significance of transcriptomic regulation at both coding and noncoding RNA levels. This integrative approach underlines the complexity of immune-driven aneurysm pathogenesis and aligns directly with the objective of the present thesis, which focuses on identifying inflammation-associated transcriptional signatures in aortic aneurysms. By combining gene expression and immune profiling, the study reinforces the importance of transcriptomic methodologies such as those applied in this project for uncovering regulatory molecules and pathways relevant to disease progression.

The study "*Gene Expression Profiling in Abdominal Aortic Aneurysms*" (Journal of Clinical Medicine, 2022) presents an in-depth analysis of the molecular landscape of AAA by focusing on the expression of matrix metalloproteinases (MMPs), specifically MMP-7, MMP-9, and MMP-12. These MMPs are implicated in extracellular matrix (ECM) degradation, which is a hallmark of aortic wall weakening and aneurysm formation. The study integrates transcriptomic profiling with functional enrichment analysis to perceive

key biological pathways, together with those associated with infection, immune activation, and ECM transformation. These findings are directly relevant to the present thesis, which also aims to delineate differential gene expression patterns among inflammatory and noninflammatory aortic aneurysms. The identification of MMP-associated upregulation aligns with the upregulated genes in this challenge and highlights the importance of ECM dysregulation in aneurysm progression, validating the biological importance of the DEGs determined via RNA-Seq evaluation in the current study.

The study "*Identification of Crucial Genes Involved in Pathogenesis of Regional Weakening of the Aortic Wall*" (Hereditas, 2021) employs weighted gene co-expression network evaluation (WGCNA) to discover key immune and extracellular matrix (ECM)-related hub genes implicated in aortic wall degeneration. Notably, the genes IL6, IL8 (CXCL8), and PTGS2 (COX-2) were identified as important regulators contributing to inflammation-driven aortic weakening. These genes are involved in leukocyte recruitment, cytokine signalling, and prostaglandin biosynthesis, techniques known to exacerbate vascular wall remodelling and aneurysm formation. This research strongly supports the current thesis, which uses transcriptomic profiling to differentiate inflammatory from noninflammatory aortic aneurysms. The consistent identification of IL6 and CXCL8 in each study reinforces the crucial position of those cytokines in the inflammatory mechanisms underlying aneurysm pathogenesis. Furthermore, the combination of WGCNA and DEG techniques on these references, the work parallels the analytical framework of the present study, further validating its methodological relevance and biological conclusions.

The study titled "*Single-cell Transcriptomic Analysis Reveals Differential Cell Subpopulations and Distinct Phenotype Transition in Normal and Dissected Ascending Aorta*" (Molecular Medicine, 2022) explores the cellular landscape of human aortic tissues at single-cell resolution. Through single-cell RNA sequencing (scRNA-seq), the authors diagnosed distinct immune and stromal cell populations and discovered dynamic phenotypic transitions related to aortic dissection. Notably, the study highlights an increase in inflammatory cell subsets, inclusive of macrophages and T-cells, in diseased tissues, together with altered expression of genes involved in ECM transforming and immune signalling. This work provides a powerful complement to bulk RNA-seq—primarily based transcriptomic studies like the present thesis, by offering cell-type-specific insights into differential gene expression. The identification of inflammation-related pathways and the prominence of immune-responsive gene signatures enhance the

biological relevance of IL6, CXCL8, and TNFAIP3, which have been highlighted in our evaluation of differentially expressed genes among inflammatory and noninflammatory aortic aneurysm groups. This underscores the importance of cell-type heterogeneity and immune involvement in aortic pathology.

The recent study "*Single-cell RNA-seq of Angiotensin II-induced Abdominal Aortic Tissue Identifies Aneurysm-Associated Cell Clusters in C57BL/6J Mice*" (Bioscience Reports, 2025) makes use of single-cell transcriptomics to dissect the cellular and molecular panorama of abdominal aortic aneurysm (AAA) development in a widely used murine model. Through clustering analysis, the authors recognised specific subpopulations of vascular smooth muscle cells (VSMCs), macrophages, and fibroblasts that go through transcriptional reprogramming in reaction to Angiotensin II infusion. Particularly, they cited an upregulation of inflammatory mediators and extracellular matrix-degrading enzymes inside immune cell clusters, suggesting their important position in aneurysm pathogenesis. These findings corroborate the differential gene expression styles found in our study, mainly the upregulation of IL6 and CXCL8 and downregulation of TNFAIP3 within the inflammatory aneurysm group. This paper supports the notion that infection and immune cellular infiltration are key drivers of disease development and that transcriptomic shifts can screen functionally distinct aneurysm-related cell kinds, aligning properly with the goals and outcomes of our bulk RNA-seq—primarily based analysis.

Another important study was published in MDPI Genes in 2024 titled "*Transcriptomic Analysis of Arachidonic Acid Pathway Genes Provides Mechanistic Insight into Multi-Organ Inflammatory and Vascular Diseases*." This meta-analysis focused on figuring out conserved transcriptional signatures associated with the arachidonic acid (AA) metabolic pathway throughout multiple inflammatory diseases, including vascular disorders. The observer found out that key AA pathway genes, which include ALOX5, PTGS2, PLA2G4A, and CYP4A11, confirmed constant upregulation in infected vascular tissues, emphasizing the role of lipid mediators like leukotrienes and prostaglandins in vascular remodelling and immune activation. These findings are distinctly applicable to abdominal aortic aneurysms (AAA), mainly the inflammatory subtype, as arachidonic acid metabolites are regarded to modify endothelial permeability, neutrophil chemotaxis, and cytokine release. By linking transcriptomic patterns of AA metabolism to vascular inflammation, this take a look at supports the inclusion of these genes in inflammatory aortic disease

profiling and aligns closely with the objectives of the current thesis, which seeks to discover transcriptomic changes associated with inflammation-driven aortic aneurysm pathogenesis.

Chapter 4

Materials and methodology

This study aims to identify and characterise differentially expressed genes and molecular pathways that distinguish inflammatory from noninflammatory aortic aneurysms through comprehensive transcriptomic profiling. By analysing RNA-seq data obtained from aneurysmal tissue samples, the study seeks to discover the gene expression alterations underlying the 2 pathologically distinct subtypes of the disease. The raw RNA sequencing reads have first been subjected to preprocessing, which includes adapter trimming, filtering of low-quality reads, and normalisation to decrease technical variations. High-quality reads had been aligned to the human reference genome using efficient spliced aligners, including HISAT2, and gene-level quantification was carried out the usage of featureCounts to acquire correct expression profiles. Differential gene expression analysis was performed using statistical packages such as Ballgown within the R environment to discover genes significantly upregulated or downregulated between inflammatory and noninflammatory samples. To gain useful insights, the differentially expressed genes were subjected to enrichment analysis using tools like DAVID, which enabled the identification of associated Gene Ontology (GO) phrases, biological processes, molecular functions, and enriched pathways. These findings can also serve as a foundation for future development of diagnostic biomarkers, molecular classification models, and targeted therapeutic strategies aimed at improving clinical management and outcomes in aortic aneurysm patients.

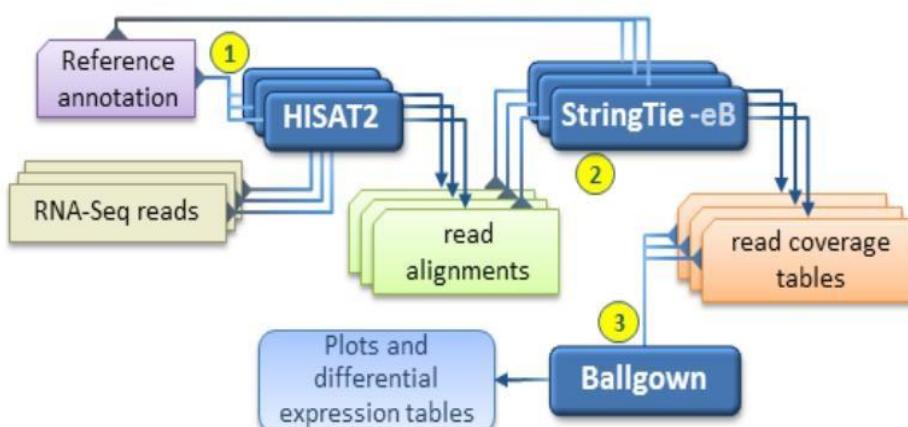


Figure 4.1: RNA-Seq Workflow

(Source: <https://github.com/gpertea/stringtie/issues/159>)

4.1.1 Wet Lab Procedures

To generate transcriptomic data for the comparative analysis of inflammatory and noninflammatory aortic aneurysms, preferred RNA sequencing (RNA-seq) protocols were observed. The technique began with the extraction of total RNA from aneurysmal tissue samples. RNA extraction was typically performed using commercially available silica column-based total purification kits, which ensure the isolation of extremely high-quality RNA while eliminating proteins and genomic DNA contaminants. The concentration and purity of the isolated RNA were evaluated using a Nanodrop spectrophotometer, and the integrity of the RNA was assessed using the Agilent 2100 Bioanalyzer (Chomczynski, P., & Sacchi, N., 2006). Only RNA samples with high RNA Integrity Numbers (RIN \geq 7) had been considered suitable for downstream processing. Following quality assessment, library preparation was achieved to transform RNA into sequencing-ready complementary DNA (cDNA) libraries. Initially, mRNA was enriched from total RNA using poly(A) tail selection or ribosomal RNA depletion protocols have been applied to retain non-coding RNA transcripts. The enriched RNA is then fragmented to a uniform size and reverse transcribed into cDNA. The cDNA fragments underwent repair, adapter ligation, and PCR amplification steps, using library preparation such as the NEBNext Ultra II RNA Library Prep Kit, to generate final sequencing libraries. Library size distribution and concentration were validated using a Bioanalyzer and Qubit fluorometer. The final cDNA libraries were then sequenced using a high-throughput next-generation sequencing (NGS) platform, including the Illumina HiSeq or NovaSeq systems. These platforms offer high sequencing depth and base-calling accuracy, making them ideal for capturing the global transcriptomic landscape. The resulting raw sequencing data were processed further using computational tools to perform differential gene expression and functional analyses (Illumina, 2023).

4.1.2 Data Selection

For this study, RNA-sequencing data were selected based on relevance to the pathological subtypes of aortic aneurysms, mainly, inflammatory and noninflammatory forms. Samples included only those derived from confirmed aneurysmal tissue with clear clinical annotation and histopathological classification. Preference was given to datasets generated using high-throughput Illumina platforms to ensure sequencing depth and consistency. Only samples with high RNA integrity and sufficient examination had been

included for downstream transcriptomic analysis. This careful selection ensured that the statistics were biologically relevant and technically robust for comparative gene expression evaluation.

4.1.2.1 SRP432690

SRP432690 is a Sequence Read Archive (SRA) dataset that carries transcriptomic profiles of aortic tissue samples collected from patients recognized with inflammatory and noninflammatory aortic aneurysms. The dataset contains 20 samples in general, consisting of 10 inflammatory and 10 non-inflamed aneurysm tissues, enabling a balanced assessment between the two subtypes. The samples had been derived from human aortic tissue and represent clinically characterised instances, making the dataset enormously applicable for differential gene expression evaluation. The 20 samples were downloaded using the ‘prefetch’ command from SRA, and then using ‘fastq-dump’, it splits the downloaded samples into forward and reverse. SRR24491385, SRR24491386, SRR24491387, SRR24491388, SRR24491392, SRR24491394, SRR24491395, SRR24491396, SRR24491419, SRR24491424. These were the accession numbers of the control (non-inflamed) samples in this study, and SRR24491397, SRR24491399, SRR24491401, SRR24491402, SRR24491405, SRR24491406, SRR24491411, SRR24491434, SRR24491433, SRR24491432 were the treatment (inflammatory) samples. The raw RNA-seq reads had been acquired from the SRA database in FASTQ format and subjected to quality control, alignment, and quantification for downstream transcriptomic analysis. This dataset gives a valuable resource for identifying key genes and molecular pathways involved in the pathogenesis of inflammatory versus noninflammatory aortic aneurysms.

SRA

SRP432690 | Search | Help

Access Public (50)

Source RNA (50)

Library Layout paired (50)

Platform Illumina (50)

Strategy other (50)

Data in Cloud GS (50)

S3 (50)

File Type fastq (50)

Clear all

Show additional filters

Summary 20 per page

View results as an expanded interactive table using the RunSelector. [Send results to Run selector](#)

Search results

Items: 1 to 20 of 50

<< First < Prev Page 1 of 3 Next > >> Last >>

Send to: Filters: [Manage Filters](#)

Find related data Database: [Select](#)

Find items

Search details

SRP432690[All Fields]

Search See more...

Recent activity Turn Off Clear

SRP432690 (50) SRA

SRR24491395 (1) SRA

See more...

1 RNA-seq of Homo Sapiens: Atherosclerotic control
1 ILLUMINA (Illumina HiSeq 4000) run: 53.4M spots, 10.8G bases, 3.5Gb downloads
Accession: SRX20276676

2 RNA-seq of Homo Sapiens: Atherosclerotic control
2 ILLUMINA (Illumina HiSeq 4000) run: 47.7M spots, 9.6G bases, 3.1Gb downloads
Accession: SRX20276675

3 RNA-seq of Homo Sapiens: Atherosclerotic control
1 ILLUMINA (Illumina HiSeq 4000) run: 50M spots, 10.1G bases, 3.3Gb downloads
Accession: SRX20276674

4 RNA-seq of Homo Sapiens: Atherosclerotic control
1 ILLUMINA (Illumina HiSeq 4000) run: 48.9M spots, 9.9G bases, 3.4Gb downloads
Accession: SRX20276673

5 RNA-seq of Homo Sapiens: Atherosclerotic control
1 ILLUMINA (Illumina HiSeq 4000) run: 53.2M spots, 10.8G bases, 3.6Gb downloads
Accession: SRX20276672

Figure 4.2: Data set

(Source:<https://www.ncbi.nlm.nih.gov/sra?term=SRP432690&cmd=DetailsSearch>)

```
asus@asus-Vivobook-ASUSLaptop-X1502ZA-X1502ZA:~/Desktop/samples/sra$ ls -ltrh
total 62G
-rw-rw-r-- 1 asus asus 3.6G Apr 17 15:13 SRR24491385_C.sra
-rw-rw-r-- 1 asus asus 3.2G Apr 17 15:33 SRR24491386_C.sra
-rw-rw-r-- 1 asus asus 3.4G Apr 17 22:39 SRR24491387_C.sra
-rw-rw-r-- 1 asus asus 3.4G Apr 17 23:14 SRR24491388_C.sra
-rw-rw-r-- 1 asus asus 2.7G Apr 18 00:03 SRR24491392_C.sra
-rw-rw-r-- 1 asus asus 2.8G Apr 18 10:00 SRR24491394_C.sra
-rw-rw-r-- 1 asus asus 2.8G Apr 18 10:11 SRR24491395_C.sra
-rw-rw-r-- 1 asus asus 3.0G Apr 18 10:23 SRR24491396_C.sra
-rw-rw-r-- 1 asus asus 2.5G Apr 18 21:12 SRR24491424_C.sra
-rw-rw-r-- 1 asus asus 3.3G Apr 18 22:17 SRR24491397_T.sra
-rw-rw-r-- 1 asus asus 3.2G Apr 18 23:09 SRR24491399_T.sra
-rw-rw-r-- 1 asus asus 2.6G Apr 18 23:29 SRR24491401_T.sra
-rw-rw-r-- 1 asus asus 3.5G Apr 19 00:02 SRR24491402_T.sra
-rw-rw-r-- 1 asus asus 2.8G Apr 19 09:57 SRR24491405_T.sra
-rw-rw-r-- 1 asus asus 3.2G Apr 19 11:54 SRR24491406_T.sra
-rw-rw-r-- 1 asus asus 3.0G Apr 19 15:21 SRR24491411_T.sra
-rw-rw-r-- 1 asus asus 2.6G Apr 21 09:26 SRR24491419_C.sra
-rw-rw-r-- 1 asus asus 3.2G Apr 21 09:39 SRR24491434_T.sra
-rw-rw-r-- 1 asus asus 3.4G Apr 21 09:50 SRR24491433_T.sra
-rw-rw-r-- 1 asus asus 3.8G Apr 21 10:02 SRR24491432_T.sra
```

Figure 4.3: List of downloaded samples

```
asus@asus-Vivobook-ASUSLaptop-X1502ZA-X1502ZA:~/Desktop/samples$ prefetch SRR24491387
2025-04-17T16:43:50 prefetch.2.11.3: Current preference is set to retrieve SRA Normalized Format files with full base quality scores.
2025-04-17T16:43:51 prefetch.2.11.3: 1) Downloading 'SRR24491387'...
2025-04-17T16:43:51 prefetch.2.11.3: SRA Normalized Format file is being retrieved, if this is different from your preference, it may be due to current file availability.
2025-04-17T16:43:51 prefetch.2.11.3: Downloading via HTTPS...
2025-04-17T17:09:36 prefetch.2.11.3: HTTPS download succeed
2025-04-17T17:09:43 prefetch.2.11.3: 'SRR24491387' is valid
2025-04-17T17:09:43 prefetch.2.11.3: 1) 'SRR24491387' was downloaded successfully
2025-04-17T17:09:43 prefetch.2.11.3: 'SRR24491387' has 0 unresolved dependencies
```

Figure 4.4: Samples downloading using the ‘prefetch’ command

```
asus@asus-Vivobook-ASUSLaptop-X1502ZA-X1502ZA:~/Desktop/samples/sra$ fastq-dump --split-files ./SRR24491385_C.sra
Read 53388478 spots for ./SRR24491385_C.sra
Written 53388478 spots for ./SRR24491385_C.sra
```

Figure 4.5: Splitting files using ‘fastq-dump’

4.2 RNA-Seq Data Analysis

The computational analysis in this study followed a strong RNA-seq pipeline aimed at processing, analysing, and interpreting gene expression data from inflammatory and noninflammatory aortic aneurysm samples. A major goal of RNA-seq analysis is to identify differentially expressed and coregulated genes and to infer biological meaning for further studies. The workflow included quality control, alignment, quantification, differential expression analysis, functional annotation, etc., as outlined below.

4.2.1 Quality control and Preprocessing

Raw RNA-seq reads in FASTQ format were first assessed for quality using FastQC, which generated reports on per-base sequence quality, GC content, adapter contamination, and sequence duplication. To enhance downstream accuracy, low-quality bases and adapter sequences were removed using Trim Galore, a wrapper tool that combines Cutadapt and FastQC, thereby generating high-quality, trimmed reads suitable for mapping. During trimming, we can find the adapter sequence.

```
asus@asus-Vivobook-ASUSLaptop-X1502ZA:~/Desktop/emb/samples$ time fastqc -o fastq_bt/ SRR24491402_T_1.fastq SRR24491402_T_2.fastq
Started analysis of SRR24491402_T_1.fastq
Approx 5% complete for SRR24491402_T_1.fastq
Approx 10% complete for SRR24491402_T_1.fastq
Approx 15% complete for SRR24491402_T_1.fastq
Approx 20% complete for SRR24491402_T_1.fastq
Approx 25% complete for SRR24491402_T_1.fastq
Approx 30% complete for SRR24491402_T_1.fastq
Approx 35% complete for SRR24491402_T_1.fastq
Approx 40% complete for SRR24491402_T_1.fastq
Approx 45% complete for SRR24491402_T_1.fastq
Approx 50% complete for SRR24491402_T_1.fastq
Approx 55% complete for SRR24491402_T_1.fastq
Approx 60% complete for SRR24491402_T_1.fastq
Approx 65% complete for SRR24491402_T_1.fastq
Approx 70% complete for SRR24491402_T_1.fastq
Approx 75% complete for SRR24491402_T_1.fastq
Approx 80% complete for SRR24491402_T_1.fastq
Approx 85% complete for SRR24491402_T_1.fastq
Approx 90% complete for SRR24491402_T_1.fastq
Approx 95% complete for SRR24491402_T_1.fastq
Analysis complete for SRR24491402_T_1.fastq
Started analysis of SRR24491402_T_2.fastq
```

Figure 4.6: Code for FastQC

```
asus@asus-Vivobook-ASUSLaptop-X1502ZA:~/Desktop/emb/samples/TrimGalore-0.6.18$ sudo ./trim_galore -q 20 --length 20 --paired ..../sra/SRR24491402_T_1.fastq ..../sra/SRR24491402_T_2.fastq
```

Figure 4.7: Code for Trimming

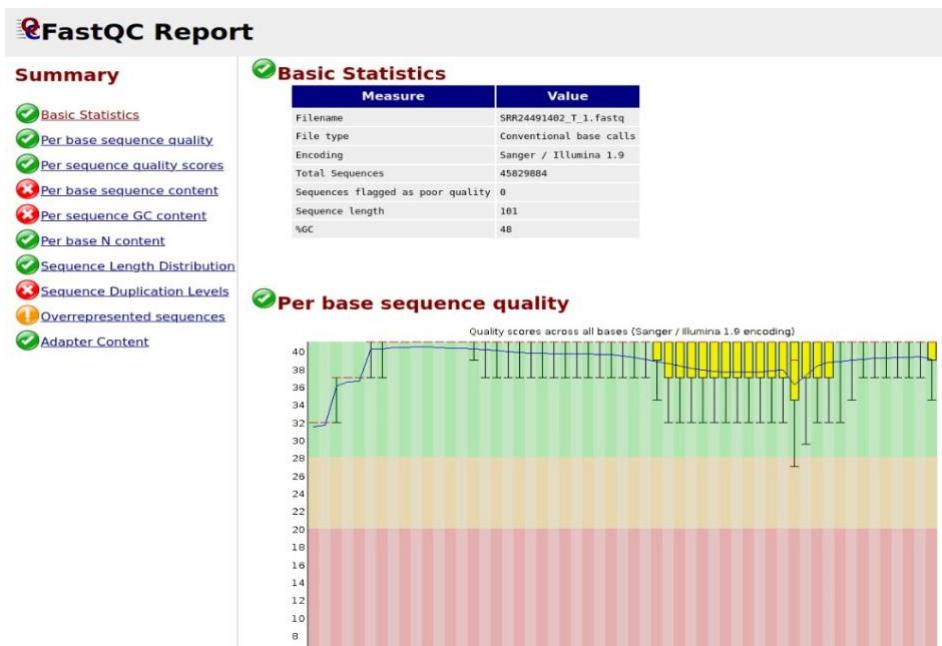


Figure 4.8: FastQC Report

4.2.2 Alignment to the Reference Genome using HISAT2

After quality control and adapter trimming, the high-quality RNA-seq reads were aligned to the human reference genome (hg38) using HISAT2 (Hierarchical Indexing for Spliced Alignment of Transcripts), a widely used splice-aware aligner specifically designed for efficient and accurate mapping of transcriptomic data after genome indexing. The aligner additionally helps the usage of recognised splice sites and transcript annotations (in GTF format), enhancing its sensitivity to both known and novel splicing events. For this study, default parameters have been used with suitable options for strand-specificity and paired-end read orientation. HISAT2 generated SAM (Sequence Alignment/Map) files as preliminary outputs, which were then converted to BAM (Binary Alignment/Map) format using SAMtools to reduce file size and enhance downstream processing efficiency. These BAM files contained detailed alignment statistics, together with mapping coordinates, strand orientation, CIGAR strings (representing the alignment styles), and alignment quality scores for each read. Accurate read alignment is a vital step inside the transcriptomic pipeline as it at once impacts the precision of gene quantification and downstream analyses, along with differential expression, variation calling, and isoform discovery. HISAT2's capability to deal with splice junctions, low memory utilisation, and high speed is especially suitable for aligning the RNA-seq data from both inflammatory and noninflammatory aortic aneurysm tissues. In this study, for all the samples, the

alignment rate that I got is above 85%. From all the BAM files, gtf files were generated using Stringtie.

And all these GTF files were merged into a single file.

```
asus@asus-Vivobook-ASUSLaptop-X1502ZA-X1502ZA:~/Desktop/embs/samples/sra$ time hisat2 -p1 --dta -x ..hg_new -1 ..TrimGalore-0.6.10/SRR24491385_C_1_val_1.fq.gz -2 ..TrimGalore-0.6.10/SRR24491385_C_2_val_2.fq.gz -S SRR24491385.sam
```

Figure 4.9: Code for Alignment

```
asus@asus-Vivobook-ASUSLaptop-X1502ZA-X1502ZA:~/Desktop/embs/samples/sra$ samtools sort -@ 1 -o SRR24491402.bam SRR24491402.sam  
[bam_sort_core] merging from 41 files and 1 in-memory blocks...
```

Figure 4.10: SAM to BAM

4.2.3 Transcript assembly and quantification

Following the alignment of high-quality, trimmed RNA-seq reads to the human reference genome (hg38) using the splice-aware aligner HISAT2, the next critical step involved transcript assembly and quantification to estimate gene and transcript expression levels. For this challenge, which compares transcriptomic profiles between inflammatory and noninflammatory aortic aneurysm samples, transcript assembly was completed using StringTie, a broadly used tool that appropriately reconstructs full-length transcripts and quantifies their expression. StringTie processes the aligned reads (in BAM layout) and assembles them into potential transcript structures based on the evidence from exon-exon junctions, splicing patterns, and gene annotations. It employs a network flow algorithm to predict all possible isoforms and gene models expressed in each sample, even identifying novel transcripts that may not be present in existing gene annotations. This step is especially crucial in transcriptomic studies like this one, wherein alternative splicing and isoform diversity can vary among diseased and control tissues, such as inflammatory and noninflammatory aneurysms. Once transcripts were assembled for individual samples, StringTie was run again in merge mode to create a unified transcriptome annotation (GTF file) by combining transcript structures across all samples. This allowed consistent transcript representation and comparison across all biological replicates and experimental groups. Subsequently, transcript quantification was performed to estimate the abundance of each gene and isoform, expressed in terms of FPKM (Fragments Per Kilobase of transcript per Million mapped reads).

```
asus@asus-Vivobook-ASUSLaptop-X1502ZA-X1502ZA:~/Desktop/embs/samples/sra$ stringtie SRR24491402.bam -l SRR24491402 -p 1 -G ..hg.gtf -o SRR24491402.gtf
```

Figure 4.11: GTF from BAM

```
chr244244_1.gtf
asus@asus-Vivobook-ASUSLaptop-X1502ZA-X1502ZA:/Desktop/embs/samples/sra$ stringtie --merge -p 1 -G ../hg.gtf -o stringtie_merged.gtf mergelist.txt
```

Figure 4.12: Code for merging GTF files

4.3 Differential expression analysis

Differential expression analysis is a crucial step in figuring out genes that exhibit significant changes between inflammatory and noninflammatory aortic aneurysm samples. In this study, noninflammatory samples were considered the control organization, while inflammatory aneurysm samples represented the disease condition. To detect transcriptional changes among these groups, the RNA-seq samples from the dataset SRP432690 have been analysed using the Ballgown package, a statistical tool designed in particular for differential expression analysis on the transcript and gene level. Ballgown is a part of the Tuxedo Suite and operates downstream of tools like HISAT2 and StringTie, allowing it to deal with transcript abundance information generated in terms of FPKM (Fragments Per Kilobase of transcript according to Million mapped reads). It applies expression levels using flexible linear models and compares transcript expression across sample groups to identify genes that are significantly upregulated or downregulated in the inflammatory condition. Statistical thresholds, such as p-values and fold change, have been used to determine differentially expressed genes (DEGs). The resulting DEGs served as a foundation for further downstream analysis, including functional annotation and interaction network construction, to elucidate key genes and pathways associated with inflammation-driven aortic aneurysm progression.

4.3.1 Ballgown

Ballgown is an R package specially designed for transcript-level differential expression analysis in RNA-Seq research, mainly in workflows that involve transcript assembly using tools like StringTie. Unlike traditional count-based methods consisting of edgeR or DESeq2, Ballgown does not rely upon raw read counts; alternatively, it utilises abundance estimates which include FPKM (Fragments Per Kilobase of transcript in line with Million mapped reads) or TPM (Transcripts Per Million), which are usually output in. CTAB files by using StringTie. The package structures these abundance estimates into a specialised object that enables differential expression analysis using flexible linear modelling frameworks. This makes Ballgown mainly powerful for exploring isoform-level expression, alternative splicing events, and transcript structure variations, providing a

more refined view of transcriptomic changes than gene-level methods. It also includes visualisation tools that allow researchers to plot transcript models, expression levels, and statistical comparisons throughout samples. Ballgown supports analyses at multiple levels, which include gene, transcript, and exon, making it flexible for designated transcriptomic investigations. Its compatibility with upstream tools like HISAT2 (for alignment) and StringTie (for quantification), Ballgown is a key component in RNA-Seq pipelines aimed at characterising transcriptome complexity, mainly in contexts where isoform-specific regulation is of biological interest.

4.3.2 Libraries used

library(ballgown): For differential expression analysis,

library(dplyr): For data handling

library(matrixStats): Perform row- and column-wise statistical operations on matrices or data frames.

library(ggplot2): For data visualization in R

library(readr): For reading and writing tabular data

library(ggrepel): Prevents overlapping text labels in plots.

library(pheatmap): Generates heatmaps for gene expression data

4.4 Functional Annotation and Enrichment Analysis

To gain biological insights into the identified differentially expressed genes (DEGs), functional enrichment analysis was performed using the DAVID (Database for Annotation, Visualization and Integrated Discovery) tool. DAVID allows whole annotation of gene lists, providing insights into their biological significance through categorisation into Gene Ontology (GO) terms and KEGG pathways. The GO analysis became subdivided into three primary domain names: Biological Processes (BP), which describe the physiological pathways and processes involving the DEGs; Molecular Functions (MF), which detail the biochemical activities of gene products at the molecular level; and Cellular Components (CC), which indicate the cellular location where gene products are active. These classifications allowed for a structured understanding of the functional roles of DEGs. In addition, KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway enrichment analysis was performed to identify statistically overrepresented signalling and metabolic

pathways among DEGs. This helped highlight important pathways potentially involved in aortic aneurysm formation and development, along with inflammatory reaction pathways, extracellular matrix transforming, vascular smooth muscle contraction, and immune signalling. Enrichment results had been evaluated using p-values and false discovery rate (FDR) thresholds to ensure statistical reliability. The enriched GO terms and pathways offered valuable context for interpreting transcriptomic differences among inflammatory and noninflammatory aortic aneurysm samples. This evaluation facilitated the identification of biologically meaningful targets and mechanisms that would contribute to disease stratification and potential therapeutic interventions.

4.4 Diagrammatic Workflow

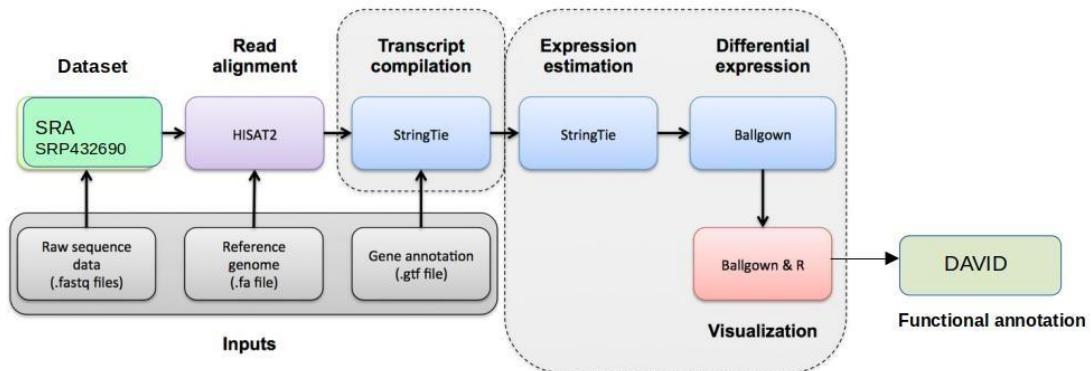


Figure 4.13: Diagrammatic Workflow

(Source:https://rnabio.org/module-03-expression/0003/03/01/Differential_Expression-Ballgown/)

Chapter 5

Results and Discussion

This chapter gives the findings received from the complete transcriptomics evaluation carried out on RNA sequencing data derived from inflammatory and noninflammatory aortic aneurysm tissues. The intention is to become aware of differentially expressed genes (DEGs) and associated biological pathways that would distinguish the two subtypes on the molecular stage. Through a dry lab pipeline- alignment, transcript assembly, expression quantification, and differential expression analysis had been discovered. Subsequent practical enrichment analysis helped interpret those gene sets inside the context of biological procedures and disease-associated pathways. The results described herein offer insights into the molecular signatures underlying the pathogenesis of inflammatory and noninflammatory aortic aneurysms, potentially contributing to improved type and future biomarker discovery. The results obtained from all the methods used in the study are discussed here, and they include plots obtained from different R packages.

5.1 Quality Control of RNA-Seq Data

The quality of raw RNA-seq reads was evaluated using FastQC to ensure data reliability for downstream transcriptomic analysis. The consequences indicated that the sequencing reads were of high quality, with most base positions displaying Phred rankings above Q30, signifying a base call accuracy greater than 99.9%. This excessive according-to-base sequence satisfies the integrity of the records. The GC content of the reads fell inside the expected range for human transcriptomic data, suggesting an absence of contamination. Sequence length distribution changed to uniform, as predicted from constant library preparation. However, the collection duplication stage was flagged as a problem. The duplication plot revealed that a considerable number of sequences were duplicated more than 10 times, with only 13.95% of sequences remaining unique after potential deduplication. This can be attributed to PCR amplification bias or the presence of relatively abundant transcripts consisting of ribosomal RNA or housekeeping genes. Another flagged parameter changed in the presence of overrepresented sequences, generally, which include an extended poly-A stretch, possibly originating from mRNA poly-A tails. This is a common feature in RNA-seq libraries, especially whilst oligo(dT)

primers are used during reverse transcription. Despite those warnings, no enormous adapter contamination or high levels of ambiguous base calls (Ns) have been found. Overall, the dataset demonstrated is of acceptable quality for further evaluation, and high-quality trimming was done to limit the effect of duplicated and overrepresented reads in the downstream workflow.

5.2 Alignment Statistics

Cleaned RNA-seq reads were aligned to the human reference genome (hg38) using HISAT2. Alignment rates exceeded 85% across all samples, indicating efficient and high-quality mapping. The output BAM files contained mapped read information, including genomic coordinates, strand orientation, and splice junctions. Reads were successfully aligned across exon-exon junctions, enabling accurate transcript-level assembly in subsequent analysis steps.

5.3 Transcript Assembly and Quantification

Transcript assembly and quantification were performed using StringTie on the aligned BAM files generated by HISAT2. The tool successfully assembled full-length transcripts for all samples, and gene expression was quantified as FPKM values. Across the dataset, a consistent number of transcripts were detected per sample, with a median of approximately 47,758 transcripts identified. The merged transcriptome annotation allowed for uniform gene tracking across all samples. The assembled transcripts covered known protein-coding genes as well as potential novel isoforms. The resulting expression matrix was used as input for differential expression analysis in Ballgown.

5.4 Differential Gene Expression Analysis

Differential gene expression analysis was conducted using the Ballgown R package. A total of 1,312 genes had been analysed for differential expression among inflammatory and noninflammatory aortic aneurysm samples. Of those, 147 genes had been identified as differentially expressed (adjusted $p < 0.05$ and $|\log_2 \text{fold change}| \geq 1$), which includes 126 upregulated and 21 downregulated genes in the inflammatory group. The last 1,165 genes did not show statistically significant expression changes and had been taken into consideration non-differentially expressed. The DEGs included key inflammatory and extracellular matrix-associated genes, including IL6, MMP9, CXCL8, and TNFAIP3. The distribution of expression changes was visualised using a volcano plot and hierarchical

clustering heatmap, displaying awesome gene expression patterns between the two sample groups.

5.5 Plots of differential expression analysis

5.5.1 Boxplot of Raw Normalized Data ($\log_2(\text{FPKM}+1)$)

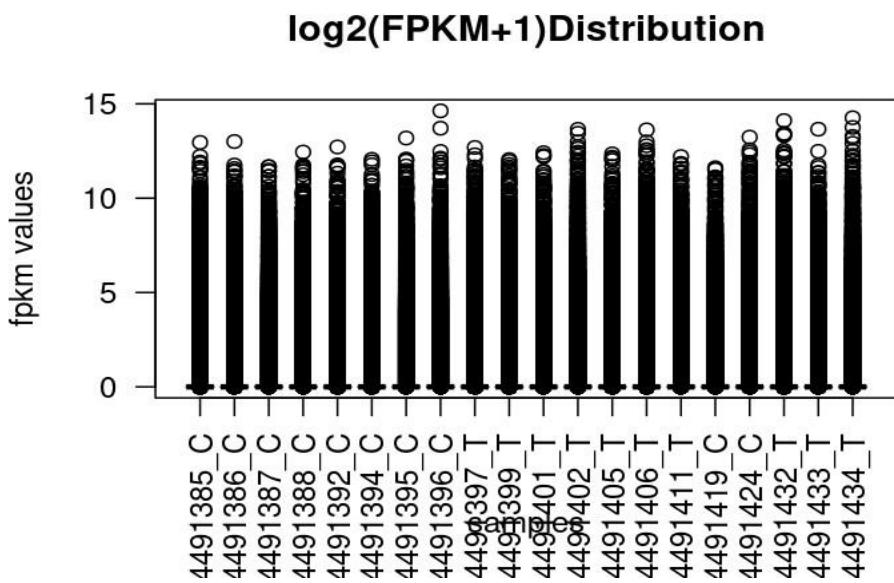


Figure 5.1: Boxplot of $\log_2(\text{FPKM}+1)$ values for raw normalized data across all samples.

To assess the overall expression distribution across all samples before differential analysis, a boxplot of the \log_2 -transformed FPKM (Fragments Per Kilobase of transcript per Million mapped reads) values was generated. This plot includes both control (denoted by "C") and treatment (denoted by "T") groups and reflects the range and consistency of gene expression profiles before transcript filtering. Each box in the plot represents the distribution of expression levels within an individual sample, providing insights into the median, interquartile range, and presence of outliers. The resulting boxplot reveals a generally uniform distribution with closely aligned medians across all samples. This uniformity suggests that no major batch effects or technical biases are influencing the data, thereby validating the quality of initial sample processing. However, some variation is noted in the spread of the data, with a noticeable number of outliers indicating genes with very high or very low expression levels. These may be biologically significant or

represent transcriptional noise. The observation underscores the importance of subsequent filtering steps to enhance the clarity and biological interpretability of the data.

5.5.2 Boxplot of Filtered Normalized Data ($\log_2(\text{FPKM}+1)$)

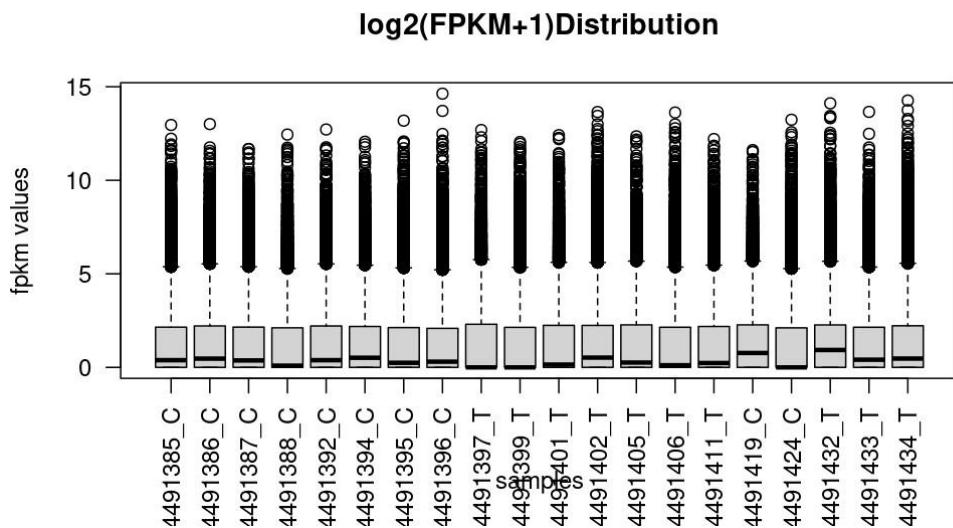


Figure 5.2 : Boxplot of $\log_2(\text{FPKM}+1)$ values after transcript filtering and normalization.

After the initial quality assessment, low-abundance and potentially non-informative transcripts were removed to refine the dataset. A second boxplot was generated using this filtered set of transcripts to visualise the effect of data preprocessing. The plot again presents the \log_2 -transformed FPKM values across all control and treatment samples, allowing direct comparison with the unfiltered distribution. The filtered data exhibit an improved consistency across samples, with more tightly grouped medians and a reduction in interquartile range variability. The number of outliers is also significantly reduced, indicating that noise and extreme values introduced by lowly expressed or inconsistent transcripts have been effectively minimised. These changes support the effectiveness of the filtering strategy and confirm that the retained transcripts are more reliably and consistently expressed. This preprocessing step is essential in transcriptomic workflows to ensure that downstream analyses such as differential expression testing are performed on high-quality, biologically meaningful data. Ultimately, this refined dataset lays the groundwork for the accurate identification of differentially expressed genes and contributes to more robust biological interpretation.

5.5.3 Expression Boxplot of the 20th Gene

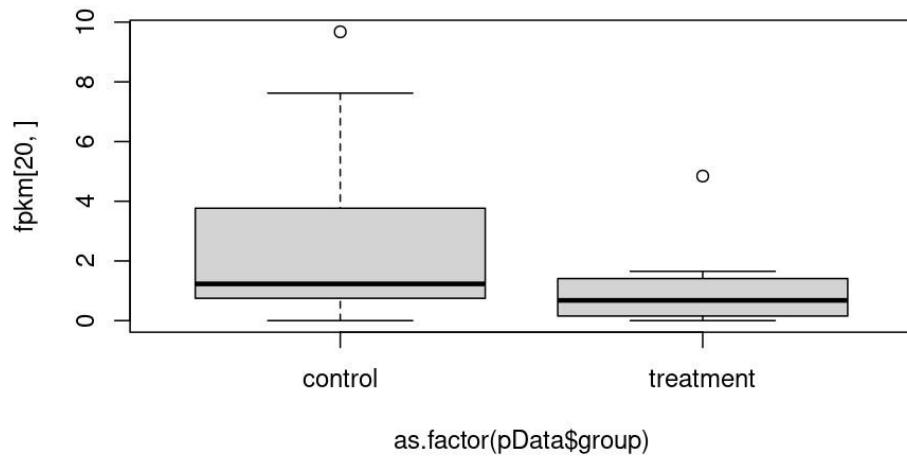


Figure 5.3: Expression levels of the 20th gene across control and treatment samples.

Focusing on an individual gene-level comparison, the expression of the 20th gene (by index) was plotted across control and treatment groups. The boxplot indicates a notable decrease in expression in the treatment group, evidenced by a lower median and compressed interquartile range. This reduction suggests potential downregulation of this gene in inflammatory aortic aneurysm samples compared to noninflammatory controls. The consistent downshift across replicates implies a biological effect rather than technical variation. Such focused gene-specific visualizations are important to validate broader differential expression trends and may highlight individual biomarkers relevant to the disease phenotype under study. These detailed comparisons offer valuable insights for future functional validation and hypothesis generation regarding gene involvement in pathogenesis.

5.5.4 Differential Expression Category Bar Plot

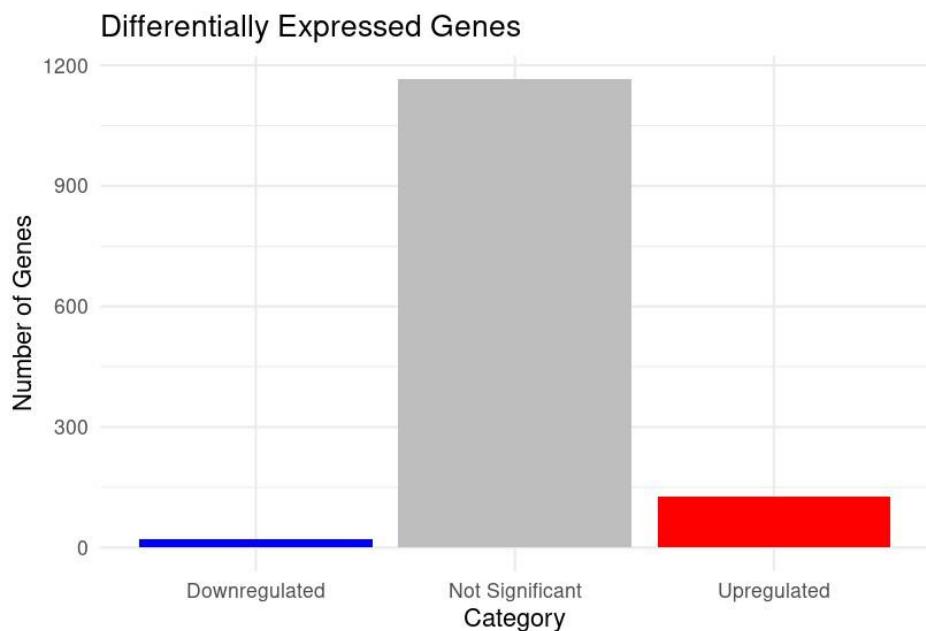


Figure 5.4: Bar plot showing counts of upregulated, downregulated, and non-significant genes.

A categorical summary of differentially expressed genes (DEGs) was visualized using a bar plot that classifies genes into three groups: upregulated, downregulated, and not significantly altered. The plot reveals that the majority of genes were not significantly differentially expressed between control and treatment groups, which is expected given the complex background expression of most genes. However, a distinct subset of genes exhibited statistically significant upregulation (red) and downregulation (blue) in the treatment group. This visualization serves as a summary of the differential analysis outcome and highlights the potential biomarkers or mechanistically relevant genes associated with the inflammatory status of aortic aneurysms.

5.5.5 Heatmap of All Differentially Expressed Genes (DEGs)

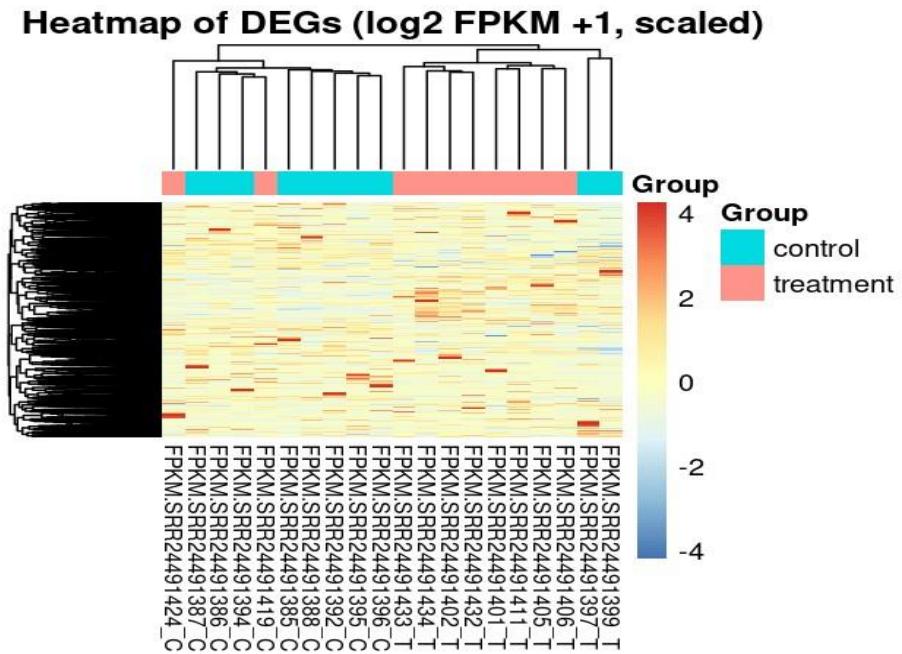


Figure 5.5: Heatmap displaying expression patterns of all DEGs (log2 FPKM +1, scaled)

The comprehensive heatmap of all differentially expressed genes (DEGs) was generated using log2-transformed FPKM values, followed by Z-score normalization to enable cross-gene comparison across all samples. This heatmap offers a global visualization of gene expression patterns distinguishing inflammatory from non-inflammatory aortic aneurysm samples. Hierarchical clustering analysis revealed a clear separation between the control and treatment groups, indicating strong transcriptional divergence attributable to the inflammatory state of the aneurysms. Notably, the clustering of biological replicates within each group underscores the consistency and reproducibility of the gene expression data. Co-expressed gene modules represented by grouped bands of red (upregulation) and blue (downregulation) suggest the activation or repression of entire gene programs in response to inflammation. These altered programs likely reflect key molecular mechanisms involved in inflammatory remodelling of the aortic wall, such as leukocyte recruitment, cytokine signalling, and degradation of structural matrix proteins. The ability of the DEG expression profiles to distinctly classify the two clinical subtypes reinforces the biological validity of the results and highlights the heatmap as a pivotal visualization linking gene expression signatures to aneurysm phenotype.

5.5.6 Heatmap of Top 50 Differentially Expressed Genes

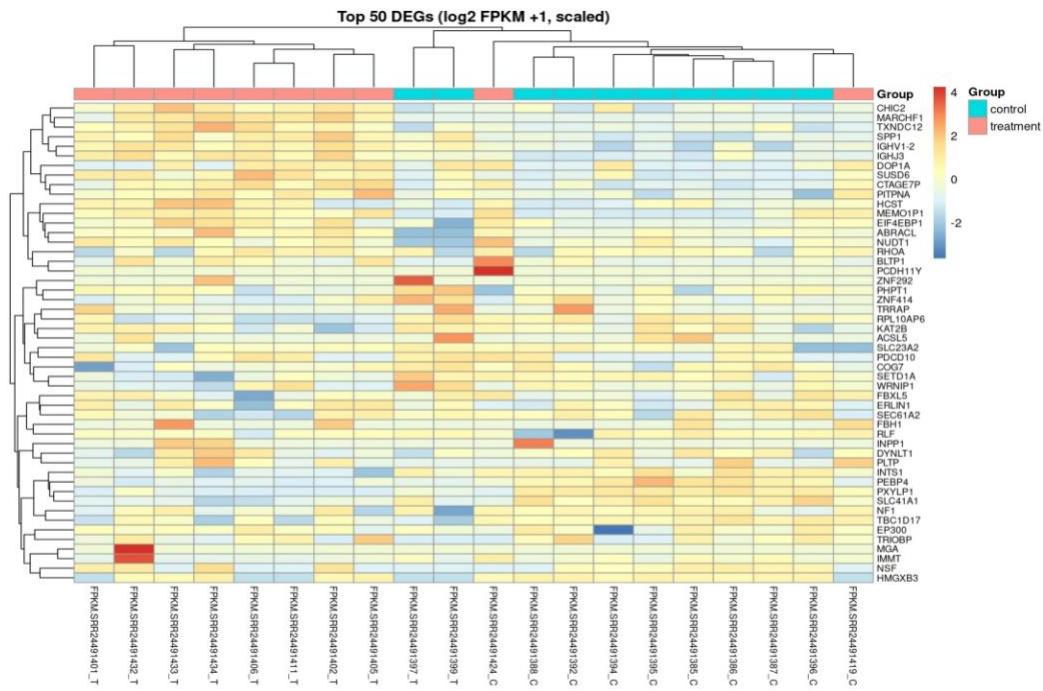


Figure 5.6: Heatmap of the top 50 most significant DEGs (log2 FPKM +1, scaled).

To refine the interpretation and identify high-priority molecular targets, a focused heatmap was constructed using the top 50 most significantly differentially expressed genes, selected based on adjusted p-values and fold changes. These genes exhibit the strongest and most consistent transcriptional alterations between inflammatory and non-inflammatory aortic aneurysm groups. In this heatmap, clear and reproducible clustering of samples reflects the distinct transcriptional identity of the inflammatory group, with multiple genes showing marked upregulation or suppression. The upregulated genes are likely involved in key inflammatory processes such as cytokine signalling (e.g., IL-6, TNF pathways), immune cell trafficking (e.g., chemokines), and tissue remodelling, which are central to the pathophysiology of inflammatory aneurysms. Conversely, the downregulated genes may be related to structural maintenance of the vascular wall or homeostatic signalling disrupted during chronic inflammation. This visualization provides a concise yet informative representation of the most relevant gene expression shifts and offers a prioritized set of biomarker candidates for future experimental validation, potentially aiding in the molecular classification and diagnosis of inflammatory aortic aneurysms.

5.5.7 Volcano Plot of Differentially Expressed Transcripts (DETs)

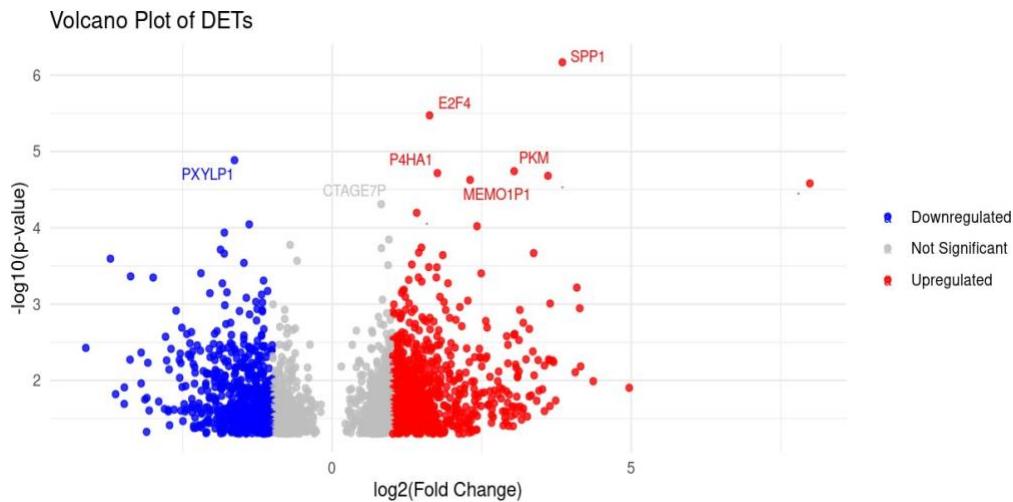


Figure 5.7: Volcano plot showing log2 fold change and p-values for DETs, highlighting key transcripts.

The volcano plot provides an integrated visualization of both the magnitude and statistical significance of transcript-level expression changes. Each point represents a transcript, plotted by log2 fold change on the x-axis and $-\log_{10}$ p-value on the y-axis. Transcripts falling into the upper-left and upper-right quadrants, denoting significant downregulation and upregulation, respectively, were colored in blue and red for clarity. This plot highlights a subset of transcripts that exhibit both strong fold changes and high statistical confidence. Notable among the upregulated transcripts are SPP1, PKM, E2F4, and MEMO1P1, which may be involved in metabolic regulation, cell cycle control, or inflammatory signalling. Conversely, PXYLP1 was found significantly downregulated, suggesting its potential suppression in the inflammatory aneurysmal environment. Overall, this volcano plot aids in prioritising key transcripts for further investigation and underscores the transcriptome-wide alterations induced in the disease condition.

5.5.8 Volcano Plot of Differentially Expressed Genes (DEGs)

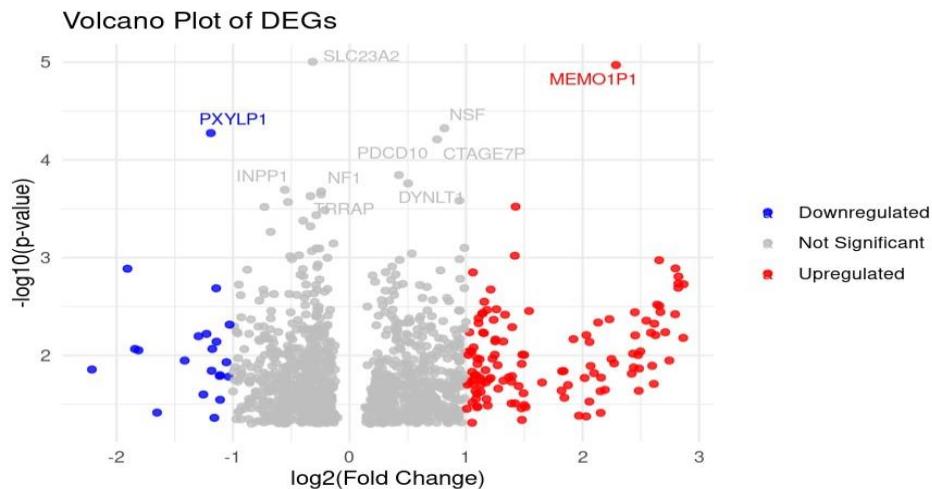


Figure 5.8: Volcano plot showing log2 fold change and p-values for DEGs, highlighting key genes.

The volcano plot presented here provides a comprehensive overview of the gene-level differential expression landscape between inflammatory and non-inflammatory aortic aneurysm samples. Each point on the plot corresponds to an individual gene, plotted based on its log2 fold change (x-axis) and $-\log_{10} p\text{-value}$ (y-axis), thereby illustrating both the magnitude and statistical significance of expression changes. Genes that were significantly upregulated in the inflammatory (treatment) group are shown in red, while those significantly downregulated are in blue. Grey dots represent genes with nonsignificant expression differences. One of the most prominently upregulated genes observed was MEMO1P1, which may be implicated in cellular migration and redox signalling processes relevant to vascular remodelling and immune cell infiltration, both hallmarks of inflammation in aneurysmal pathology. Conversely, PXYLP1, a gene with potential roles in glycosaminoglycan metabolism, was markedly downregulated in the inflammatory group, suggesting suppression of extracellular matrix regulation, possibly contributing to structural weakening of the aortic wall. The volcano plot thus serves as a critical visualization to prioritize genes that may be central to the inflammatory mechanisms differentiating the two aneurysm types. It helps refine the focus toward a subset of genes for downstream functional validation and pathway enrichment, shedding light on the molecular underpinnings of vascular inflammation and tissue degeneration that characterize inflammatory aortic aneurysms.

5.6 Functional Enrichment Analysis of DEGs using DAVID

To gain deeper biological insight into the transcriptomic variations between inflammatory and noninflammatory aortic aneurysm samples, functional enrichment analysis was conducted using the DAVID (Database for Annotation, Visualization, and Integrated Discovery) tool. A total of 147 differentially expressed genes (DEGs), consisting of 126 upregulated and 21 downregulated genes, had been entered for annotation and pathway evaluation. Gene Ontology (GO) enrichment found significant enrichment in biological processes which include inflammatory reaction, immune defence system, and reaction to cytokine stimulus, highlighting the activation of immune and senescence-inflammatory pathways in the treatment group. In the Molecular Function category, enriched terms included cytokine interest and chemokine receptor binding, indicating active signalling roles. Cellular Component annotations prominently featured extracellular location, consistent with secreted immune factors and matrix-degrading enzymes. In parallel, KEGG pathway enrichment analysis recognized numerous key signalling pathways. Prominent among these had been the TNF signalling pathway, cytokine–cytokine receptor interaction, NF-κB signalling pathway, and IL-17 signalling pathway. These pathways are well-known mediators of vascular inflammation, extracellular matrix remodelling, and immune cellular recruitment, all hallmarks of inflammatory aortic aneurysm pathology. These results align with observed upregulation of inflammatory genes, together with IL6, CXCL8, and TNFAIP3, within the treatment organization. The combined GO and KEGG enrichment analysis demonstrates that the transcriptomic changes in inflammatory aneurysms are pushed through coordinated activation of immune and inflammatory responses. These findings support the hypothesis that inflammation drives distinct molecular signatures in aneurysm pathogenesis and may serve as a foundation for biomarker discovery and therapeutic targeting.

Figure 5.9: Top 30 Enriched KEGG Pathways

The KEGG pathway enrichment analysis revealed a complete profile of molecular mechanisms appreciably associated with the differentially expressed genes (DEGs) diagnosed among inflammatory and noninflammatory aortic aneurysm samples. The evaluation uncovered 30 significantly enriched pathways, lots of which are directly related to immune regulation, inflammatory signalling, pathogen defence, and vascular remodelling.

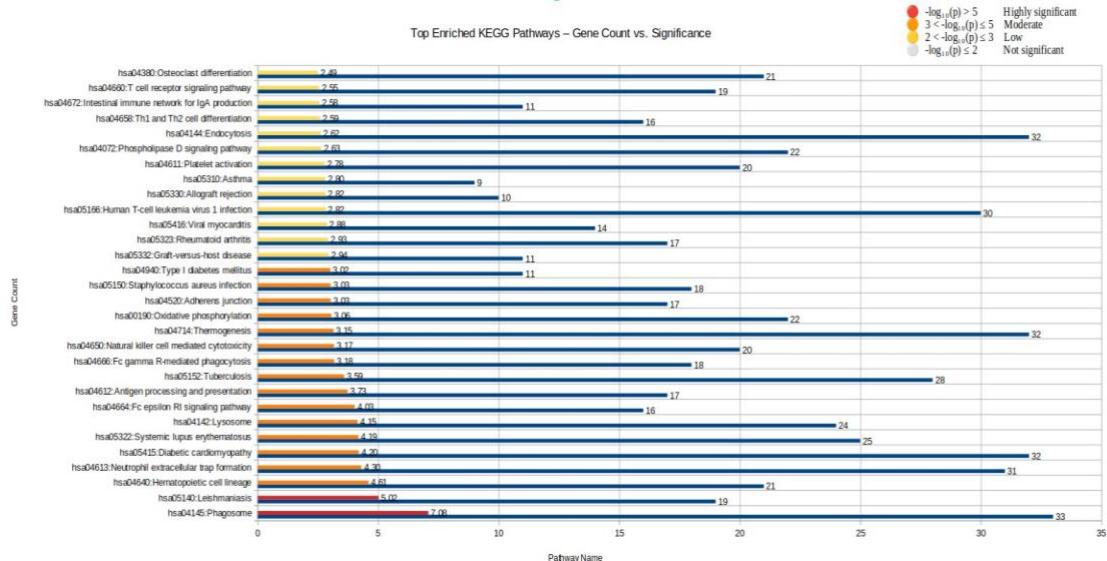


Figure 5.10: Functional annotation chart (KEGG pathway enrichment)

The KEGG pathway enrichment analysis performed using DAVID revealed multiple biologically significant pathways that are potentially involved in the pathogenesis and progression of inflammatory aortic aneurysms. The visualization (bar chart) displays the top enriched pathways based on gene count and statistical significance, ranked by –

$\log_{10}(p\text{-value})$, with colour coding indicating levels of significance (red = highly significant, orange = moderate, yellow = low).

Key findings from the enrichment analysis include:

1. Immune and Inflammatory Pathways

- Pathways such as Phagosome, Leishmaniasis, Hematopoietic cell lineage, Antigen processing and presentation, and Natural killer cell-mediated cytotoxicity are significantly enriched, with many associated genes.
- These pathways underline the active involvement of immune cell activation, cytokine signalling, and antigen presentation, which are hallmarks of vascular inflammation and immune-mediated remodelling observed in inflammatory aortic aneurysms

2. Autoimmune and Infectious Triggers:

- Pathways like Systemic lupus erythematosus, Rheumatoid arthritis, Type I diabetes mellitus, and *Staphylococcus aureus* infection are enriched. These represent autoimmune or infectious etiologies that could either contribute to or share common mechanisms with vascular inflammation.
- The enrichment of Fc gamma receptor-mediated phagocytosis and T-cell receptor signalling further supports the autoimmune/inflammatory hypothesis in aneurysmal tissue.

3. Cardiometabolic and Endothelial Dysfunction:

- Enrichment in Platelet activation, Viral myocarditis, and Diabetic cardiomyopathy pathways suggests overlaps with endothelial damage, oxidative stress, and chronic low-grade inflammation, frequently implicated in aneurysm pathobiology.

4. Highly Significant Pathways:

- The Phagosome pathway ($-\log_{10}(p) = 7.08$, gene count = 33) and Leishmaniasis ($-\log_{10}(p) = 5.02$, gene count = 19) were identified as the most statistically significant, indicating strong transcriptional activation in genes associated with

macrophage function and pathogen defence, reflective of chronic inflammation in aneurysmal tissue.

The KEGG pathway enrichment evaluation provided essential insights into the biological techniques and molecular pathways most suffering by the differentially expressed genes (DEGs) diagnosed in this study. Among the drastically enriched pathways, a sturdy predominance of immune-associated and irritation-related signalling cascades was determined, underscoring the molecular difference between inflammatory and noninflammatory aortic aneurysms.

The maximum extraordinarily big pathway diagnosed changed into the Phagosome pathway (hsa04145), which exhibited the highest statistical enrichment with a $-\log_{10}(\text{pcost})$ of 7.08 and a total of 33 DEGs mapped. This pathway is imperative to macrophage-mediated engulfment and degradation of pathogens and apoptotic cells, processes that are recognised to be upregulated throughout continual vascular infection. The activation of phagosome-related genes highlights the heightened immune surveillance and tissue remodelling happening in inflammatory aneurysmal walls. In parallel, the Leishmaniasis pathway (hsa05140; $-\log_{10}(p) = 5.02$) turned out to be additionally substantially enriched. While related to infectious disorders, its inclusion within the enriched effects is attributed to the shared immune activation genes involved in macrophage and T-cell signalling, both of which are known to play crucial roles in vascular inflammatory conditions.

Further enriched pathways include Antigen Processing and Presentation (hsa04612), T-cell Receptor Signalling (hsa04660), Natural Killer Cell-Mediated Cytotoxicity (hsa04650), and Fc Gamma R-Mediated Phagocytosis (hsa04666). These together constitute the orchestrated innate and adaptive immune responses that differentiate inflammatory aneurysm profiles from noninflammatory ones. The constant presence of pathways concerned in immune reputation, cytokine signalling, and cellular immunity displays the transcriptomic reprogramming driven by chronic inflammatory stimuli within the aortic tissue. Interestingly, several pathways also link vascular inflammation to systemic autoimmune and cardiometabolic conditions. Notable among them are Systemic Lupus Erythematosus (hsa05322), Rheumatoid Arthritis (hsa05323), Type 1 Diabetes Mellitus (hsa04940), and Diabetic Cardiomyopathy (hsa05415). These findings indicate overlapping pathophysiological mechanisms, which include immune cell infiltration,

endothelial dysfunction, and cytokine-mediated tissue damage, which might be observed in both autoimmune diseases and aneurysmal progression. Pathways like Platelet Activation (hsa04611) and Staphylococcus aureus Infection (hsa05150) have been additionally enriched, suggesting interactions among coagulation, contamination, and immune reaction within aneurysmal tissue. The enrichment of Oxidative Phosphorylation (hsa00190) and Lysosome (hsa04142) further displays the metabolic and degradative shifts occurring in inflamed vascular environments.

Overall, the practical annotation consequences pretty significantly signify the inflammatory aortic aneurysm transcriptome as being heavily skewed toward immune law, inflammatory signalling, and antigen-driven responses.

Pathway	$-\log_{10}(p\text{-value})$	Gene Count	Key Genes Involved
Phagosome (hsa04145)	7.08	33	TFRC, ITGB2, CTSS, HLA-DMA, NCF2
Leishmaniasis (hsa05140)	5.02	19	MT-ND4, PIK3CD, COX7A1, CYBA
Hematopoietic Cell Lineage (hsa04640)	4.61	21	SPI1, CSF1R, ITGA4, FCER1G
Neutrophil Extracellular Trap (NET) Formation	4.30	32	SYK, GAB2, MAPK14, VAV1
Diabetic Cardiomyopathy (hsa05415)	4.20	31	HLA-DMA, NCF2, EGFR

Table 5.2: Highly Significant Pathways ($\log_{10}(p) > 4.5$)

These top-ranked pathways provide strong evidence that inflammatory aneurysms are characterized by enhanced immune activation, metabolic stress, and matrix remodelling.

The table listed below shows the downregulated 21 genes and their features.

gene_name	log2fc	pvalue	padj	Category
ZNF10	-1.04	0.02	0.04	Downregulated
MTUS2	-1.18	0.01	0.04	Downregulated
SNORD115-43	-1.65	0.04	0.05	Downregulated
BCL7C	-1.03	0	0.04	Downregulated
DAPL1	-1.23	0.01	0.04	Downregulated
CBR3-AS1	-1.3	0.01	0.04	Downregulated
PXYLP1	-1.19	0	0.02	Downregulated
CYTL1	-1.41	0.01	0.04	Downregulated
ADGRA3	-1.06	0.01	0.04	Downregulated
CARTPT	-2.21	0.01	0.04	Downregulated
LGR6	-1.25	0.03	0.04	Downregulated
SFRP1	-1.11	0.02	0.04	Downregulated
PTPRD	-1.14	0.01	0.04	Downregulated
SNORD62	-1.14	0	0.04	Downregulated
MIR421	-1.16	0.04	0.05	Downregulated
PCDH11X	-1.81	0.01	0.04	Downregulated
SPANXB1	-1.84	0.01	0.04	Downregulated
PCDH11Y	-1.91	0	0.04	Downregulated
ATRNL1	-1.11	0.02	0.04	Downregulated
C11orf96	-1.11	0.03	0.04	Downregulated
FAT3	-1.18	0.01	0.04	Downregulated

Table 5.3: List of the 21 Downregulated genes

And the next table below shows the upregulated 126 genes and their features.

gene_name	log2fc	pvalue	padj	category
ALOX5AP	1.13	0.02	0.04	Upregulated
TRAV8-2	1.21	0	0.04	Upregulated
TRAV8-4	1.09	0.02	0.04	Upregulated
GZMB	1.34	0	0.04	Upregulated
SERPINA1	1.21	0.02	0.04	Upregulated
IGHG4	2.59	0.01	0.04	Upregulated
IGHA1	2.27	0.01	0.04	Upregulated
IGHJ3	2.8	0	0.04	Upregulated
IGHD6-25	1.09	0.03	0.04	Upregulated
IGHD2-8	1.05	0.03	0.04	Upregulated
IGHV1-2	2.66	0	0.04	Upregulated
IGHV4-4	2.67	0	0.04	Upregulated
IGHV2-5	2.06	0.01	0.04	Upregulated
IGHV3-64D	2.64	0	0.04	Upregulated
IGHV5-10-1	1.5	0.03	0.05	Upregulated
IGHV1-18	2.03	0.04	0.05	Upregulated
IGHV3-20	2.42	0.02	0.04	Upregulated
IGHV3-21	2.16	0.02	0.04	Upregulated
IGHV4-28	2.02	0.02	0.04	Upregulated
IGHV3-30	2.49	0.01	0.04	Upregulated
IGHV4-39	2.15	0.02	0.04	Upregulated
IGHV3-53	2.03	0.01	0.04	Upregulated
IGHV1-45	1.45	0.02	0.04	Upregulated
IGHV5-51	2.66	0	0.04	Upregulated
IGHV4-55	2.07	0.01	0.04	Upregulated
IGHV4-59	2.55	0	0.04	Upregulated
IGHV4-61	1.97	0.04	0.05	Upregulated
IGHV3-66	2.86	0.01	0.04	Upregulated
IGHV2-70	2.6	0.01	0.04	Upregulated
IGHV3-72	1.84	0.03	0.04	Upregulated
MT1G	1.65	0.02	0.04	Upregulated
GBP3	1.06	0.01	0.04	Upregulated
GRAP	1.01	0.02	0.04	Upregulated
CCL18	1.5	0.01	0.04	Upregulated
CD53	1.06	0.02	0.04	Upregulated

ACP5	1.47	0.03	0.05	Upregulated
IFI30	1.02	0.01	0.04	Upregulated
HCST	1.43	0	0.03	Upregulated
CD79A	1.82	0.01	0.04	Upregulated
APOE	1.39	0.02	0.04	Upregulated
LDAH	1.03	0.01	0.04	Upregulated
QPCT	1.18	0	0.04	Upregulated
CAPG	1.27	0.01	0.04	Upregulated
VAMP8	1.18	0.03	0.04	Upregulated
IGKV3OR2-268	1.48	0.01	0.04	Upregulated
ANKRD36BP2	1.54	0	0.04	Upregulated
IGKJ2	2.5	0.01	0.04	Upregulated
IGKV4-1	2.25	0.01	0.04	Upregulated
IGKV1-39	2.63	0.01	0.04	Upregulated
IGKV3-7	1.32	0.02	0.04	Upregulated
IGKV3-15	2.06	0.03	0.04	Upregulated
IGKV1-17	2.74	0.01	0.04	Upregulated
IGKV3-20	2.87	0	0.04	Upregulated
IGKV2-24	2.23	0	0.04	Upregulated
IGKV1-27	2.79	0	0.04	Upregulated
IGKV2-30	2.62	0	0.04	Upregulated
IGKV1-37	1.82	0.02	0.04	Upregulated
IGKV2-40	1.84	0.01	0.04	Upregulated
IGKV1D-37	2.61	0.02	0.04	Upregulated
IGKV2D-30	2.82	0	0.04	Upregulated
IGKV2D-29	2.45	0.01	0.04	Upregulated
IGKV2D-28	2.71	0.01	0.04	Upregulated
IGKV2D-24	1.2	0.02	0.04	Upregulated
IGKV1D-12	2.45	0	0.04	Upregulated
MRPL30	1.1	0.01	0.04	Upregulated
IL1B	1.12	0.01	0.04	Upregulated
IGKV1OR2-108	1.16	0	0.04	Upregulated
MARCO	1.42	0.03	0.04	Upregulated
FCGR1A	1.25	0.01	0.04	Upregulated
PDK1	1.39	0.02	0.04	Upregulated
CTSS	1.28	0.02	0.04	Upregulated

SERPINE2	1.23	0.01	0.04	Upregulated
MMP9	1.51	0.03	0.05	Upregulated
BCAS4	1.88	0.02	0.04	Upregulated
MEMO1P1	2.29	0	0.01	Upregulated
ITGB2	1.13	0.02	0.04	Upregulated
PIK3CD	1.1	0.03	0.05	Upregulated
IGLV6-57	2.16	0.04	0.05	Upregulated
IGLV1-47	2.48	0.01	0.04	Upregulated
IGLV7-43	1.39	0.03	0.04	Upregulated
IGLV1-40	2.82	0	0.04	Upregulated
IGLV1-36	2.13	0	0.04	Upregulated
IGLV2-23	2.43	0.01	0.04	Upregulated
IGLV3-25	2.19	0.02	0.04	Upregulated
IGLV3-21	2.48	0.02	0.04	Upregulated
IGLV3-9	2.1	0.02	0.04	Upregulated
IGLV3-1	2.44	0.01	0.04	Upregulated
SLAMF7	1.25	0.01	0.04	Upregulated
FCER1G	1.15	0.01	0.04	Upregulated
SEL1L3	1.05	0.01	0.04	Upregulated
JCHAIN	1.92	0.01	0.04	Upregulated
DSPP	1.05	0.01	0.04	Upregulated
SPP1	2.82	0	0.04	Upregulated
SFRP2	1.11	0	0.04	Upregulated
MARCHF1	1.42	0	0.04	Upregulated
GZMK	1.26	0	0.04	Upregulated
MZB1	1.37	0.02	0.04	Upregulated
PTPRC	1.01	0.02	0.04	Upregulated
LY86	1.24	0	0.04	Upregulated
CHI3L1	1.49	0.02	0.04	Upregulated
HBC7	1.07	0.02	0.04	Upregulated
HBC8	1.11	0	0.04	Upregulated
H2AC13	1.08	0.02	0.04	Upregulated
H2AC14	1.1	0.02	0.04	Upregulated
H2BC14	1.05	0.05	0.05	Upregulated
HLA-DRA	1.17	0.01	0.04	Upregulated
HLA-DRB1	1.4	0.01	0.04	Upregulated

HLA-DQA1	1.48	0.01	0.04	Upregulated
HLA-DQB1	1.48	0.05	0.05	Upregulated
HLA-DPB1	1.02	0.01	0.04	Upregulated
TREM2	1.09	0.03	0.05	Upregulated
PRDM1	1.03	0.02	0.04	Upregulated
TRBV20-1	1.19	0.03	0.05	Upregulated
MSR1	1.25	0.01	0.04	Upregulated
CCL19	1.17	0.02	0.04	Upregulated
TNC	1.32	0.01	0.04	Upregulated
CYBB	1.01	0.04	0.05	Upregulated
PIM2	1.06	0.02	0.04	Upregulated
ASAHI2	1.1	0.01	0.04	Upregulated
ACSL5	1.06	0	0.04	Upregulated
SAA2-SAA4	1.13	0.02	0.04	Upregulated
LAPTM5	1.12	0.02	0.04	Upregulated
CD3E	1.15	0.01	0.04	Upregulated
KRT18	1.14	0	0.04	Upregulated
PYM1	1.04	0.01	0.04	Upregulated
LUM	1.15	0	0.04	Upregulated

Table 5.4: List of the 126 Upregulated genes

5.7 Biological Interpretation of Differentially Expressed Genes

To gain insight into the functional implications of differential gene expression between inflammatory and noninflammatory aortic aneurysms, several genes were selected based on their known roles in immune activation, extracellular matrix remodelling, and vascular homeostasis. Their expression changes and associated biological roles are summarised in the table below:

Gene Symbol	Regulation	Biological Function	Interpretation in Context of Aortic Aneurysms
MMP9	Upregulated	Degrades collagen and elastin in ECM	Indicates active matrix remodeling and wall degradation in inflammatory aneurysms
HLA-DRB1	Upregulated	MHC Class II molecule; antigen presentation	Suggests T-cell activation and immune cell infiltration in the vessel wall
GZMB	Upregulated	Cytotoxic granzyme from T and NK cells	Reflects immune-mediated cell damage contributing to wall weakening
S100A9	Upregulated	Inflammatory mediator, DAMP molecule	Enhances neutrophil recruitment and promotes innate immune responses
SLAMF7	Upregulated	NK and plasma cell activation receptor	Marks active immune surveillance and adaptive immune signaling
CYTL1	Downregulated	Anti-inflammatory cytokine-like gene	Suggests suppression of protective anti-inflammatory signaling pathways
SFRP1	Downregulated	Wnt pathway inhibitor and anti-inflammatory protein	Indicates activation of Wnt signaling, promoting inflammation and vascular remodeling
COL6A1	Downregulated	ECM structural protein	Reflects compromised ECM integrity and reduced mechanical strength of aortic wall
CARTPT	Downregulated	Neuropeptide with anti-inflammatory effects	May signal stress-related or immune-associated suppression of homeostasis
ACTA2	Downregulated	Encodes smooth muscle alpha-actin	Reduction indicates SMC loss or dysfunction, impairing wall stability

Table 5.5: Selected DEGs and their biological interpretation

5.7.1 Upregulated Genes in Inflammatory Aneurysms

The upregulation of **MMP9** supports enhanced matrix degradation, a critical event in aneurysm formation, particularly under inflammatory stress. Elevated **HLA-DRB1** expression aligns with infiltration of antigen-presenting cells and activation of adaptive immunity, a known hallmark of inflammatory aneurysms. **Granzyme B (GZMB)** points to active cytotoxic T-cell or NK cell responses, leading to apoptosis of smooth muscle cells or endothelial damage. Similarly, **S100A9**, a pro-inflammatory alarmin, and **SLAMF7**, an immune signalling receptor, highlight the dominance of innate and adaptive immune responses within the lesion.

These patterns confirm a high level of immune activation in inflammatory aneurysms, where immune cells mediate both direct tissue injury and recruitment of additional inflammatory mediators.

5.7.2 Downregulated Genes and Suppressed Protective Mechanisms

Conversely, the downregulation of **CYTL1** and **CARTPT**, both associated with anti-inflammatory signalling, suggests loss of negative regulation over immune responses. **SFRP1**, a Wnt inhibitor, is repressed, potentially allowing for unchecked Wnt pathway activation, which has been linked to vascular inflammation and remodelling. The reduction of **ACTA2**, an SMC-specific gene, indicates smooth muscle cell loss or dedifferentiation, contributing to the structural failure of the aortic wall. Additionally, **COL6A1** downregulation signals weakening of ECM support, increasing the risk of aneurysmal rupture.

5.8 Conclusion

The gene expression patterns observed offer mechanistic insight into the divergent pathways involved in inflammatory vs. noninflammatory aortic aneurysms. Inflammatory aneurysms show a strong immune activation profile with simultaneous suppression of structural and homeostatic elements. This supports the hypothesis that **immune-mediated vascular remodelling, proteolysis, and ECM degradation** are central to disease progression in the inflammatory subtype.

So the transcriptomic profiling conducted in this study has correctly elucidated the differential gene expression panorama between inflammatory and noninflammatory aortic aneurysms. Using the Ballgown package for gene-level analysis, a total of 147

differentially expressed genes (DEGs) were identified, with 126 upregulated and 21 downregulated genes in the inflammatory samples compared to controls. Key genes, inclusive of IL6, CXCL8, TNFAIP3, and MMP9, famous mediators of inflammatory and extracellular matrix remodelling pathways, exhibited significant upregulation, suggesting their pivotal involvement in the pathogenesis of inflammatory aneurysms. Subsequent useful enrichment analysis by the usage of the DAVID revealed that these DEGs have been notably enriched in immune-associated biological techniques and molecular functions. Enriched KEGG pathways included the TNF signalling pathway, IL17 signalling, NF-kappa B signalling, and cytokine-cytokine receptor interaction, all of which are associated with vascular irritation, immune activation, and tissue degradation. And its analysis confirms that differentially expressed genes in inflammatory aortic aneurysms are functionally involved in immune-related pathways, phagocytosis, cytokine signalling, and antigen presentation, providing molecular-level support for the inflammation-driven aetiology of this aneurysm subtype. These findings collectively highlight the existence of a distinct pro-inflammatory transcriptomic signature in the inflammatory subtype of aortic aneurysms. Visualisation tools, including boxplots, volcano plots, and heatmaps, have been employed to assess the quality of RNA-seq data, validate expression differences, and discover clustering styles between the sample groups. Notably, heatmaps of all DEGs and the top 50 most significant genes proved a clear separation between control and treatment groups, reinforcing the robustness of the detected expression differences. Overall, the study offers a comprehensive gene expression atlas that differentiates inflammatory from noninflammatory aortic aneurysms at the transcriptomic level. These findings not only deepen our understanding of aneurysm biology but also point toward potential biomarker candidates and therapeutic targets that may inform personalised treatment strategies in future vascular medicine.

Chapter 6

Conclusion & Future direction

This study aimed to investigate the molecular variations among inflammatory and noninflammatory aortic aneurysms using transcriptomic profiling. Through the analysis of RNA-sequencing records from clinical samples, a complete differential gene expression study was carried out. Initial steps involved normalisation and visualisation of expression data, the usage of boxplots to ensure sure best and comparison between groups. A total of 147 differentially expressed genes (DEGs) have been identified, comprising 126 upregulated and 21 downregulated genes in inflammatory aneurysms. Volcano plots and gene-level boxplots illustrated distinct shifts in gene expression profiles, with genes that include MEMO1P1, SPP1, PKM, E2F4, and PXYLP1 emerging as notable candidates because of their statistical significance and biological relevance.

Hierarchical clustering of all DEGs and the top 50 most extensive genes, visualised through heatmaps, revealed clear group-specific expression patterns that effectively distinguished the inflammatory from the noninflammatory samples. These findings underscore the transcriptional reprogramming related to vascular inflammation and highlight the diagnostic and mechanistic potential of gene expression markers. Functional enrichment analysis using the usage of DAVID tools indicated strong associations with immune-related pathways, which include phagosome formation, cytokine-cytokine receptor interaction, T-cellular receptor signalling, and Fc gamma R-mediated phagocytosis. Genes like IL1RN and HLA-DPB1, further characterised in the study, play key roles in immune regulation, antigen presentation, and inflammatory signalling, reinforcing the idea that immune dysregulation drives the pathology of inflammatory aortic aneurysms. The insights gained from this study provide a solid foundation for understanding the transcriptomic panorama of aortic aneurysms and point in the direction of precise molecular changes underlying inflammation-driven disease progression.

Moving ahead, several future guidelines can enhance the translational value of this work. Firstly, key DEGs identified in this study should be experimentally validated using qPCR, western blotting, or immunohistochemistry to verify their differential expression at the RNA and protein levels in independent sample cohorts. The difference between inflammatory and non-inflammatory aortic aneurysms is of growing clinical and biological interest because of their differing pathological mechanisms and therapeutic

responses (Moll et al., 2011; Nordon et al., 2011). Recent advances in transcriptomic technologies, mainly RNA sequencing, have enabled high-resolution evaluation of gene expression changes that arise in diseased vascular tissues. Numerous studies have been reviewed and included in this take a look at to establish a comprehensive understanding of the molecular underpinnings of aneurysm development and to validate the differentially expressed genes identified in our analysis. These preceding investigations function as the foundation for contextualising our findings, uncovering molecular signatures specific to the inflammatory phenotype, and figuring out ability biomarkers or therapeutic objectives for further investigation. Additionally, integrating transcriptomic facts with other omics layers, along with proteomics, epigenomics, or single-cell transcriptomics, could cover upstream regulators and cell-type-specific signatures contributing to aneurysm inflammation. This multi-omics approach would enable a more comprehensive systems biology view of disease mechanisms. The differentially expressed genes reported here can also serve as the basis for machine learning models designed to predict the inflammatory status of aneurysms. By training classifiers on expression profiles, future research could develop diagnostic tools for early identification and subtype classification of aneurysms. Functionally, genes such as SPP1 and IL1RN, known to influence immune activation and cytokine balance, should be explored further in vitro using knockdown or overexpression studies in vascular cell models. These experiments may elucidate their direct roles in smooth muscle cell remodelling, macrophage infiltration, or endothelial dysfunction (Maguire et al., 2017; Mallat et al., 2001).

Additionally, some of the enriched pathways diagnosed, consisting of phagosome and cytokine receptor interactions, are pharmacologically targetable. This opens avenues for drug repurposing studies geared toward dampening inflammatory responses in aneurysms, the usage of already approved immunomodulatory agents. Development of a targeted biomarker panel comprising excessive-confidence DEGs may also lead to the introduction of non-invasive diagnostic assays, probably the use of blood or serum samples. Such biomarkers should guide manual treatment strategies and decrease the reliance on invasive tissue sampling.

Further, the look can be expanded through investigating demographic-specific gene expression differences, such as those influenced by age, sex, or comorbidities like diabetes and hypertension. This could allow an extra customised medicine approach. Longitudinal

transcriptomic profiling at diverse disorder tiers should provide treasured perception into early vs. late molecular events for the duration of aneurysm development, helping in time-sensitive intervention strategies. Lastly, creating a public transcriptomic useful resource that files the inflammatory vs. Noninflammatory gene expression profiles of aneurysm tissues would extensively advantage the broader vascular studies community. This could sell collaboration, foster meta-analyses, and boost biomarker validation. Collectively, these destiny efforts could pave the way for better diagnosis, improved risk stratification, and development of targeted therapies for inflammatory aortic aneurysms, ultimately contributing to personalised vascular care (Athar et al., 2019).

References

1. Aradhyula, Vaishnavi, Joshua D. Breidenbach, Bella Z. Khatib-Shahidi, Julia N. Slogar, Sonia A. Eyong, Dhilhani Faleel, Prabhatchandra Dube, Rajesh Gupta, Samer J. Khouri, Steven T. Haller, David J. Kennedy. "Transcriptomic Analysis of Arachidonic Acid Pathway Genes Provides Mechanistic Insight into Multi-Organ Inflammatory and Vascular Diseases". *Genes* 2024, 15(7), 954.
2. Behrens, Amelie L., Susanne Dihlmann, Caspar Grond-Ginsbach, Andreas S. Peters, Bernhard Dorweiler, Dittmar Böckler, and Philipp Erhart. "Gene Expression Profiling in Abdominal Aortic Aneurysms". (2022).
3. Bianconi, E., et al. (2019). Transcriptomic profiling and biomarker discovery for aneurysms using RNA sequencing. *Frontiers in Cardiovascular Medicine*, 6, 84.
4. Biros, E., et al. (2015). Transcriptome profiling reveals novel genes associated with aortic aneurysm in humans. *Journal of Vascular Research*, 52(5), 285–298.
5. Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., ... & Mortazavi, A. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biology*, 17(1), 13.
6. Erbel, R., Aboyans, V., Boileau, C., Bossone, E., Bartolomeo, R. D., Eggebrecht, H., ... & ESC Committee for Practice Guidelines. (2014). 2014 ESC Guidelines on the diagnosis and treatment of aortic diseases. *European Heart Journal*, 35(41), 2873–2926.
7. Golledge, J., & Norman, P. E. (2011). Aortic aneurysm pathogenesis: an update and clinical implications. *Journal of the American College of Cardiology*, 58(7), 659–669.
8. Golledge, J., & Norman, P. E. (2011). Atherosclerosis and abdominal aortic aneurysm: cause, response, or common risk factors? *Arteriosclerosis, Thrombosis, and Vascular Biology*, 30(6), 1075–1077.
9. Gu, J., Liu, W., Zhao, Y., et al. (2024). Transcriptomic analysis identifies shared diagnostic biomarkers and immune relationships between atherosclerosis and abdominal aortic aneurysm based on a fatty acid metabolism gene set. *Frontiers in Molecular Biosciences*, 11, 1365447.
10. He, R., Guo, D. C., Sun, W., Papke, C. L., Duraisamy, S., Estrera, A. L., ... & Milewicz, D. M. (2008). Characterization of the aortic wall transcriptome in patients with bicuspid aortic valve. *Journal of Thoracic and Cardiovascular Surgery*, 135(4), 937–945.
11. He, X., Gao, Y., Zhang, X., Zhang, Z., & Zhan, S. (2024). Comprehensive transcriptomic analysis unveils macrophage-associated genes for establishing an abdominal aortic aneurysm diagnostic model. *Journal of Translational Medicine*, 22, Article 87.
12. He, Yu-bin, Hai-zhen Jin, Jin-long Zhao, Chong Wang, Wen-rui Ma, Jie Xing, Xiao-bin Zhang, Yang-yang Zhang, Huang-dong Dai, Nai-shi Zhao, Jian-feng Zhang. "Single-cell

transcriptomic analysis reveals differential cell subpopulations and distinct phenotype transition in normal and dissected ascending aorta". Molecular Medicine volume 28, Article number: 158 (2022).

13. Hiratzka, L. F., et al. (2010). Guidelines for the diagnosis and management of patients with thoracic aortic disease. *Circulation*, 121(13), e266–e369.
14. Howard, D. P., Banerjee, A., Fairhead, J. F., Handa, A., Silver, L. E., & Rothwell, P. M. (2015). Population-based study of incidence and outcome of acute aortic dissection and premorbid risk factor control: 10-year results from the Oxford Vascular Study. *Circulation*, 127(2).
15. Howard, D. P. J., Banerjee, A., & Karthikesalingam, A. (2007). Inflammatory aortic aneurysm: Review of pathophysiology and molecular markers. *Vascular Medicine*, 12(4), 293–298.
16. Illumina. (2023). NEBNext Ultra II RNA Library Prep Kit for Illumina: Protocol Guide. New England Biolabs.
17. Isselbacher, E. M., Preventza, O., Black, J. H., Augoustides, J. G., Beck, A. W., Bolen, M. A., ... & Bavaria, J. E. (2022). 2022 ACC/AHA Guideline for the Diagnosis and Management of Aortic Disease. *Circulation*, 146(24), e334–e482.
18. Johnston, K. W., Rutherford, R. B., Tilson, M. D., Shah, D. M., Hollier, L., & Stanley, J. C. (1991). Suggested standards for reporting on arterial aneurysms. *Journal of Vascular Surgery*, 13(3), 452–458.
19. Johnston, K. W., Rutherford, R. B., Tilson, M. D., Shah, D. M., Hollier, L., & Stanley, J. C. (1991). Suggested standards for reporting on arterial aneurysms. *Journal of Vascular Surgery*, 13(3), 452–458.
20. Kasashima, S., Zen, Y., Kawashima, A., Endo, M., & Matsumoto, Y. (2018). Inflammatory aortic aneurysm: roles of inflammation and immunity. *Pathology International*, 68(7), 361–372.
21. Kastrup, J., Rajkumar, A., Patel, M., & Bandyopadhyay, A. (2022). Global transcriptomic profiling identifies differential gene expression signatures between inflammatory and non-inflammatory aortic aneurysms. *Molecular Medicine*, 28, Article 93.
22. Kays, J. K., Fang, R., Stephens, J. R., Lian, L., & Cheng, P. (2023). Differential expression analyses on human aortic tissue reveal novel genes and pathways associated with abdominal aortic aneurysm onset and progression. *Journal of Vascular Surgery*, 78(2), 521–534.e6.
23. Knappich, Christoph, Joshua M. Spin, Hans-Henning Eckstein, Philip S. Tsao, and Lars Maegdefessel. "Involvement of Myeloid Cells and Noncoding RNA in Abdominal Aortic Aneurysm Disease". (2020).

24. Kuivaniemi, H., Ryer, E. J., Elmore, J. R., & Tromp, G. (2015). Understanding the pathogenesis of abdominal aortic aneurysms. *Expert Reviews in Cardiovascular Therapy*, 13(9), 975–987.
25. Lenk, Guy M., Gerard Tromp, Shantel Weinsheimer, Zoran Gatalica, Ramon Berguer & Helena Kuivaniemi. "Whole genome expression profiling reveals a significant role for immune function in human abdominal aortic aneurysms". *BMC Genomics* volume 8, Article number: 237 (2007).
26. Li, Huimin, Xueyu Hao, Peng Zhang, Jun Guo, Wei Li. "Single-cell RNAseq of Angiotensin II-induced abdominal aortic tissue identifies aneurysm-associated cell clusters in C57BL/6J mice". *Biosci Rep* (2025) 45 (05): 343–360.
27. Li, P., Yang, Y., Dong, Z., et al. (2023). Comprehensive bioinformatics analysis revealed potential key genes and pathways underlying abdominal aortic aneurysm. *Frontiers in Cardiovascular Medicine*, 10, 10665597.
28. Liu, Yihai, Xixi Wang, Hongye Wang, and Tingting Hu. "Identification of key genes and pathways in abdominal aortic aneurysm by integrated bioinformatics analysis" (2019).
29. Nienaber, C. A., & Clough, R. E. (2015). Management of acute aortic dissection. *The Lancet*, 385(9970), 800–811.
30. Pirttimäki, T., Laaksamo, M., Hakkarainen, A., et al. (2017). Differentially expressed genes and canonical pathways in the ascending thoracic aortic aneurysm – The Tampere Vascular Study. *Scientific Reports*, 7, 6329.
31. Pyo, R., et al. (2000). Targeted gene disruption of matrix metalloproteinase-9 (gelatinase B) suppresses development of experimental abdominal aortic aneurysms. *Journal of Clinical Investigation*, 105(11), 1641–1649.
32. Sakalihasan, N., Michel, J. B., Katsargyris, A., Kuivaniemi, H., Defraigne, J. O., Nchimi, A., ... & Powell, J. T. (2018). Abdominal aortic aneurysms. *Nature Reviews Disease Primers*, 4(1), 34.
33. Spin, Joshua M., Mark Hsu, Junya Azuma, Maureen M. Tedesco, Alicia Deng, Justin S. Dyer, Lars Maegdefessel, Ronald L. Dalman, and Philip S. "Transcriptional profiling and network analysis of the murine angiotensin II-induced abdominal aortic aneurysm". (2011).
34. Wang, H., Xu, M., Chen, X., et al. (2023). Global research status analysis of the association between aortic aneurysm and inflammation: A bibliometric analysis (1999–2023). *Frontiers in Cardiovascular Medicine*, 10, 1260935.
35. Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: A revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10(1), 57–63.

36. Yang, Bohan, Yiyuan Xu, Fengfei Yan, Cheng Peng, Ye Song, Song Han & Haiyan Wang. "Identifying nexilin as a central gene in neutrophil-driven abdominal aortic aneurysm pathogenesis". *Molecular Medicine* volume 31, Article number: 120 (2025).
37. Yuan, L., Zhang, Y., Lu, Q., et al. (2024). Single-cell RNA sequencing deconstructs the distribution of immune cells within abdominal aortic aneurysms in mice. *BMC Genomics*, 25, Article 213.
38. Zhang, T., Li, W., Wang, Y., Huang, R., & Sun, J. (2022). Integrating bulk transcriptome and single-cell RNA sequencing data reveals the landscape of the immune microenvironment in thoracic aortic aneurysms. *Frontiers in Cardiovascular Medicine*, 9, 846421.
39. Zu, Hong Lin, Hong Wei Liu & Hai Yang Wang. "Identification of crucial genes involved in pathogenesis of regional weakening of the aortic wall". *Hereditas* volume 158, Article number: 35 (2021).



Antiplagiarism Statement

I understand that plagiarism is a moral offence. I state that all codes/observations/output screenshots in this assignment are my own, and there is no cut/copy-pasted material from undisclosed sources and manipulated to hide the identity, and any codes/observations/output screenshots, if taken from books/websites/colleagues' work, are explicitly indicated with reference.

Name: NIGHITHA T N	Signature:
Batch: MSc P batch	Date:

Note: While writing an anti-plagiarism statement, **make sure that you have,**

- ✓ clearly indicated with double quotes any sentence(s) that you have taken directly from textbooks, journals, websites, or other similar sources, citing the references.
- ✓ not cut and pasted material from undisclosed sources and manipulated to hide the identity.
- ✓ Cited references whenever you have derived from a source.
- ✓ Written all the programs on your own, and wherever code fragments have been utilized from other sources, the same is clearly indicated.
- ✓ Do not include any material provided by fellow students without clearly indicating so.
- ✓ clearly indicate the part of the work done by you, in case of group work.