



Introduction

Crohn's disease (CD) and **Ulcerative Colitis (UC)** are the two main forms of **Inflammatory Bowel Disease (IBD)**, a group of chronic, relapsing inflammatory disorders of the gastrointestinal (GI) tract. CD is a discontinuous inflammatory response that can affect the whole GI tract whereas UC is restricted to the colon and rectum with continuous inflammation of the mucosal layer. The incidences and prevalence of IBD have risen steadily worldwide. Despite advances in treatment options, a significant proportion of patients does not respond or loses response over time, emphasizing the need for a better understanding of molecular mechanisms driving the disease (Porter, 2020). In order to better understand these diseases and how they might be related, we will compare gene expression profiles between CD patients, UC patients and **non-IBD patients (HC)**. Transcriptomic profiling of intestinal tissue offers a powerful way to investigate these mechanisms. Here we analyse the bulk RNA sequencing (RNA-seq) expression data from colon biopsies of the intestinal mucosa of HC, UC patients and CD patients, comparing gene expression between healthy controls and clinical pathology. **We aim to identify genetic changes between these groups and in this way distinguish different molecular pathways.**

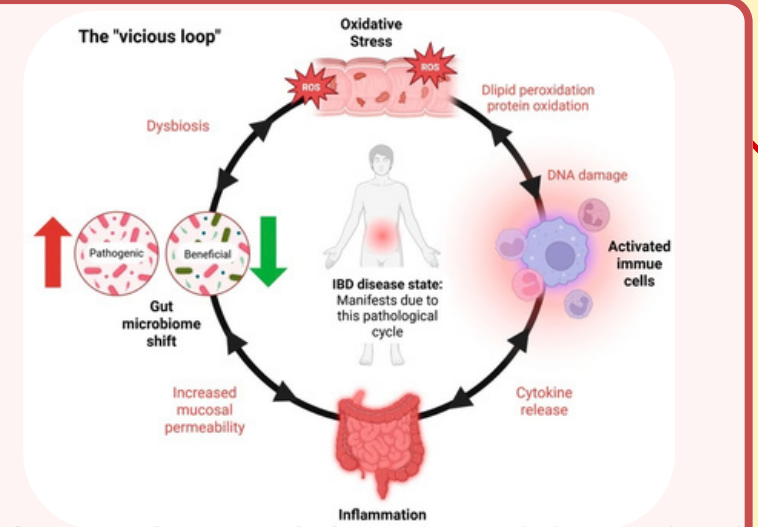


Figure 1. Disease pathology of IBD. Shoham, S. (2025)

Differential Gene Expression UC & CD VS Healthy

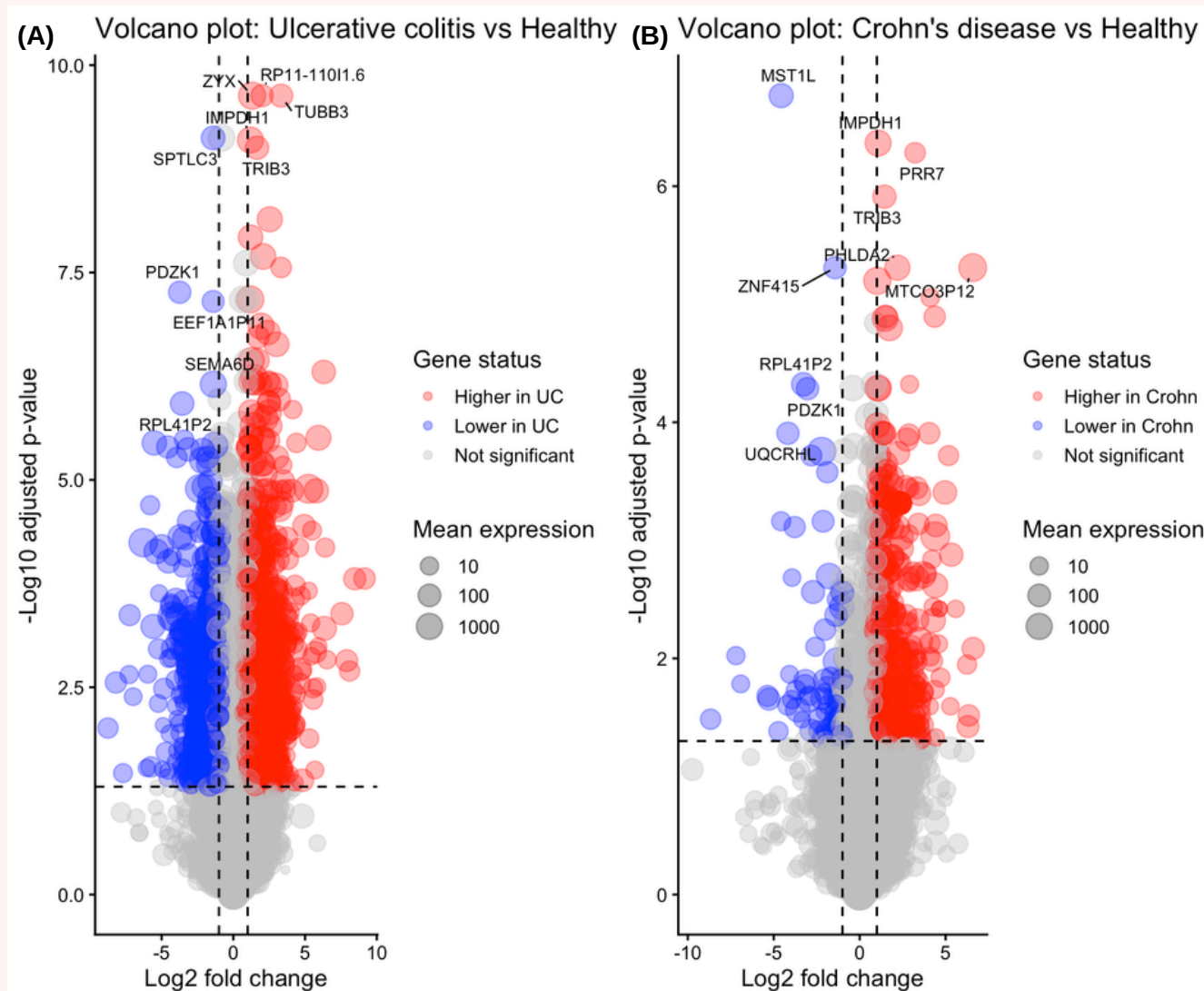


Figure 2. Volcano Plots of differential gene expression. (A) UC compared to healthy & (B) CD compared to healthy. ZYX, RP11-11011.6, TUBB3, IMPDH1 and TRIB3 are significantly most upregulated in UC, SPTLC3, PDZK1, EEF1A1P1, SEMA6D and RPL41P2 are significantly most downregulated compared to healthy. PRR7, IMPDH1, PHLDA2, MTCO3P12 and TRIB3 are significantly most upregulated in CD, MST1L, ZNF415, PDZK1, UQCRHL and RPL41P2 are significantly most downregulated compared to healthy (adjusted $p < 0.05$ and $|\log_2 \text{fold change}| > 1$). X-axis indicates the change in gene expression ($\log_2 \text{fold change}$), Y-axis indicates significance ($-\log_{10} \text{adjusted } p\text{-value}$), sphere size indicates mean expression.

Data preparation

Data sets from NCBI (2023) were used for analysis. The libraries were prepared from a total RNA TruSeq stranded mRNA kit from Illumina according to manufacturers instructions. 21 Samples out of 56 samples from the total sample list were selected: 7 healthy controls, 7 with Ulcerative Colitis, and 7 with Crohn's disease. All biopsies were taken from the sigmoid colon. The groups were assigned to "Healthy", "UC", and "CD". Groups were normalised and insignificant genes were filtered out. The remaining genes from the diseased groups were then compared to the healthy control. Rstudio was used for data analysis and visualisation.

Heatmap

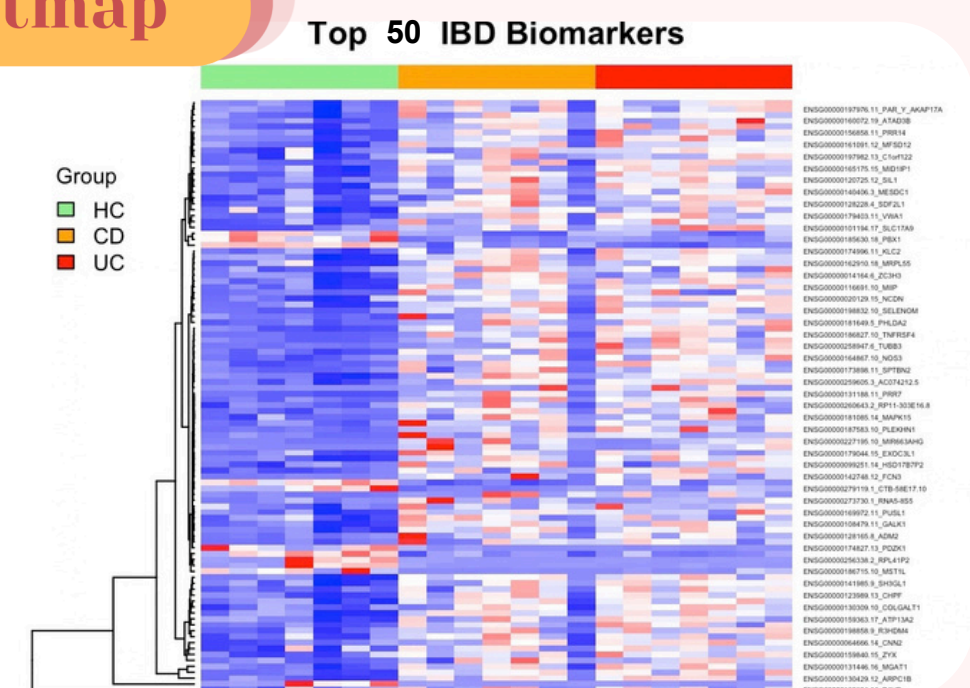


Figure 3. Heatmap, indicating upregulated (red) and downregulated (blue) genes in healthy controls, Crohn's disease and Ulcerative colitis. The top 100 most significant genes ($p=0.05$) were filtered by $\text{cpms} > 1$ and only included if expressed in all patients per group.

Overlap and Distinctions in UC and CD

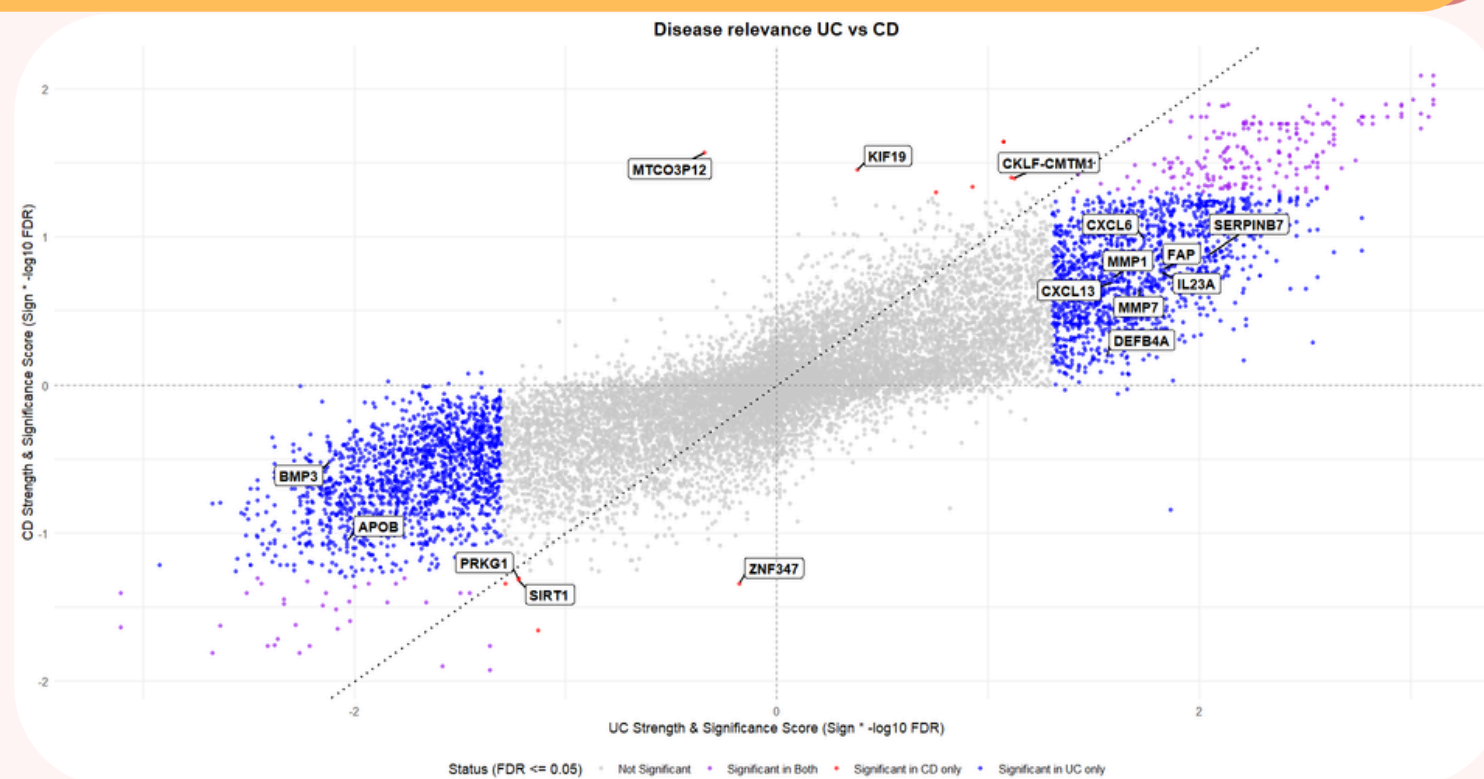


Figure 4. Scatterplot with highlighted genes of interest for UC (blue), CD (red) and overlapping (purple). In UC, upregulation of genes involved in ECM breakdown and fibrosis like MMP1, MMP7, FAP and pro-inflammatory cytokines IL23a, chemoattractants CXCL6 and CXCL13 was observed. In CD a downregulation of pro-barrier genes like SIRT1 and PRKG1 was observed. Overlapping genes showed no disease relevant pathways. Normalization with TMM analysis and quasilielihood test were performed and Disease relevance of UC and CD compared to healthy controls was scored and compared based on expression levels and FDR rate ($p=0.05$). g:profiler was used to identify disease relevant pathways.

Pathway analysis

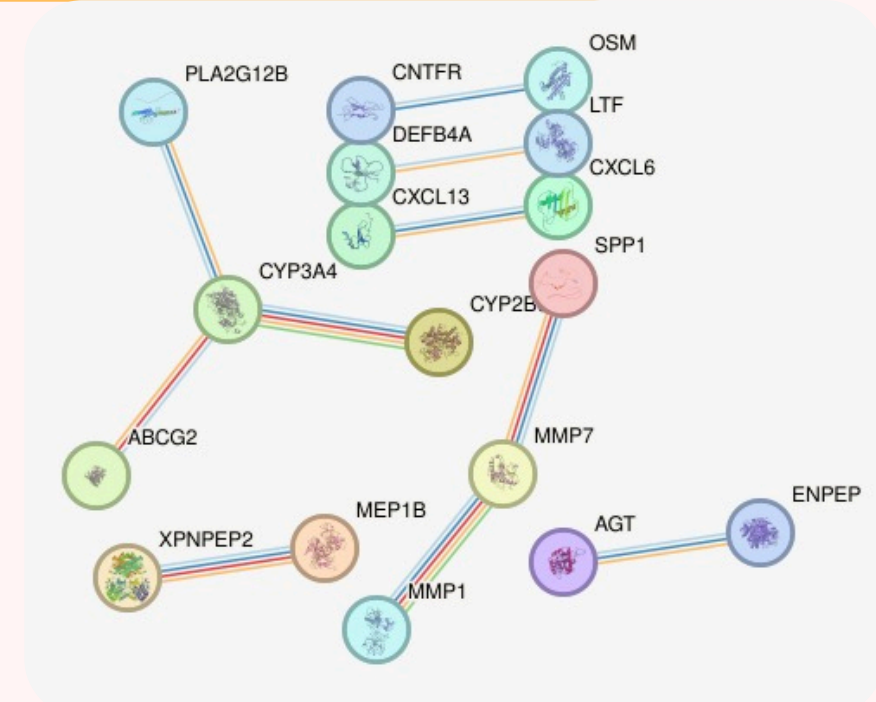


Figure 5. STRING functional association network of IBD-associated genes. The network shows 17 of 53 genes identified in the analysis, including 5 genes associated with CD and 48 genes associated with UC (minimum interaction score 0.7). Nodes represent proteins with highlighted clusters from the following functional pathways: peptidase activity & extracellular matrix degradation (MMP7, MMP1, SPP1, MEPIB), inflammation and immune response (CXCL6&13, OSM, DEFB4A), and the renin-angiotensin system (AGT, ENPEP). The clusters involved in extracellular matrix degradation, inflammation and immune response are particularly relevant to IBD, as these processes are associated with tissue damage and intestinal inflammation.

Conclusion

From the differential gene expression we found that the IMPDH1 and TRIB3 genes are both among the five most significantly upregulated in UC and CD compared to healthy. Of these genes, TRIB3 has been shown to link inflammatory and metabolic signals to cell survival (LifeMap Sciences, 2026). The PDZK1 and RPL41P2 genes are both among the five most significantly downregulated in UC and CD compared to healthy. PDZK1 plays a role in cholesterol homeostasis and has been linked to metabolic syndrome (LifeMap Sciences, 2026). Furthermore, the disease comparison analyses showed a downregulation of anti-inflammatory SIRT1 and pro-barrier protein PRKG1 in CD while in UC an upregulation of chemoattractants CXCL6- and -13, pro-inflammatory cytokine IL23A and upregulation of genes involved in ECM breakdown and fibroblast activation FAP, MMP1 & MMP7. The upregulation of genes involved in pro-inflammatory processes and protein breakdown with a heightened immune response to bacterial products shows us that there is likely a dysfunction in intestinal barrier integrity which contributes to a chronic inflammatory cycle. Our data suffered from limitations in sample size while heavily relying on large colon biopsy collections, future analyses should be performed with biopsies from the small intestine as UC exclusively presents in the large colon but Crohn's disease presents throughout the entire GI tract.

REFERENCES

- Porter, R. J., Kalla, R., & Ho, G. T. (2020). Ulcerative colitis: Recent advances in the understanding of disease pathogenesis. F1000Research, 9, F1000 Faculty Rev-294. <https://doi.org/10.12688/f1000research.20805.1>
- Garrido-Trigo, A., & Salas, A. (2023). Bulk RNAseq expression data of colon biopsies from intestinal mucosa of non-IBD controls and patients with Crohn's disease or ulcerative colitis [Data set]. NCBI Gene Expression Omnibus. <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE235236>
- Shoham, S., Pintel, N., & Avni, D. (2025). Oxidative Stress, Gut Bacteria, and Microalgae: A Holistic Approach to Manage Inflammatory Bowel Diseases. Antioxidants, 14(6). <https://doi.org/10.3390/ANTIOX14060697>
- LifeMap Sciences, Inc. Version 6.1. (z.d.). TRIB3. GeneCards. Accessed on the 8th of 2026, of <https://www.genecards.org/card/TRIB3>
- LifeMap Sciences, Inc. Version 6.1. (z.d.). PDZK1. GeneCards. Accessed on the 8th of 2026, of <https://www.genecards.org/card/PDZK1>

Zoë Hamstra S5173973

Ilse Wals S5047390

Gijs Hoogendoorn S6130372

Rob Franzel S5257506

Jade Hofmeester S4976770

Niels Tuit S4744373



university of
groningen

COPD and disease stages: a comparative analysis of disease stages

Andriana Panoutsou, Tania Agathangelidi Trakatelli, Fianne Kroes, Jason Tsarmaklis, Thom Vaes and Meike Wurkum

Based on data from: Kim, W. J., Lim, J. H., Lee, J. S., Lee, S.-D., Kim, J. H., & Oh, Y.-M. (2015). Comprehensive Analysis of Transcriptome Sequencing Data in the Lung Tissues of COPD Subjects. International Journal of Genomics, 2015, 1–9. <https://doi.org/10.1155/2015/206937>

STAGE 0 VS. STAGE 1

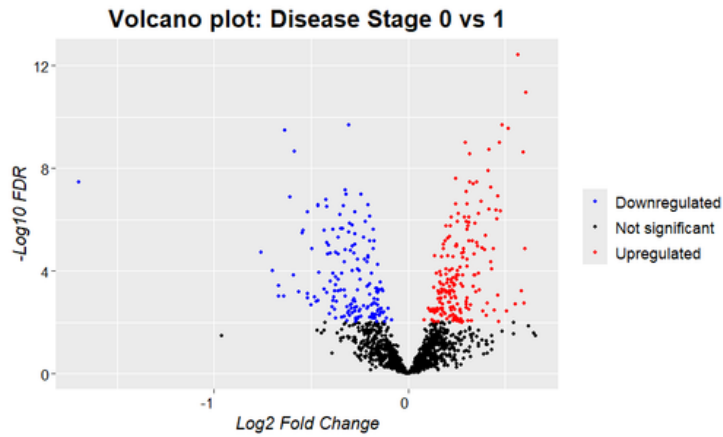


Figure 1. Volcano plot comparing gene expression between stage 0 and stage 1 COPD. Each point represents a gene, with the horizontal axis showing the change in gene expression (log2 fold change) and the vertical axis showing the statistical significance of the change ($-\log_{10}$ FDR). Values to the right of $x = 0$ indicate higher expression in disease state 1, whereas values to the left of $x = 0$ indicate lower expression in disease state 1. Higher points indicate lower FDR values and therefore stronger statistical evidence for differential gene expression. Red and blue points represent significantly upregulated and downregulated genes, respectively ($FDR \leq 0.01$), while black points represent genes that are not significantly differentially expressed.

Profiling of upregulated genes

ID	Source	Term ID	Term Name	P_{adj} (query_1)
1	GO:MF	GO:0003712	transcription coregulator activity	6.109×10^{-5}
2	GO:MF	GO:0061629	RNA polymerase II-specific DNA-binding transcript...	4.102×10^{-3}
3	GO:MF	GO:0015662	P-type ion transporter activity	1.099×10^{-2}
4	GO:BP	GO:0048522	positive regulation of cellular process	1.022×10^{-2}
5	GO:BP	GO:0003407	neural retina development	1.898×10^{-2}
6	GO:BP	GO:0061061	muscle structure development	1.705×10^{-2}
7	GO:BP	GO:0009895	negative regulation of catabolic process	1.866×10^{-2}
8	GO:BP	GO:0060235	lens induction in camera-type eye	2.479×10^{-2}
9	GO:BP	GO:0002026	regulation of the force of heart contraction	3.690×10^{-2}
10	GO:CC	GO:0005815	microtubule organizing center	5.442×10^{-4}
11	GO:CC	GO:0099080	supramolecular complex	6.033×10^{-4}
12	GO:CC	GO:0034967	Set3 complex	9.178×10^{-4}
13	GO:CC	GO:0070210	Rpd3L-Expanded complex	9.178×10^{-4}
14	GO:CC	GO:0016020	membrane	1.225×10^{-2}
15	GO:CC	GO:0005654	nucleoplasm	2.045×10^{-2}
16	GO:CC	GO:0030315	T-tubule	3.008×10^{-2}

Table 1: Upregulated genes between stage 0 and stage 1 COPD, generated in g:Profiler. P-values in the table output are color-coded from yellow (low significance) to green (highly significant), shown in column ' P_{adj} '.

Profiling of downregulated genes

ID	Source	Term ID	Term Name	P_{adj} (query_1)
1	GO:MF	GO:0003735	structural constituent of ribosome	3.531×10^{-11}
2	GO:MF	GO:0003723	RNA binding	3.479×10^{-6}
3	GO:MF	GO:1990948	ubiquitin ligase inhibitor activity	5.330×10^{-4}
4	GO:MF	GO:0045296	cadherin binding	2.153×10^{-2}
5	GO:MF	GO:0042605	peptide antigen binding	3.304×10^{-2}
6	GO:BP	GO:0002181	cytoplasmic translation	1.000×10^{-16}
7	GO:BP	GO:0042254	ribosome biogenesis	5.172×10^{-6}
8	GO:CC	GO:0022626	cytosolic ribosome	8.726×10^{-22}
9	GO:CC	GO:0070062	extracellular exosome	8.453×10^{-6}
10	GO:CC	GO:0042611	MHC protein complex	3.421×10^{-3}
11	GO:CC	GO:0016020	membrane	4.298×10^{-2}

Table 2: Downregulated genes between stage 0 and stage 1 COPD, generated in g:Profiler. P-values in the table output are color-coded from yellow (insignificant) to green (highly significant), shown in column ' P_{adj} '.

INTRODUCTION

- Chronic obstructive pulmonary disease (COPD) is a lung disease characterized by a restriction in airflow and breathing problems
- The dataset used for this analysis is the largest lung tissue sample collection, gathered from cancer patients undergoing lung transplantation
- We attempted to identify molecular pathways that are implicated in specific stages of COPD, by performing gene expression profiling of lung tissue

STAGE 1 VS. STAGE 2

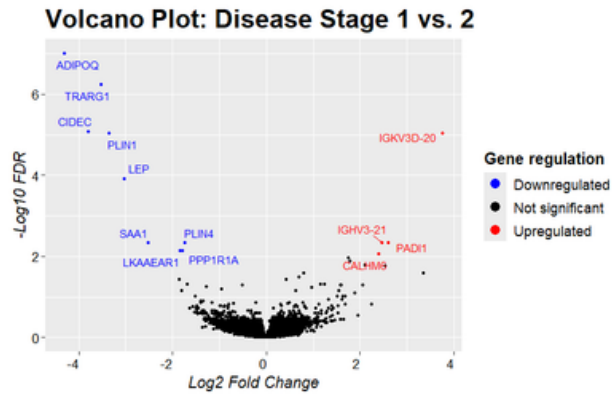


Figure 2. Volcano plot comparing gene expression between stage 1 and stage 2 COPD. Each point represents a gene, with the horizontal axis showing the change in gene expression (log2 fold change) and the vertical axis showing the statistical significance of the change ($-\log_{10}$ FDR). Values to the right of $x = 0$ indicate higher expression in disease state 2, whereas values to the left of $x = 0$ indicate lower expression in disease state 2. Higher points indicate lower FDR values and therefore stronger statistical evidence for differential gene expression. Red and blue points represent significantly upregulated and downregulated genes, respectively ($FDR \leq 0.01$), while black points represent genes that are not significantly differentially expressed.

Profiling of downregulated genes

ID	Source	Term ID	Term Name	P_{adj} (query_1)
1	GO:BP	GO:0098708	D-glucose import across plasma membrane	9.043×10^{-4}
2	GO:BP	GO:0016042	lipid catabolic process	2.523×10^{-3}
3	GO:CC	GO:0005811	lipid droplet	8.324×10^{-4}

Table 3: Downregulated genes between stage 1 and stage 2 COPD, generated in g:Profiler. P-values in the table output are color-coded from yellow (insignificant) to green (highly significant), shown in column ' P_{adj} '.

CONCLUSIONS

Stage 0 vs. 1

Comparing gene expression between stage 0 and stage 1 showed significant difference in regulation of certain pathways related to RNA binding and ribosomal translation. With this in mind, we concluded that changes in the expression of genes are important in the development of COPD

Stage 0 vs. 2

Comparing gene expression between stage 0 and stage 2 COPD showed similar altered gene expression to the comparison between stage 0 and stage 1. Pathways that were similarly altered in Ov1 and Ov2 include: Downregulation of ribosomal, RNA-binding, and translational pathways. Upregulation of positive regulation of cellular pathways and RNA polymerase. In addition, there are stage 2-specific pathways that are upregulated and downregulated (e.g., stress granule assembly).

Stage 1 vs. 2

The number of significant differentially expressed genes between the two different diseased Stages was low (13 genes total). Comparing these genes to the common appearance of others in Stages Ov1 and Ov2 show no commonality, possibly indicating that the extracted dataset is too small to be used in a meaningful comparison or that differences in expression might be either minimal or close to the significance threshold.

Take home message

Comparison of significantly differentially expressed genes across the three COPD populations found similarities in upregulation and downregulation of certain pathways only in the Ov1 to Ov2 comparison. This would indicate that there are certain genes, especially ones involved in RNA binding and ribosomal translation, which are implicated in COPD development and are not observed in healthy individuals. The differences noted between Ov1 and Ov2 could imply that there are genes which increase exponentially with disease progression and thus are not present during mild stage of COPD.

STAGE 0 VS. STAGE 2

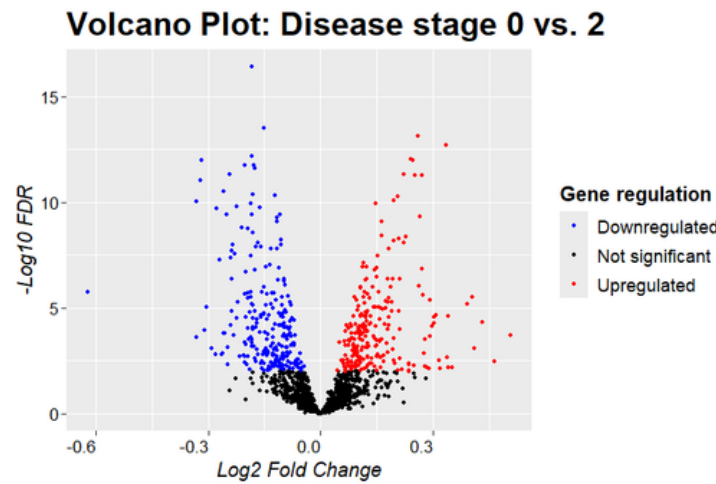


Figure 3. Volcano plot comparing gene expression between stage 0 and stage 2 COPD. Each point represents a gene, with the horizontal axis showing the change in gene expression (log2 fold change) and the vertical axis showing the statistical significance of the change ($-\log_{10}$ FDR). Values to the right of $x = 0$ indicate higher expression in disease state 2, whereas values to the left of $x = 0$ indicate lower expression in disease state 2. Higher points indicate lower FDR values and therefore stronger statistical evidence for differential gene expression. Red and blue points represent significantly upregulated and downregulated genes, respectively ($FDR \leq 0.01$), while black points represent genes that are not significantly differentially expressed.

Profiling of upregulated genes

ID	Source	Term ID	Term Name	P_{adj} (query_1)
1	GO:MF	GO:0030674	protein-macromolecule adaptor activity	3.269×10^{-4}
2	GO:MF	GO:0061629	RNA polymerase II-specific DNA-binding transcript...	1.341×10^{-3}
3	GO:MF	GO:0140945	histone H3K4 monomethyltransferase activity	3.956×10^{-2}
4	GO:MF	GO:0003677	DNA binding	4.944×10^{-2}
5	GO:BP	GO:0048522	positive regulation of cellular process	5.821×10^{-6}
6	GO:BP	GO:0051716	cellular response to stimulus	1.272×10^{-2}
7	GO:BP	GO:0009894	regulation of catabolic process	1.558×10^{-2}
8	GO:BP	GO:0048468	cell development	2.226×10^{-2}
9	GO:BP	GO:0061061	muscle structure development	3.514×10^{-2}
10	GO:BP	GO:0045893	positive regulation of DNA-templated transcription	3.782×10^{-2}
11	GO:BP	GO:2001237	negative regulation of extrinsic apoptotic signaling...	4.577×10^{-2}
12	GO:BP	GO:0034063	stress granule assembly	4.936×10^{-2}
13	GO:CC	GO:0005654	nucleoplasm	2.532×10^{-4}
14	GO:CC	GO:0005829	cytosol	1.022×10^{-3}
15	GO:CC	GO:0032991	protein-containing complex	2.655×10^{-3}
16	GO:CC	GO:0043232	intracellular membraneless organelle	4.896×10^{-3}
17	GO:CC	GO:0044666	MLL3/4 complex	4.231×10^{-2}

Table 4: Upregulated genes between stage 0 and stage 2 COPD, generated in g:Profiler. P-values in the table output are color-coded from yellow (insignificant) to green (highly significant), shown in column ' P_{adj} '.

Profiling of downregulated genes

ID	Source	Term ID	Term Name	P_{adj} (query_1)
1	GO:MF	GO:0003735	structural constituent of ribosome	2.467×10^{-10}
2	GO:MF	GO:0003723	RNA binding	5.189×10^{-9}
3	GO:BP	GO:0002181	cytoplasmic translation	6.211×10^{-16}
4	GO:BP	GO:0042274	ribosomal small subunit biogenesis	2.246×10^{-3}
5	GO:CC	GO:0022626	cytosolic ribosome	1.176×10^{-20}
6	GO:CC	GO:0070062	extracellular exosome	7.452×10^{-7}
7	GO:CC	GO:0012505	endomembrane system	1.837×10^{-4}
8	GO:CC	GO:0016020	membrane	4.155×10^{-2}

Table 5: Downregulated genes between stage 0 and stage 2 COPD, generated in g:Profiler. P-values in the table output are color-coded from yellow (low significance) to green (highly significant), shown in column ' P_{adj} '.

Introduction

Cell proliferation is essential for tissue growth, maintenance, and repair. When cells stop proliferating, they can enter different states of cell-cycle arrest, including quiescence and senescence (Fig. 1). Quiescence is a reversible resting state, whereas senescence is generally considered a stable state of proliferative arrest (Marescal & Cheeseman, 2020; Wang et al., 2024). Although both quiescence and senescence are characterised by a lack of proliferation, they are distinct cellular states with different molecular and functional characteristics. Understanding the differences between these states is important for determining how cells transition from reversible cell-cycle arrest towards a more stable senescent state. This raises the question of whether proliferating, quiescent, and irradiation-induced senescent cells also have distinct gene-expression profiles. To investigate this, gene-expression data from proliferating, quiescent, and irradiation-induced senescent human foreskin fibroblasts from Hernandez-Segura et al. (2017) were compared using PCA, heatmap and Gene Ontology (GO) analysis. The aim was to identify transcriptional differences and biological processes that distinguish proliferating, quiescent, and irradiation-induced senescent cells, which may help identify markers and mechanisms involved in the transition towards senescence (Terzi et al., 2016).

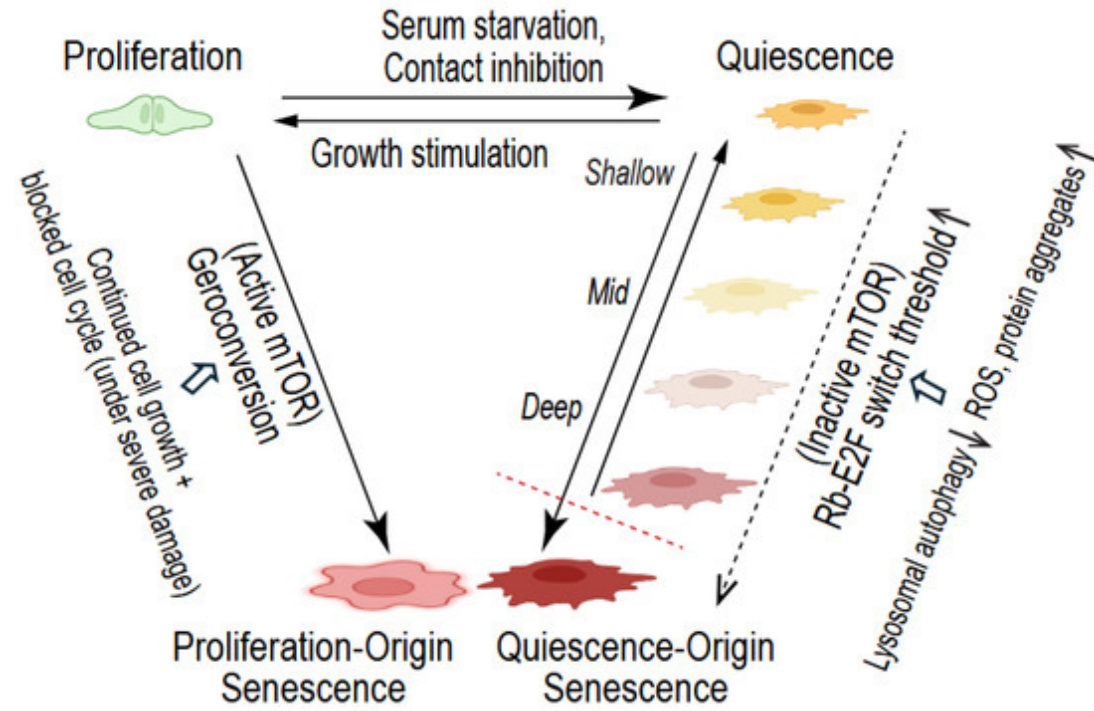


Figure 1. Proliferation and Quiescence balance (Terzi et al., 2016).

PCA plot

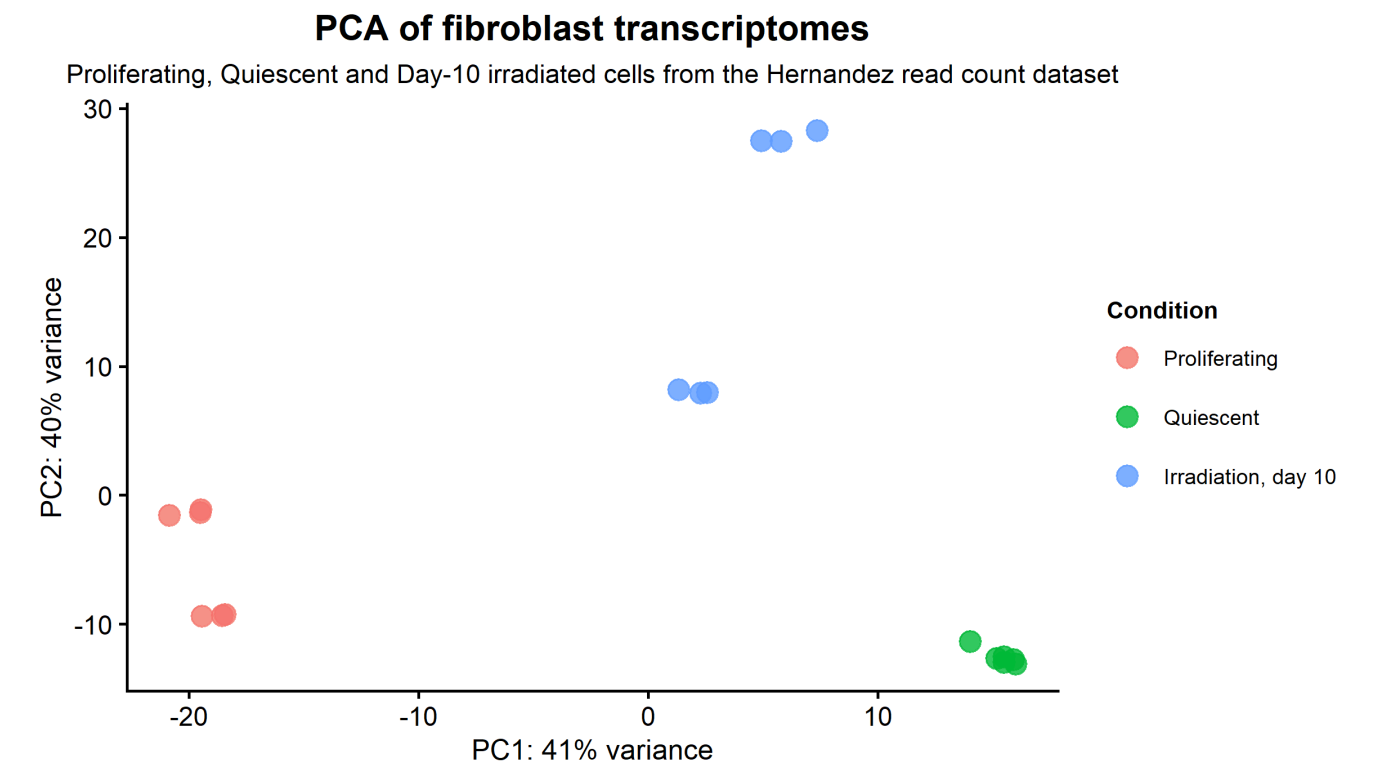


Figure 2. Principal component analysis (PCA) of the Hernandez read-count dataset of foreskin fibroblasts under proliferating, quiescent, and irradiation at day 10 conditions. Data were filtered to retain genes with at least 10 reads in at least 6 samples, irrespective of condition. From this list, the top 500 most variable genes were selected to perform PCA. Each point represents a sample from the Hernandez dataset. PC1 explains 41%, while PC2 explains 40% of the total variance. Colors indicate the experimental condition. The PCA was performed using R.

GO enrichment

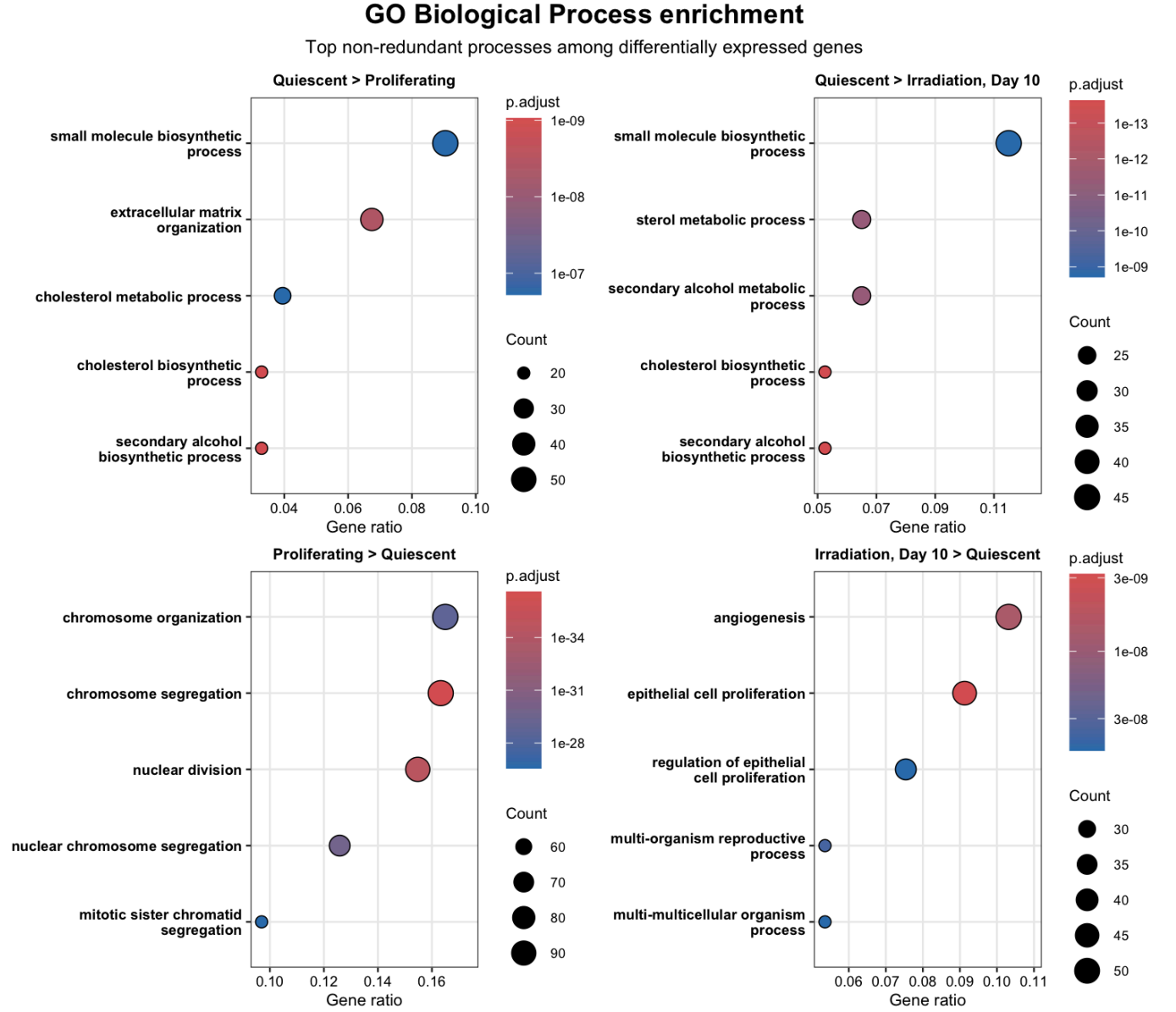


Figure 4. GO Biological Process enrichment of differentially expressed genes. GO enrichment analysis was performed separately for genes significantly upregulated in each pairwise comparison (adjusted $p < 0.05$, $|\log_2FC| > 1$). Gene ratio represents the proportion of genes in the analyzed gene set associated with each GO term, dot size indicates the number of contributing genes, and color represents the adjusted p -value. Quiescent cells showed enrichment of biosynthetic, cholesterol/sterol metabolic, and extracellular-matrix processes, whereas proliferating cells were strongly enriched for chromosome organization and cell-division processes. Compared with quiescence, Day-10 irradiated cells showed a distinct enrichment profile, including angiogenesis and epithelial-proliferation-associated processes. GO enrichment was performed using R.

Heatmap

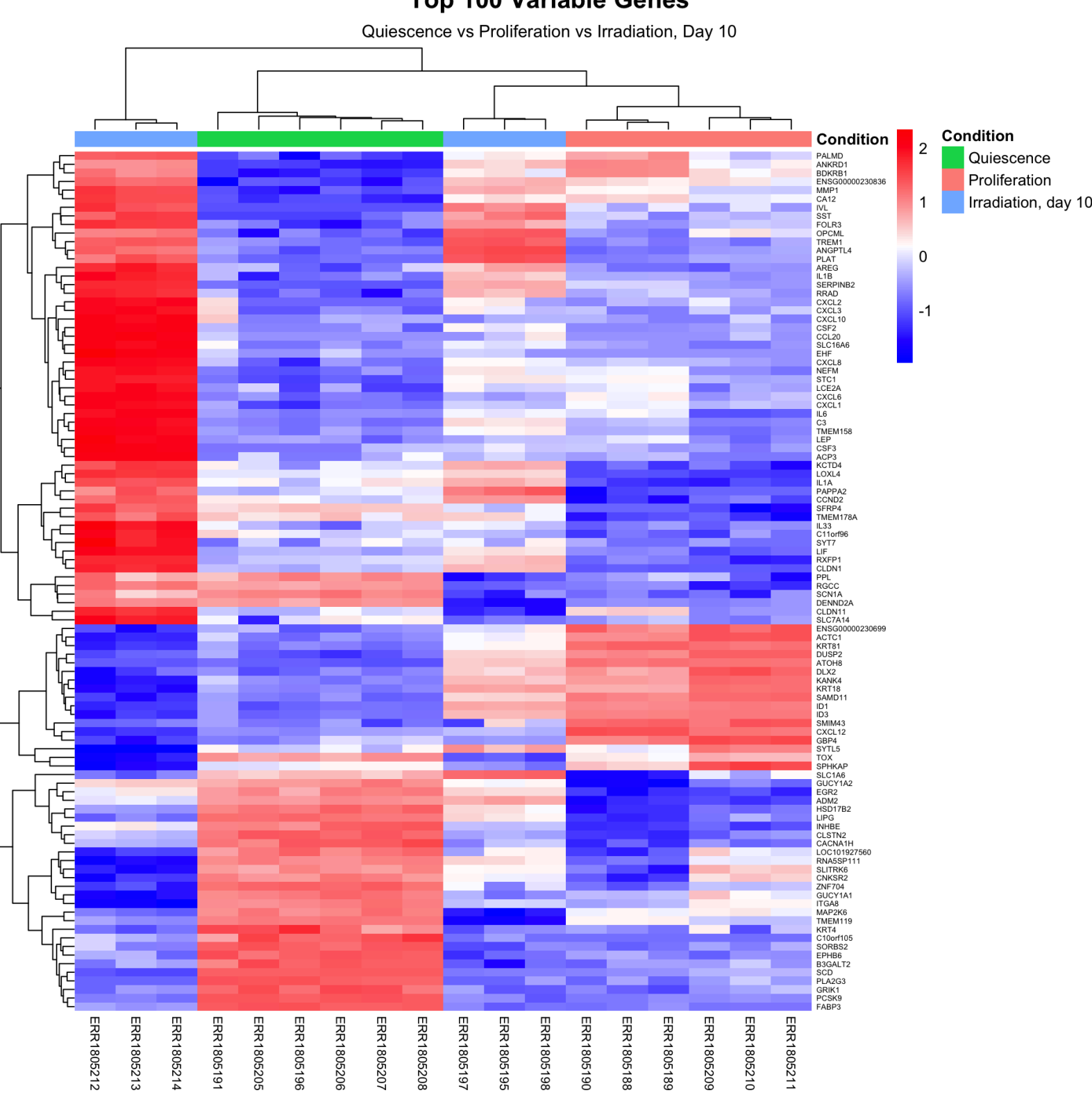


Figure 3. Comparison of gene-expression patterns between quiescent, proliferating, and irradiation day 10 foreskin fibroblasts. Based on the PCA, quiescent cells clustered more closely with a subset of the irradiation day 10 cells. However, the heatmap analysis of the top 100 variable genes did not show a clear similarity in gene-expression patterns between these two groups, as illustrated by the green column and the second blue column. Proliferating cells were included as a control and displayed a distinct expression pattern compared with the quiescent and irradiation day 10 cells. Overall, these results are consistent with the PCA, which indicated that the three cellular states have distinct transcriptional profiles. The heatmap was performed using R.

Conclusion

PCA analysis shows distinct transcriptional profiles between proliferating, quiescent and irradiation-treated fibroblasts. Proliferating and quiescent samples formed clearly separated clusters, while irradiation samples (day 10) separated into two subclusters that nevertheless remained distinct from both proliferating and quiescent cells. The heatmap confirms distinct gene-expression patterns between the cell states, while also showing some heterogeneity within the irradiation group. The key findings of the GO data analysis highlight the differences between quiescence, proliferative, and senescent cells. The main difference shows between proliferating and quiescent fibroblasts a strong activation of cell-cycle and mitotic processes. Although quiescent cells are no longer proliferating they still remain metabolically active. Genes in processes like sterol metabolics, small molecule, cholesterol and secondary alcohol biosynthesis remain expressed. Senescent fibroblasts were enriched for processes related to tissue signaling. Our results indicate that quiescent fibroblasts represent a distinct state: they have reduced cell-cycle activity compared to proliferating cells, while remaining metabolically active, but lacking the tissue-signalling phenotype.

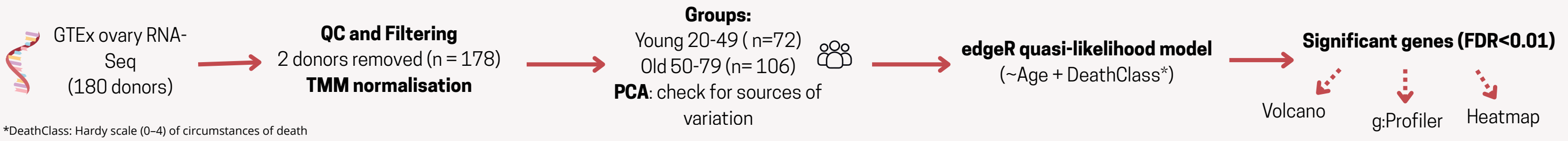
INTRODUCTION

The ovaries play essential roles in fertility and hormone production and are among the earliest organs to show age-related functional deterioration. Female fertility decline accelerates around 35 years of age. Natural menopause typically occurs between 45 and 55 years, though influenced by genetic, hormonal, and environmental factors (Cavalcante et al., 2023; Liu et al., 2025). Declining ovarian function is also associated with broader health consequences, including increased risks of osteoporosis and cardiovascular disease (Liu et al., 2025).

At the cellular and molecular levels, ovarian ageing involves extensive alterations, including fibrosis, inflammation, oxidative stress, immune-cell changes, and altered gene expression (Isola et al., 2024; Wang et al., 2020). Although recent studies have begun to characterise human ovarian ageing using single-cell RNA sequencing and spatial transcriptomic approaches (Jin et al., 2025; Wu et al., 2024), our understanding of age-associated molecular changes in the human ovary remains limited.

We leveraged bulk RNA-seq data from the GTEx project to characterise age-associated transcriptional differences in the human ovary and identify biological processes potentially involved in ovarian ageing.

METHODS



RESULTS

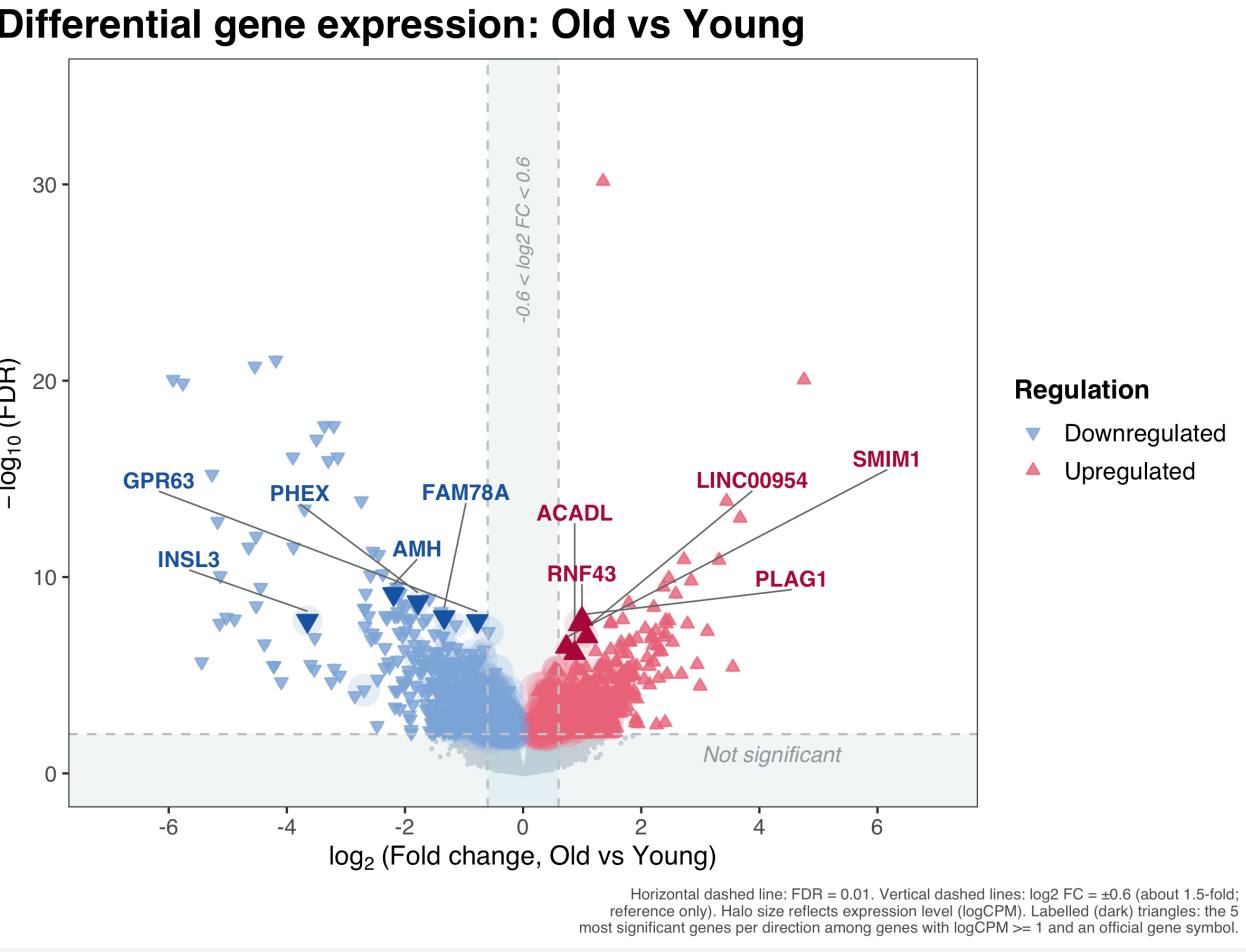


Figure 1. Ageing was associated with both up- and downregulation of gene expression in the ovary. Volcano plot of bulk RNA sequencing (RNA-seq) data from GTEx ovary donors aged 50-79 years (Old, n = 106) vs 20-49 years (Young, n = 72), analysed with an edgeR quasi-likelihood model adjusted for cause of death (DeathClass, Hardy scale) (26,448 genes tested). x-axis: log₂ fold change (Old vs Young); y-axis: -log₁₀ FDR (Benjamini-Hochberg). Genes with FDR ≤ 0.01 (horizontal dashed line) were considered differentially expressed: 496 were higher (pink ▲), and 593 were lower expressed (blue ▼) in older donors; grey dots below the line were not significant. Vertical dashed lines (log₂FC = ±0.6, ~1.5-fold) and the vertical shaded band are visual references only and were not used to define differential expression. Halo size reflects expression level (logCPM). The five most significant genes in each direction with logCPM ≥ 1 were labelled (larger, darker triangles) and ranked by p-value, with |log₂FC| used to break ties. FDR, false discovery rate; FC, fold change; CPM, counts per million; GTEx, Genotype-Tissue Expression.

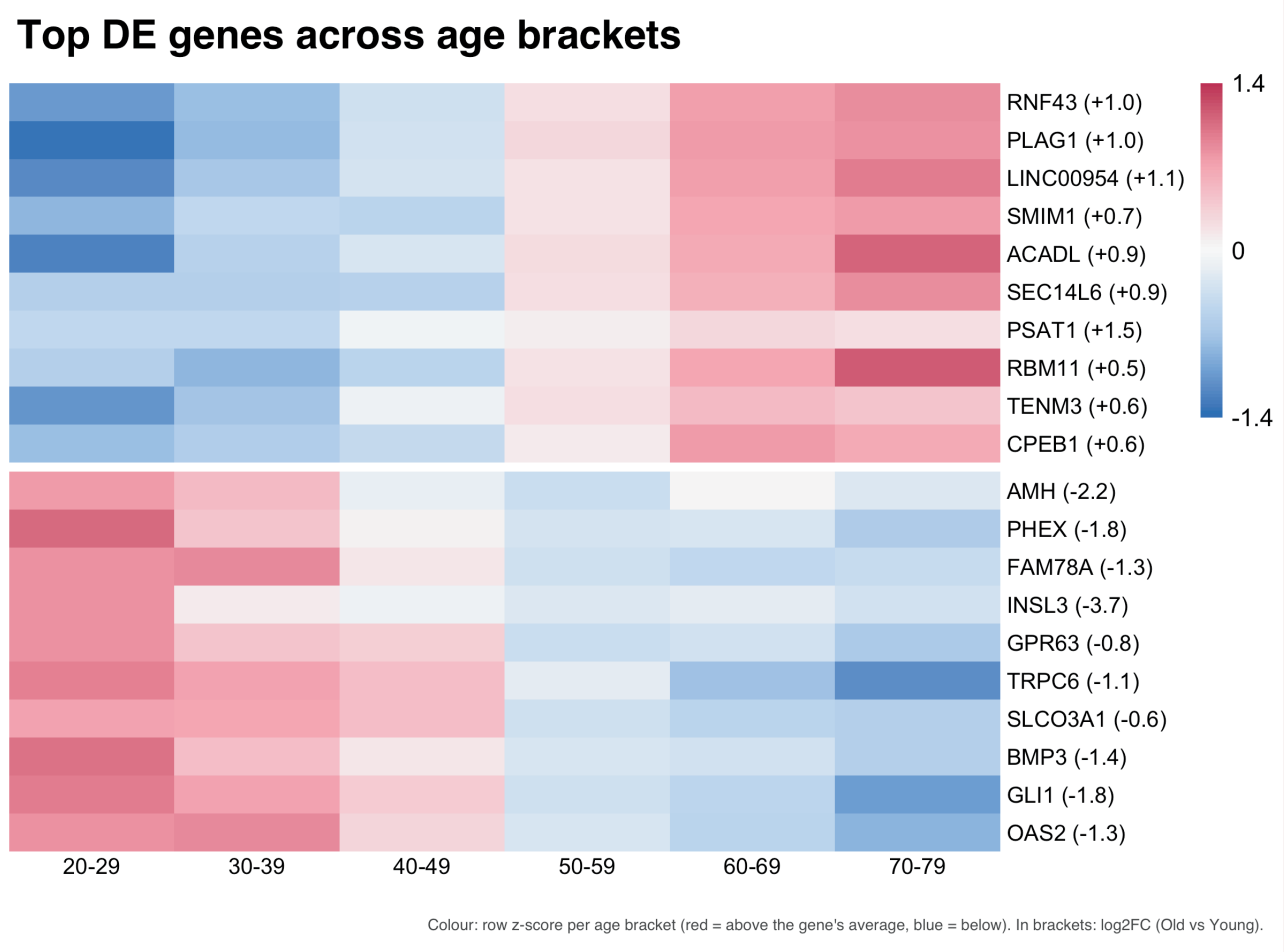


Figure 2. Top differentially expressed genes changed gradually across age brackets. Heatmap of the 20 most significantly differentially expressed (DE) genes: 10 with higher expression in older donors (top block) and 10 with lower expression (bottom block). Genes were selected among those with FDR ≤ 0.01 and logCPM ≥ 1, and ranked by p-value. Numbers in brackets indicate log₂ fold changes (Old, 50-79 years vs Young, 20-49 years). Each column is a donor age bracket (20-29; n = 22; 30-39; 14; 40-49; 36; 50-59; 60; 60-69; 42; 70-79; 4). Colour represents the mean row z-score of TMM-normalised log₂ CPM for each age bracket: red indicates expression above the gene's average and blue indicates expression below it. The five most significant genes in each direction are also labelled in the volcano plot. FDR, false discovery rate; CPM, counts per million.

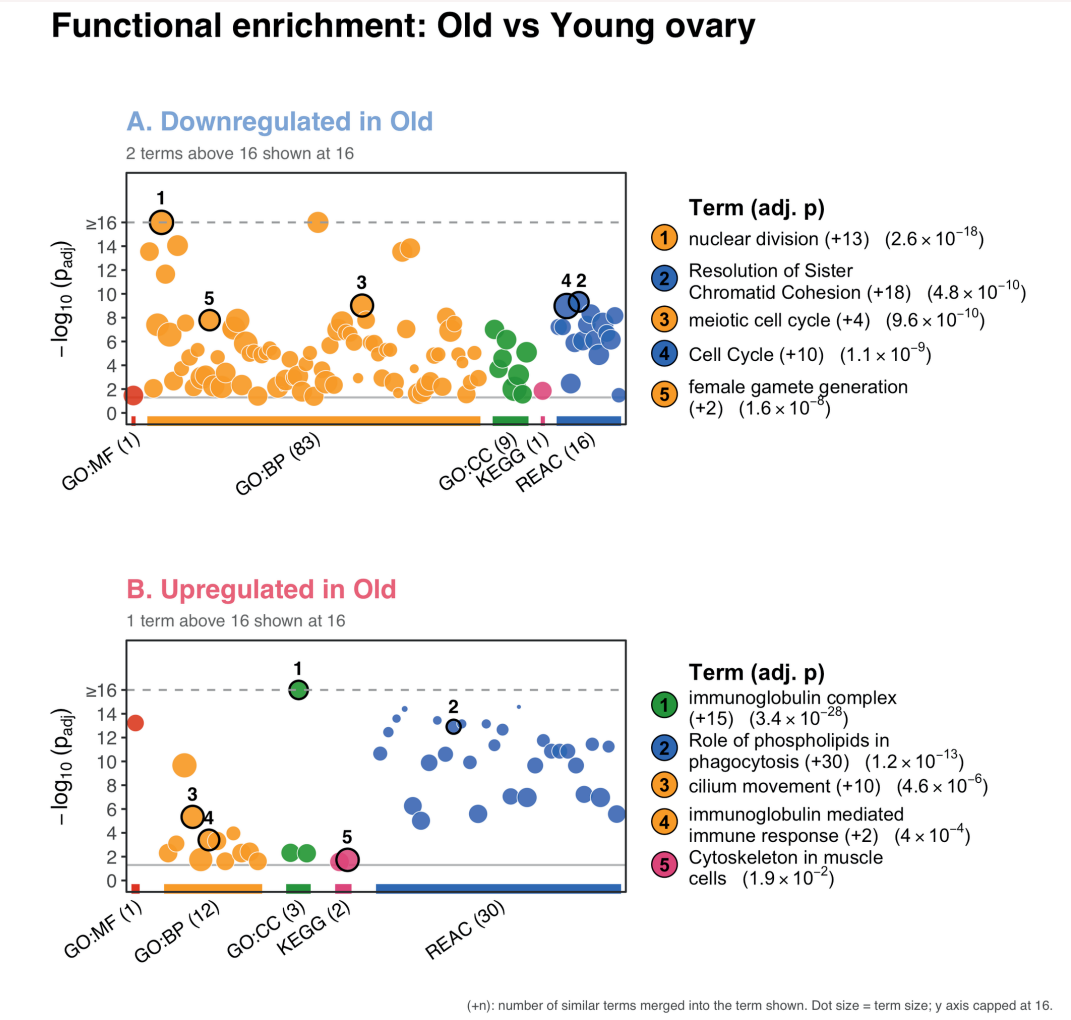


Figure 3. Genes downregulated with age were enriched for cell-cycle and meiosis terms, and upregulated genes for immune-related terms. Over-representation analysis using g: Profiler of the genes downregulated (A) and upregulated (B) in older donors (Old, 50-79 years vs Young, 20-49 years; FDR ≤ 0.01; 593 and 496 genes, respectively), against the 26,448 genes tested in the differential expression analysis as background. Terms from GO, KEGG and Reactome were tested. Each dot is a significant term (adjusted p < 0.05, g:SCS, g: Profiler's default multiple-testing correction; solid grey line), coloured by source database and sized by term size. y-axis: -log₁₀ adjusted p-value, capped at 16 (dashed line; terms above it were plotted at 16). The five most significant non-redundant terms per direction were numbered and listed with their adjusted p-values; terms sharing ≥ 50% of their genes with a better-ranked term were merged for display (+n), number of terms merged). Downregulated genes were most enriched for nuclear division (adjusted p = 2.6 × 10⁻¹⁸), and upregulated genes for immunoglobulin complex (adjusted p = 3.4 × 10⁻²⁸). Most top terms were driven by lowly expressed genes (logCPM < 1), which limits their interpretation. GO, Gene Ontology; BP, biological process; MF, molecular function; CC, cellular component; KEGG, Kyoto Encyclopedia of Genes and Genomes; REAC, Reactome; CPM, counts per million.

DISCUSSION & CONCLUSION

Lower expression of **AMH** and **INSL3** in older ovaries is consistent with age-associated alterations in follicular function. AMH is primarily produced by granulosa cells of growing follicles and is widely used as a marker of ovarian reserve, whereas INSL3 is associated with theca-cell function and ovarian steroidogenesis (Karaviti et al., 2025; Ivell & Anand-Ivell, 2018). Together, their lower expression may reflect changes in follicular abundance, cellular composition, or activity during ovarian ageing.

Among genes with higher expression in older ovaries, **RNF43** and **PLAG1** suggest possible alterations in regulatory signalling. RNF43 regulates WNT/β-catenin signalling, which has been implicated in follicular homeostasis in experimental mouse models (Fu et al., 2024). PLAG1 is associated with the IGF2 regulatory axis in human oocytes; however, reduced PLAG1 expression has been reported in oocytes from women with diminished ovarian reserve (Li et al., 2024), contrasting with its higher expression in our bulk ovarian tissue. This difference may reflect the distinct cellular populations analysed, and the roles of increased RNF43 and PLAG1 expression in human ovarian ageing remain unclear.

At the pathway level, genes with lower expression in older ovaries were enriched for **cell-division, chromosome-segregation, and meiotic processes**, consistent with age-associated alterations in reproductive and proliferative cellular activity. Conversely, genes with higher expression were strongly enriched for **immunoglobulin-related processes**, suggesting possible changes in the ovarian immune environment. Age-associated immune and inflammatory changes have previously been described in ovarian tissue (Isola et al., 2024), including a shift towards an immune-related transcriptional state in post-reproductive mouse ovaries (Converse et al., 2026).

As GTEx provides bulk RNA-seq data, these findings cannot distinguish changes in gene expression within individual cells from age-associated changes in tissue cellular composition. Moreover, the strongest enrichment signals were largely driven by lowly expressed genes, so these results must be interpreted cautiously. Finally, chronological age does not establish menopausal status, and the small number of donors aged 70-79 (n = 4) limits interpretation of patterns in the oldest age group.

Overall, ovarian ageing was associated with reduced reproductive activity and broader changes in regulatory and immune-related processes, although some pathway-level findings were sensitive to expression filtering.

References: Cavalcante MB, Sampaio OGM, Câmara FEA, Schneider A, de Ávila BM, Proszek J, et al. Ovarian aging in humans: potential strategies for extending reproductive lifespan. *GenoScience*. 2023;45(4):2121-33. Liu X, Zhao Y, Feng Y, Wang S, Zhang J. Ovarian Aging: Mechanisms, Age-Related Disorders, and Therapeutic Interventions. *MolCell* (2020). 2025;6(12):e70481. Ito J, Jyu, Ocasio SR, Hubbert CR, Ko S, Mondal SA, Hense JD, et al. A single-cell atlas of the aging mouse ovary. *Nature Aging*. 2023;4(1):145-62. Wang S, Zheng Y, Li J, Yu Y, Zhang W, Song M, et al. Single-Cell Transcriptomic Atlas of Primate Ovarian Aging. *Cell*. 2020;180(3):585-600.e19. Wu M, Tang W, Chen Y, Xue L, Dai J, Li Y, et al. Spatiotemporal transcriptomic changes of human ovarian aging and the regulatory role of FOXR1. *Nature Aging*. 2023;4(4):377-45. Jin C, Wang X, Yang J, Kim S, Hudgens AD, Camille A, et al. Molecular and genetic insights into human ovarian aging from single-nuclei multi-omics analyses. *Nature Aging*. 2025;5(2):275-90. Karaviti E, Karaviti D, Kani E-R, Chatzandreu E, Paschou SA, Psaltopoulou T, et al. The role of anti-Müllerian hormone: insights into ovarian reserve, primary ovarian insufficiency, and menopause prediction. *Endocrine*. 2025;89(2):338-55. Ivell R, Anand-Ivell R. Insulin-like peptide 3 (INSL3) is a major regulator of female reproductive physiology. *Human Reproduction Update*. 2018;24(6):619-51. Fu Y, Zhang M, Sui B, Yuan F, Zhang W, Weng Y, et al. Mesenchymal stem cell-derived apoptotic vesicles ameliorate impaired ovarian folliculogenesis in polycystic ovary syndrome and ovarian aging by targeting WNT signaling. *Theranostics*. 2024;14(8):3385-403. Li X, Wu X, Zhang H, Liu P, Xia L, Zhang N, et al. Analysis of single-cell RNA sequencing in human oocytes with diminished ovarian reserve uncovers mitochondrial dysregulation and translation deficiency. *Reproductive Biology and Endocrinology*. 2024;22(1):146. Converse A, Dipali SS, Schowe IP, Kelly EB, Jambunathan SS, Ocasio SR, et al. The post-reproductive ovary shifts from a reproductive to an immune-like organ. *Molecular Human Reproduction*. 2026;32(2).

AI disclaimer: AI (Claude and CHATGPT) was used to support R coding, namely detect mistakes and generating specific commands, and improve plot visuals.

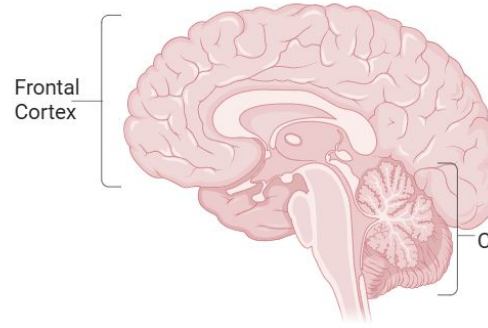
Postmortem Differential Gene Expression Across Brain Areas and Death Types

Angeliki Alexandridou, Noa van 't Hof, Ioanna Koupepia, Simone Molendijk, Kuhu Sheel, Anelia Todorova

Introduction

The Genotype-Tissue Expression (GTEx) project consists of a large database on gene expression and genetic variation in humans (1). RNA-sequencing was performed on 54 types of non-diseased tissues from just under 1000 deceased donors, including tissues from specific brain areas (1). Information about the donors was also recorded and included in the dataset, including the nature of the death according to the Hardy death scale (2). This posed the question if differential gene expression in specific brain areas could be correlated with the type of death. In order to answer this question gene expression patterns in the amygdala, cerebellum and frontal cortex were analysed comparing violent and fast deaths (Hardy score 1) to intermediate deaths (Hardy score 3).

Anatomy of the Brain



Coronal Section of the Brain (Cut at Basal Ganglia)

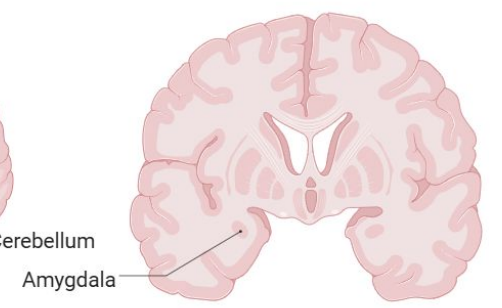


Figure 1: Location of the brain areas examined during this study. Image made using BioRender.

Results: DGE

Differential Expression: Frontal cortex

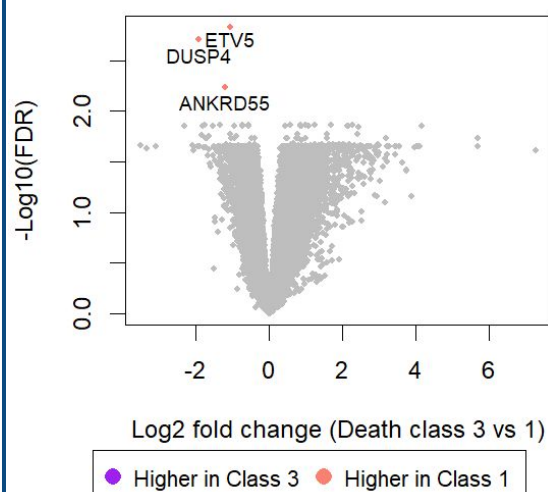


Figure 2: Differential expression between class 1 and class 3 deaths in the frontal cortex. Genes highlighted in purple are significantly more expressed in individuals of class 3 relative to class 1, of which there were none for the frontal cortex. Genes highlighted in orange are significantly more expressed in class 1 relative to class 3. 3 genes are significantly more expressed in class 1, which were identified as ETS5, DUSP4, and ANKRD55 using g:Profiler.

Differential Expression: Cerebellum

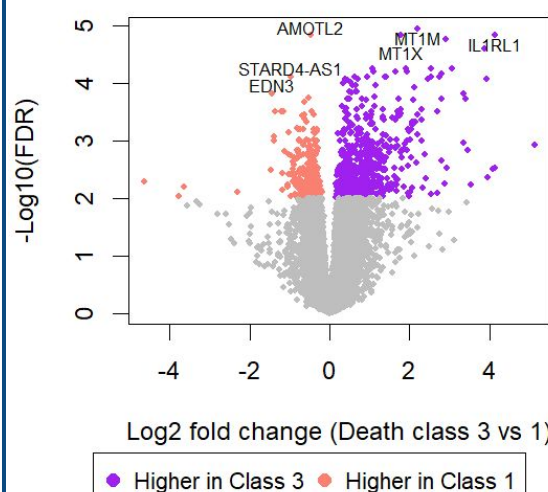


Figure 3: Differential expression between class 1 and class 3 deaths in the Cerebellum. Genes highlighted in purple (609) are significantly more expressed in individuals of class 3 relative to class 1. Genes highlighted in orange (274) are significantly more expressed in class 1 relative to class 3. Top 3 genes in each class are labelled in the plot with the gene names extracted from g:Profiler.

Differential Gene Expression: Amygdala

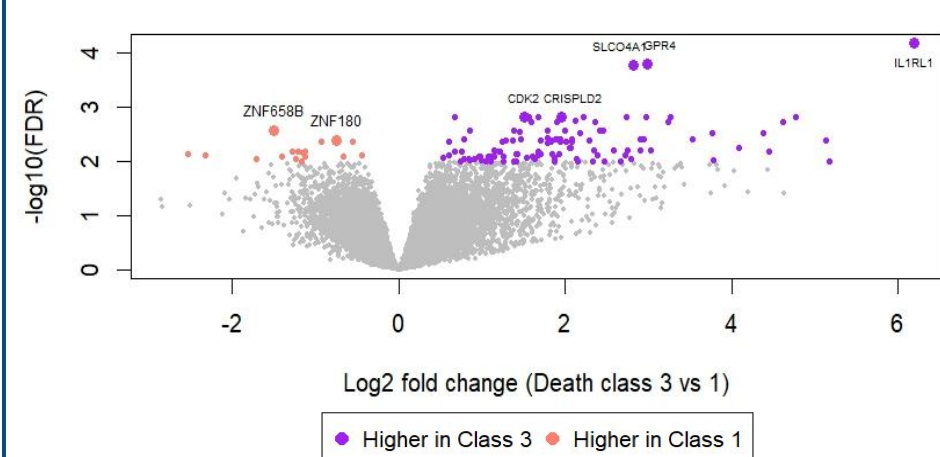


Figure 4: Differential expression between class 1 and class 3 deaths in the Amygdala. Genes highlighted in purple (110) are significantly more expressed in individuals of class 3 relative to class 1. Genes highlighted in orange (17) are significantly more expressed in class 1 relative to class 3. The top 5 and top 2 genes in each class are labelled in the plot with the gene names extracted from g:Profiler.

Results: PCA

PCA of gene expression across brain regions

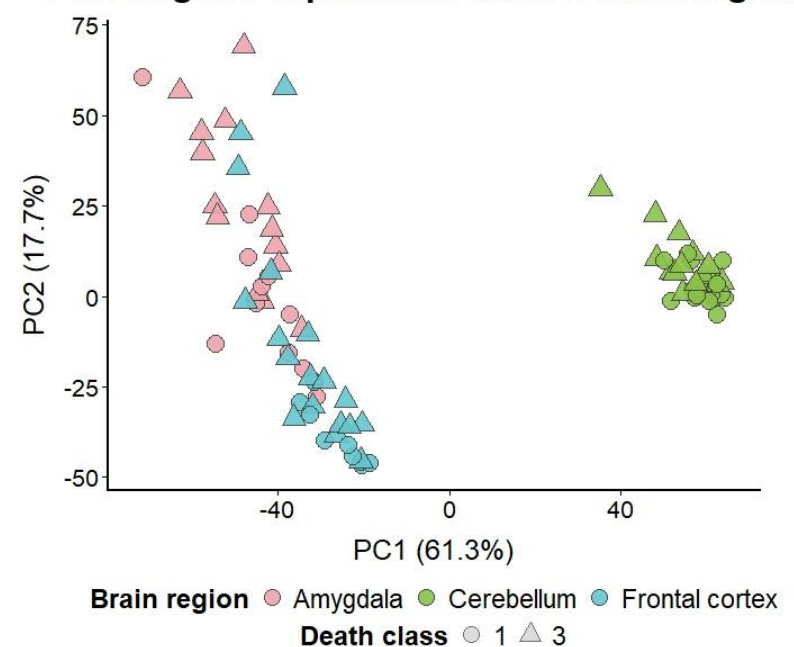


Figure 5: PCA plot representing the variance between the three different brain regions. Cerebellum shows a very distinct gene expression pattern in contrast to the Amygdala and Frontal cortex.

Discussion

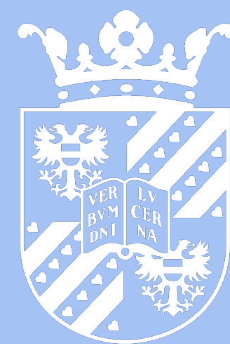
Using the DGE analysis we identified multiple genes that were differentially expressed between death class 1 and 3. In the cerebellum, EDN3 was more highly expressed in death class 1. This gene has been associated with neural cell differentiation and traumatic brain injury, which could precede sudden, violent death (3). Similarly, ETV5, which is more highly expressed in the frontal cortex of class 1, regulates neural cell differentiation (4). In contrast, GPR4, involved in vascular responses, was more highly expressed in the amygdala of class 3, potentially suggesting differences in vascular regulation between death classes (5,6). Higher variation between the two death classes was observed in the cerebellum as compared to the other two regions.

Limitations include small sample sizes and lack of information about the previous conditions of patients of the class 3 cohort. Future research could investigate whether the observed differences in gene expression reflect distinct biological processes associated with the different death classes. These findings could potentially contribute to forensic approaches for distinguishing death classes. Additionally, if there is a large difference in gene expression patterns, death classification could be considered as a potential confounding factor in postmortem transcriptomic studies. Studying the effect of death classes could help researchers distinguish disease-associated changes from those related to the dying process.

References:

- 1) Aguet, F., Anand, S., Ardlie, K. G., Gabriel, S., Getz, G. A., Graubert, A., Hadley, K., Handsaker, R. E., Huang, K. H., Kashin, S., Li, X., MacArthur, D. G., Meier, S. R., Nedzel, J. L., Nguyen, D. T., Segrè, A. V., Todres, E., Balliu, B., Barbeira, A. N., . . . Volpi, S. (2020). The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science*, 369(6509), 1318–1330. <https://doi.org/10.1126/science.aaz1776>
- 2) GTEx Portal. (n.d.). <https://www.gtexportal.org/home/>
- 3) Schober, K., Ondruschka, B., Dreßler, J., & Abend, M. (2014). Detection of hypoxia markers in the cerebellum after a traumatic frontal cortex injury: a human postmortem gene expression analysis. *International Journal Of Legal Medicine*, 129(4), 701–707. <https://doi.org/10.1007/s00414-014-1129-3>
- 4) Liu, Y., & Zhang, Y. (2019). ETV5 is Essential for Neuronal Differentiation of Human Neural Progenitor Cells by Repressing NEUROG2 Expression. *Stem Cell Reviews And Reports*, 15(5), 703–716. <https://doi.org/10.1007/s12015-019-09904-4>
- 5) Silva, P. H. I., & Wagner, C. A. (2022). Physiological relevance of proton-activated GPCRs. *Pflügers Archiv - European Journal Of Physiology*, 474(5), 487–504. <https://doi.org/10.1007/s00424-022-02671-1>
- 6) Cheng, Z., Nie, W., Leng, J., Yang, L., Wang, Y., Li, X., & Guo, L. (2024). Amygdala and cognitive impairment in cerebral small vessel disease: structural, functional, and metabolic changes. *Frontiers in Neurology*, 15, 1398009. <https://doi.org/10.3389/fneur.2024.1398009>

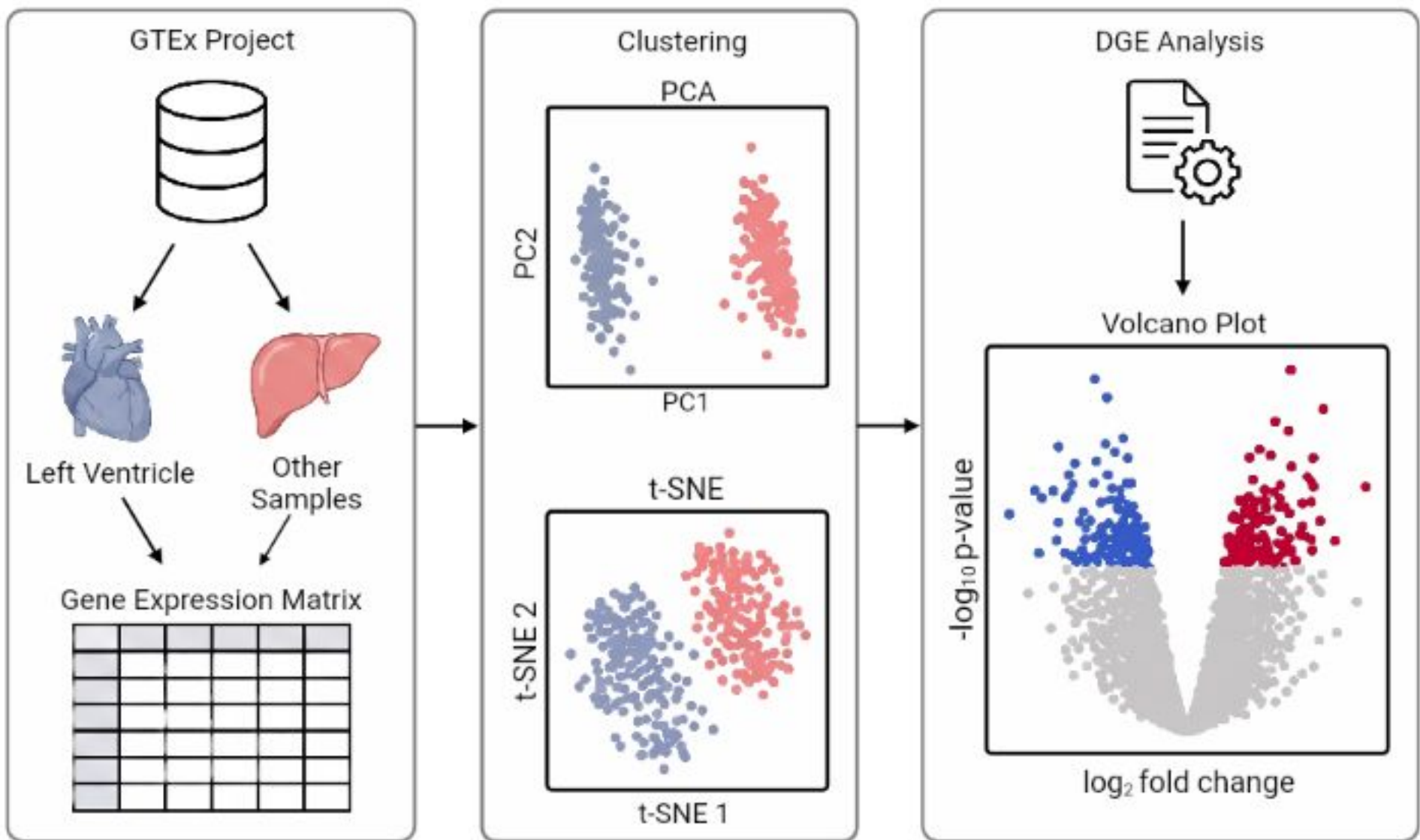
Analysing grouping in sudden cardiac death victims with post-mortem transcriptomics



Koert Jansonius (S4902998), Viktor Koetsier (S5575095), Tristan Meinen (S5576261), Jack Nicholls (S5508193), Ruben van de Rozenberg (S5597919), Erika Sonko (S5605423)

Background

- Cardiovascular disease is the world's leading cause of death, sudden cardiac death (SCD) being a type that is hard to predict (Markwerth et al., 2021).
- The Genotype Tissue Expression project (GTEx) provides post-mortem transcriptomics data of many human tissues, ordered by type of death (Hardy scale) (Lonsdale et al., 2013).
- We studied Hardy scale death type 2 (sudden natural death of <1 hour), of which ~90% are related to SCD (Sessa et al., 2021).
- RQ: To what extent is there subgrouping of gene expression patterns across different organ tissues following SCD, and are these subgroups involved in pathogenesis?



Created with BioRender.com, <https://app.biorender.com/illustrations/6ac60cd6744106d431a1a9ed>

Left ventricle grouping is not present in transcriptome of blood and other tissues

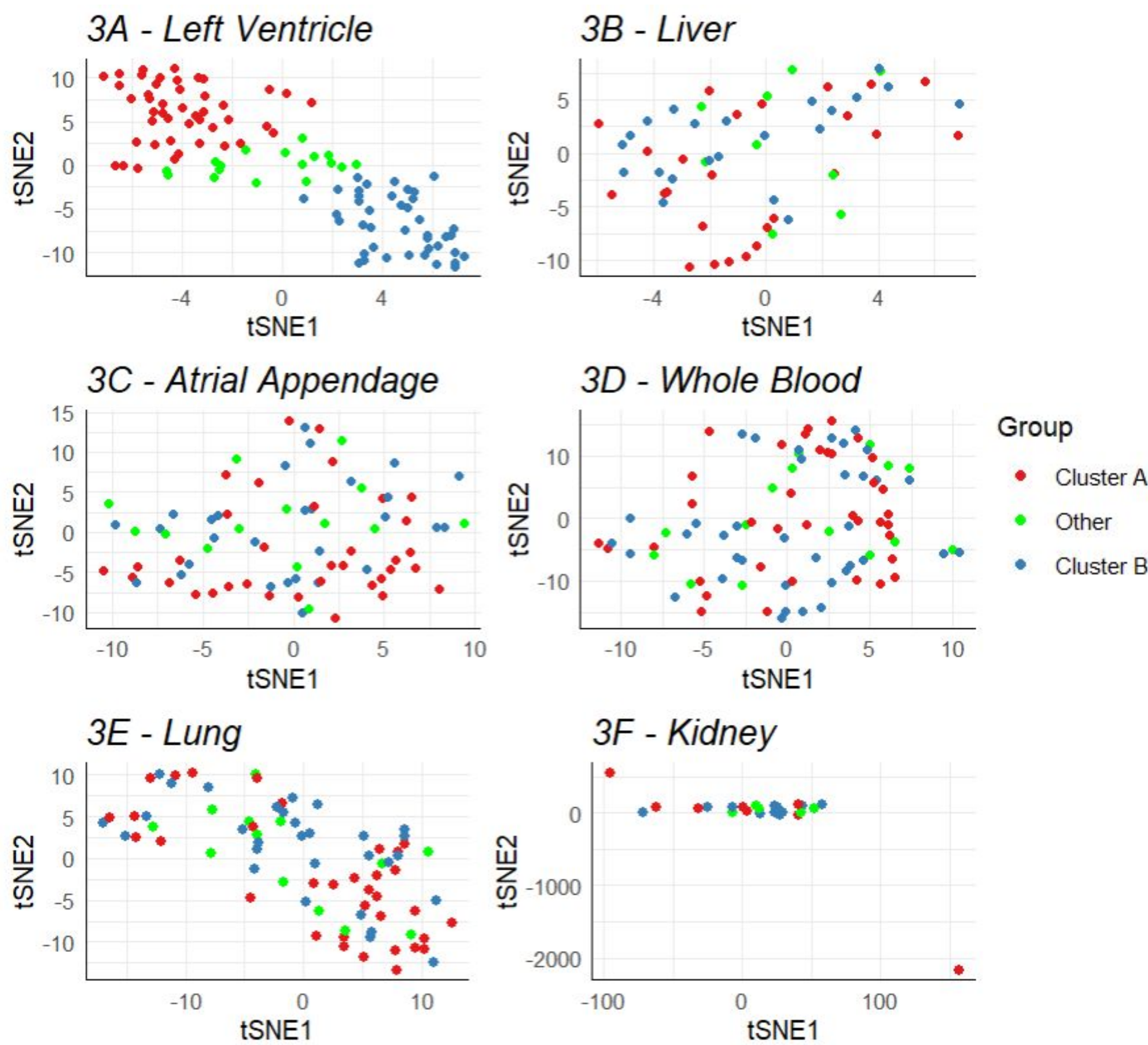


Figure 3A-F. Heart expression groups do not predict expression patterns in other tissues. t-SNE of GTEx RNA-seq (log2 CPM, 5,000 most variable genes, seed = 42) for (A) left ventricle, the main pumping chamber of the heart and the reference tissue for all comparisons, (B) liver, the main metabolic organ, linked to the heart through the liver-heart axis, in which fatty liver disease accelerates coronary artery disease (C) atrial appendage, Heart muscle from the upper chambers that receive bloods, and the closest comparison to the ventricle (D) whole blood, connecting all organs allows for testing also on living people (E) lung and (F) kidney. Each point is one donor. Donors were assigned to three groups based on their position in the death class 2 left ventricle t-SNE (A): top-left (red, n = 35), bottom-right (blue, n = 29) and middle (green, remaining donors). These heart-defined groups are shown on the t-SNE of each other tissue (B-F). No tissue showed a significant association between heart group and expression in that tissue (principal components 1-10, t-SNE axes, and a distance permutation test; all $P > 0.05$ after multiple-testing correction).

Left ventricle transcriptomics shows two clusters uncorrelated to sex and age

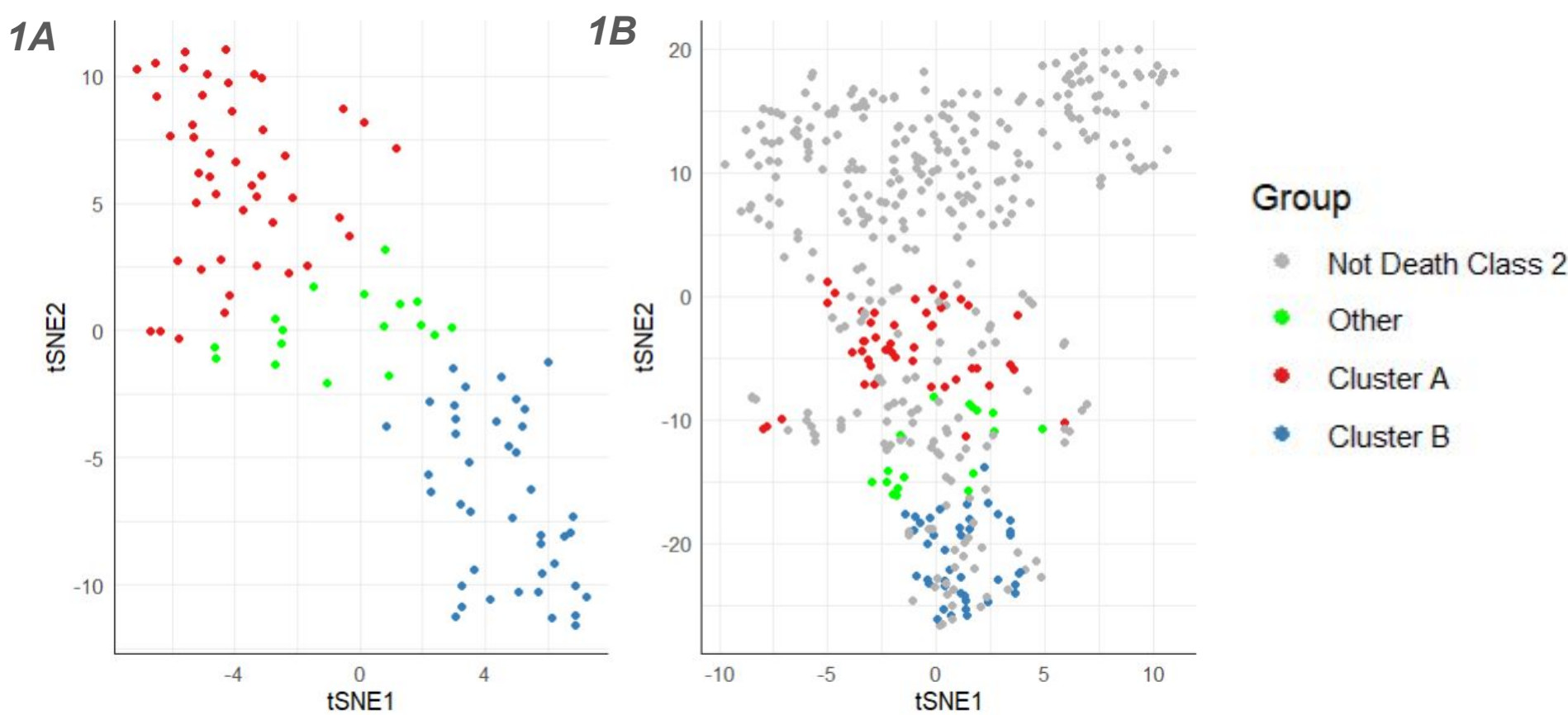


Figure 1A-B. Left ventricle gene expression separates death class 2 (fast death of natural causes) samples from other donors and reveals two groups within class 2. t-SNE of GTEx left ventricle RNA-seq (log2 CPM, 5,000 most variable genes). (A) t-SNE of death class 2 samples only (n = X; perplexity = 15). Samples are colored by position along the main axis of variation (top-left: red, middle: green, bottom-right: blue). (B) t-SNE of all left ventricle samples (n = Y; perplexity = 30). Death class 2 samples are colored using the groups from (A); all other samples are grey. Each point is one sample.

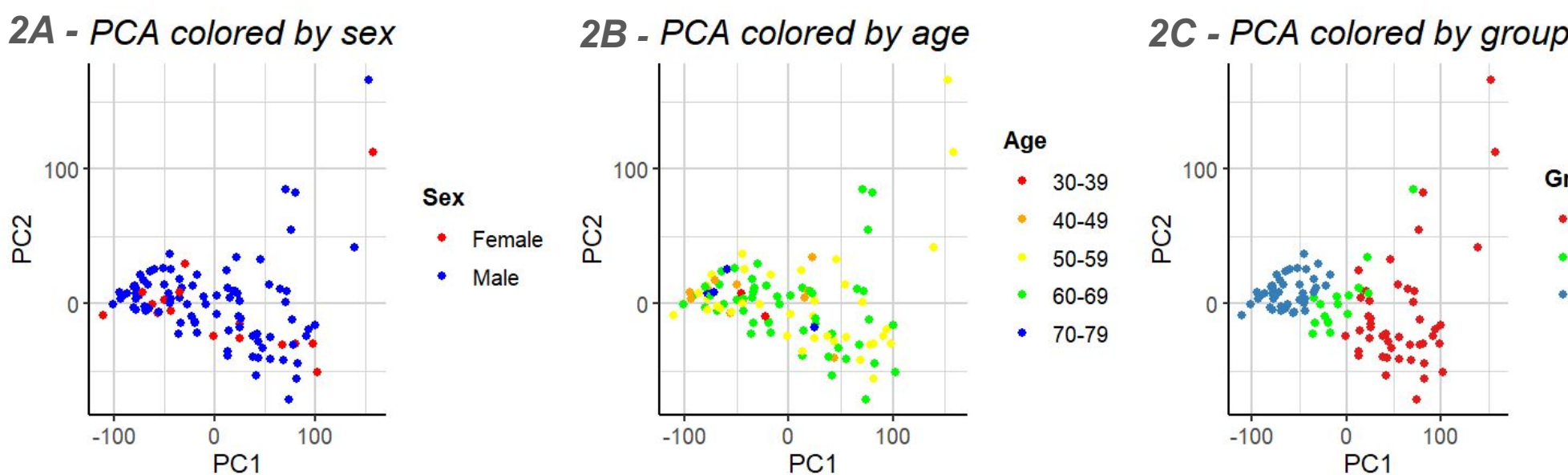


Figure 2A-C: To control that the two groups in figure 1A are not caused by the variation in sex and age, we performed principal component analysis (PCA). PC1 & PC2 represent 55.02 and 14.56% of variation respectively. Figure 2A shows that male and female samples are distributed randomly throughout the groups. Figure 2B shows that age groups are also distributed randomly throughout the sample group. Figure 2C shows that the two groups that were observed in t-SNE clustering are also distinct in PCA analysis.

DGE shows mitochondrial genes are differentially expressed in both heart patient clusters

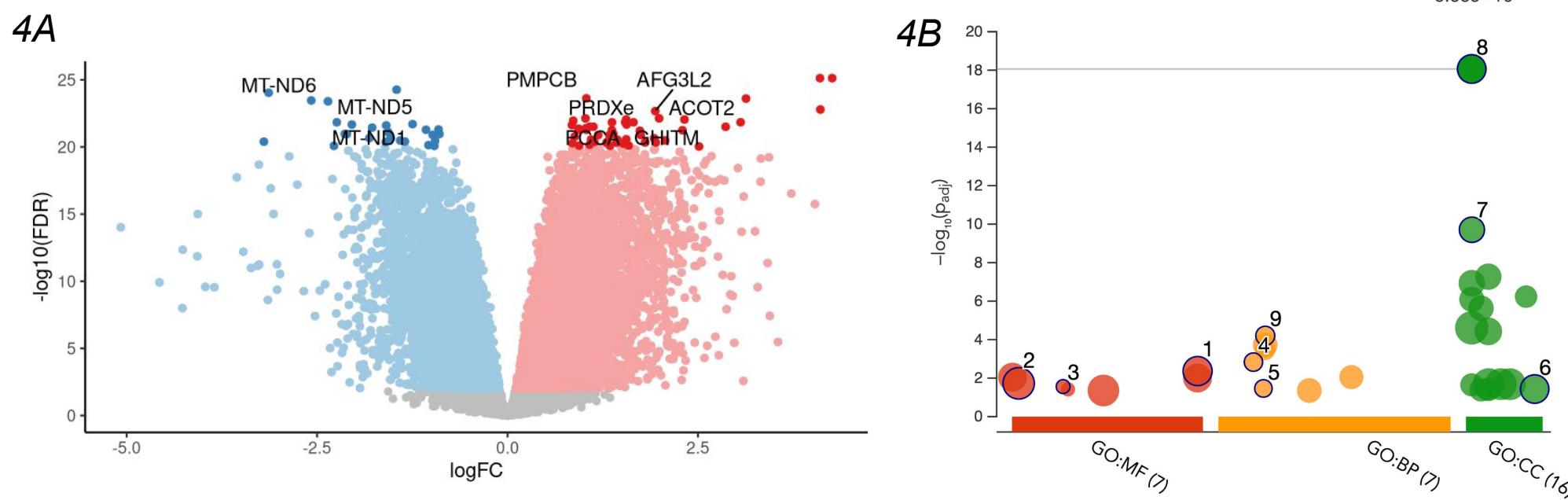


Figure 4A-B: Comparison of the transcriptome of cluster A to B. Because those clusters were already based on transcriptomics data, there are >10,000 significant genes. Significant genes (FDR < 0.01) are coloured (red = upregulated in cluster A). The most significant genes (FDR < 10e-20) are coloured brighter. They were used as input for a gene ontology analysis (figure 4B). The most significant genes related to mitochondrial function are labelled.

Perspectives & conclusion

- Left ventricle transcriptome of Sudden Death (SD) victims shows little resemblance with transcriptome of other organs, indicating that predictors for SD will not be found in other tissues or blood
- In SD patients, 2 clusters are observable in left ventricle transcriptome data
- These clusters might functionally relate to mitochondrial health
- Further research into causes of death will improve understanding

References:

Di Cesare, M., Perel, P., Taylor, S., Kabudula, C., Bixby, H., Gaziano, T. A., McGhie, D. V., Mwambi, J., Pervan, B., Narula, J., Pineiro, D., & Pinto, F. J. (2024). The Heart of the World. *Global Heart*, 19(1), 11. <https://doi.org/10.5334/gh.1288>

Fang, Z., Jia, S., Mou, X., Li, Z., Hu, T., Tu, Y., Zhao, J., Zhang, T., Lin, W., Lu, Y., Feng, C., & Xia, S. (2024). Shared genetic architecture and causal relationship between liver and heart disease. *iScience*, 27(4), 109431. <https://doi.org/10.1016/j.isci.2024.109431>

Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters, G., Garcia, F., Young, N., Foster, B., Moser, M., Karasik, E., Gillard, B., Ramsey, K., Sullivan, S., Bridge, J., Magazine, H., Syron, J., ... Moore, H. F. (2013). The Genotype-Tissue Expression (GTEx) project. *Nature Genetics*, 45(6), 580–585. <https://doi.org/10.1038/ng.2653>

Markwerth, P., Bajonowski, T., Tzimas, I., & Dettmeyer, R. (2021). Sudden cardiac death—Update. *International Journal of Legal Medicine*, 135(2), 483–495. <https://doi.org/10.1007/s00414-020-02481-z>

Sessa, F., Esposito, M., Messina, G., Di Mizio, G., Di Nunno, N., & Salerno, M. (2021). Sudden Death in Adults: A Practical Flow Chart for Pathologist Guidance. *Healthcare*, 9(7), 870. <https://doi.org/10.3390/healthcare9070870>

Introduction

- The gut-brain axis is a bidirectional communication network between the gut and the brain (Mayer et al., 2022)
- It plays a role in diseases and conditions (irritable bowel syndrome, neurodegenerative diseases, etc.) (Mayer et al., 2022)
- The hypothalamus and cerebellum are involved in hunger regulation, but only the hypothalamus is part of the gut-brain axis (Iosif et al., 2023; Wachsmuth et al., 2022)

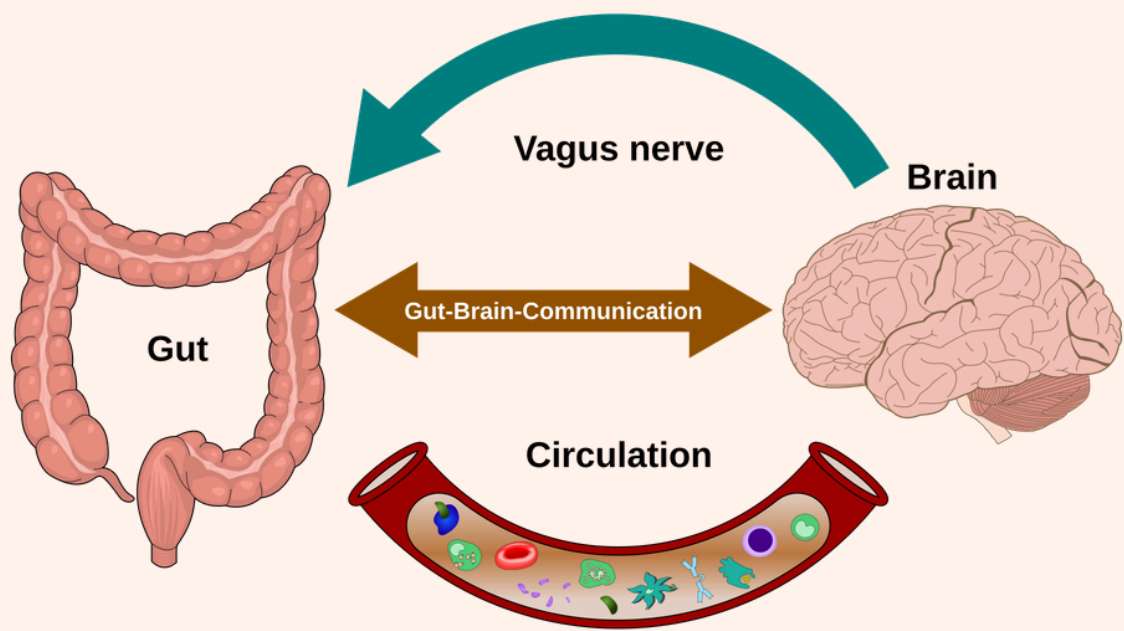


Figure 1. The bidirectional gut-brain axis signalling. Gut-derived signals reach the brain through neural and hormonal pathways, including the leptin-melanocortin pathway. Figure adapted from Suganya et al., 2020.

Our focus: The leptin-melanocortin pathway

The six genes we study: GHSR, LEPR, GLP1R, POMC, AGRP, MC4R (Yeo et al., 2021)

Research question: Does the hypothalamus share more gene expression with the gut than the cerebellum shares with the gut?

Principal Component Analysis

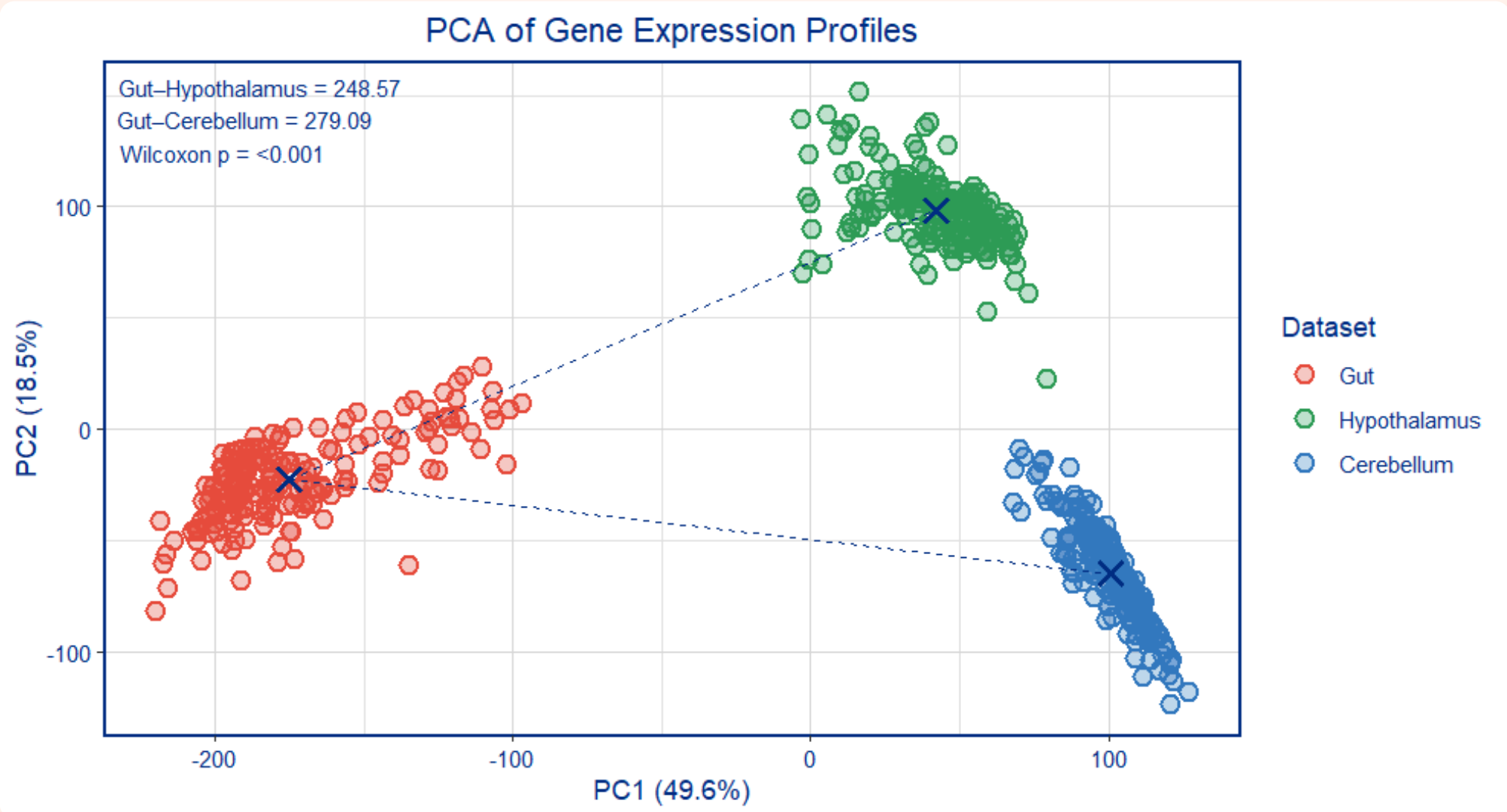


Figure 2. Principal component analysis (PCA) revealed distinct gene-expression profiles among tissues. Each point represents a sample from the GTEx dataset: Gut (red, n = 187), Hypothalamus (green, n = 202), or Cerebellum (blue, n = 241). Genes with counts per million (CPM) > 1 in at least half of the samples were retained for the PCA. The distance between cluster centroids (marked with an X) was slightly smaller for Gut-Hypothalamus (248.57) than for Gut-Cerebellum (279.09), indicating greater similarity in overall gene-expression profiles between the gut and hypothalamus. Distances to the gut centroid differed significantly between the hypothalamus and cerebellum (Wilcoxon rank-sum test, $p < 0.001$). PC1 and PC2 explained 49.6% and 18.5% of the total variance, respectively. The figure was created using R and the ggplot2 package.

Correlation Analysis

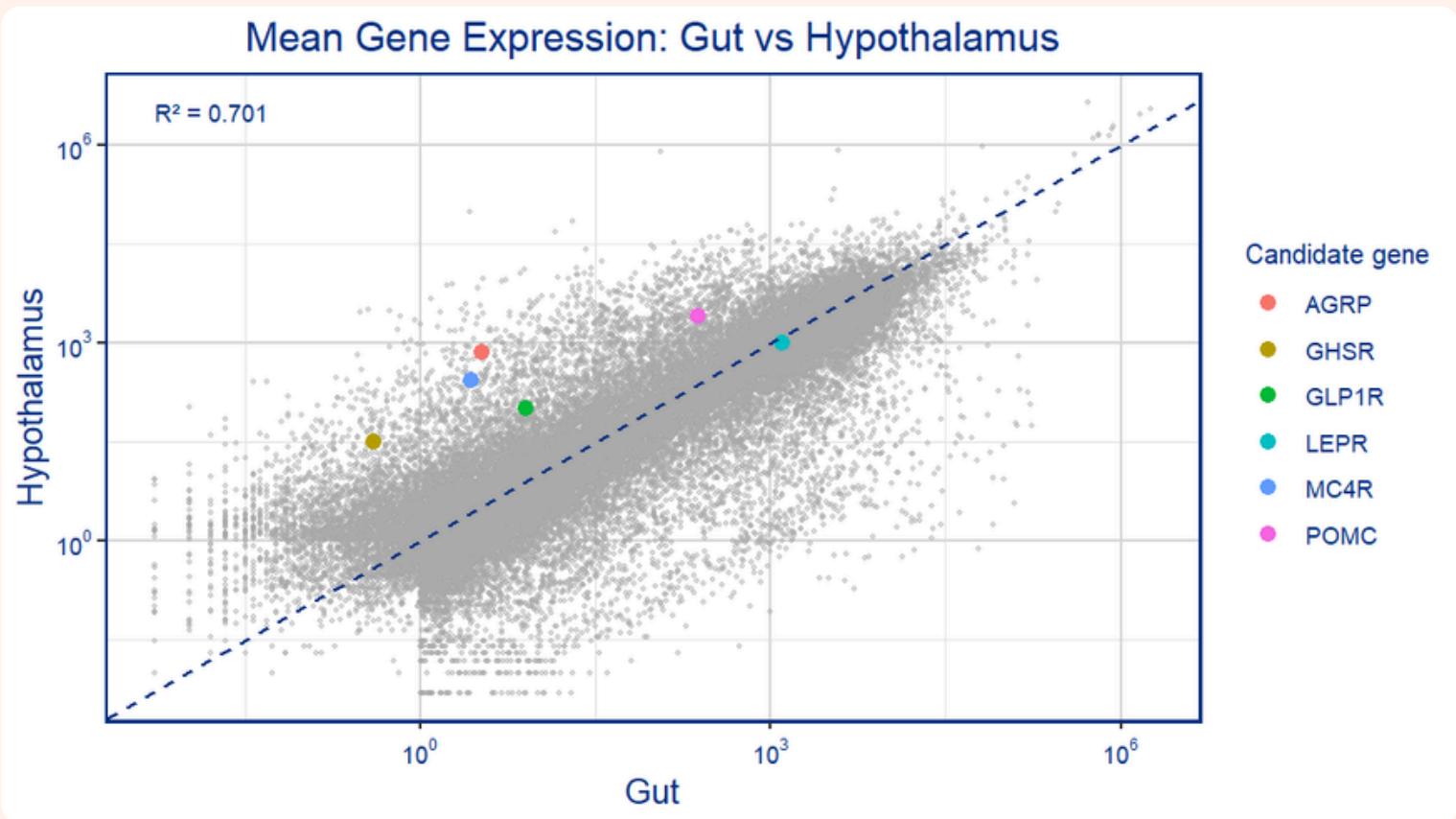


Figure 3. Scatter plot of mean gene expression between the gut and the hypothalamus shows a strong correlation, with LEPR showing a particular strong correlation. Comparison of log10-transformed mean gene expression levels in gut tissue (x-axis) versus hypothalamic tissue (y-axis) derived from the GTEx dataset (n = 35917). Each gray point represents a gene, and highlighted coloured points indicate key candidate genes involved in gastrointestinal tract signal perception: AGRP (orange), GHSR (yellow), GLP1R (green), LEPR (light blue), MC4R (dark blue), and POMC (pink). The dashed line represents the linear regression line ($y = x$) where the expression is equal in the gut and in the hypothalamus. The R^2 value (0.701) represents how strongly the gene expression is correlated between the gut and the hypothalamus for all genes.

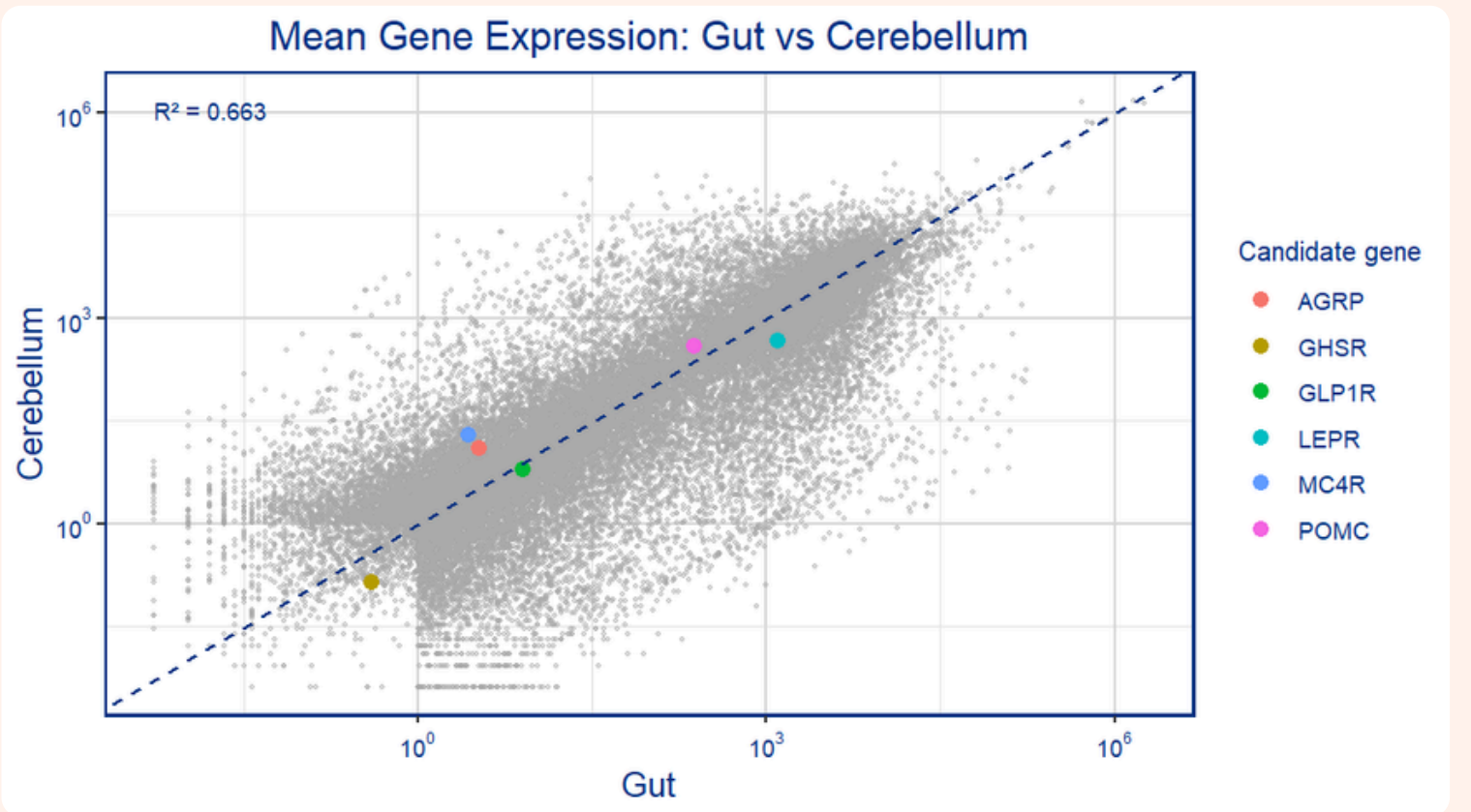


Figure 4. Scatter plot of mean gene expression between the gut and the cerebellum shows correlation for candidate genes. Comparison of log10-transformed mean gene expression levels in gut tissue (x-axis) versus cerebellar tissue (y-axis) derived from the GTEx dataset (n = 35917). Each gray point represents a gene, and highlighted coloured points indicate key candidate genes involved in gastrointestinal tract signal perception: AGRP (orange), GHSR (yellow), GLP1R (green), LEPR (light blue), MC4R (dark blue), and POMC (pink). The dashed line represents the linear regression line ($y = x$) where the expression is equal in the gut and in the cerebellum. The R^2 value (0.663) represents how strongly the gene expression is correlated between the gut and the cerebellum for all genes. The selected candidate genes show a higher correlation than the mean gene expression correlation observed between the gut and hypothalamus (Figure 3).

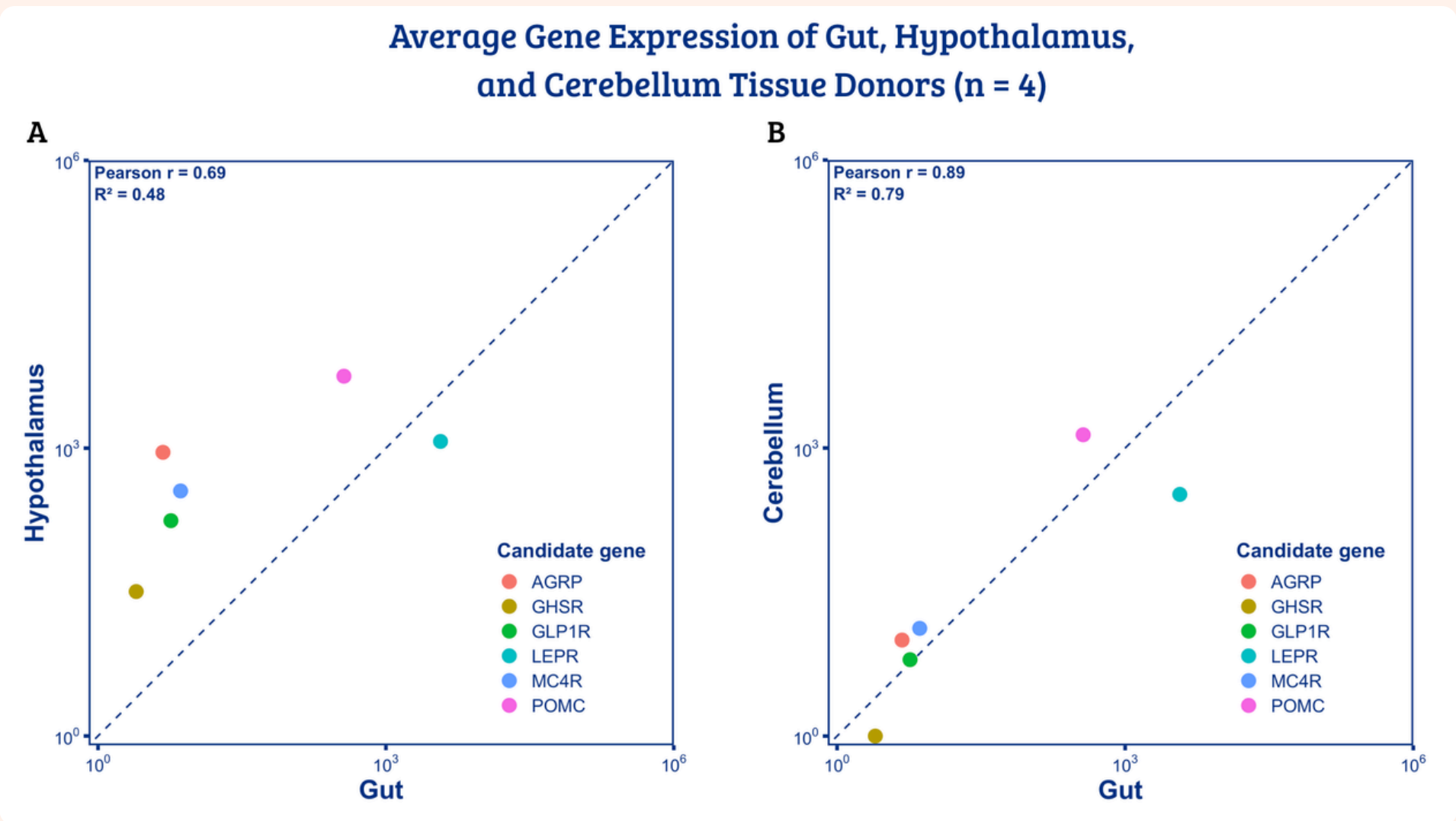


Figure 5. Correlation of mean candidate gene expression is higher between the gut and cerebellum compared to the gut and hypothalamus. Comparison of log10-transformed mean gene expression levels across 4 donors in (A) the hypothalamus versus gut ($r = 0.69$, $R^2 = 0.48$) and (B) the cerebellum versus gut ($r = 0.89$, $R^2 = 0.79$) for six key candidate genes derived from the GTEx dataset. Each point represents a candidate gene involved in gastrointestinal signal perception: AGRP (orange), GHSR (yellow), GLP1R (green), LEPR (light blue), MC4R (dark blue), and POMC (pink). The dashed diagonal line represents equal expression ($y = x$), where the expression is equal between the gut and either hypothalamus or cerebellum.

Functional Analysis

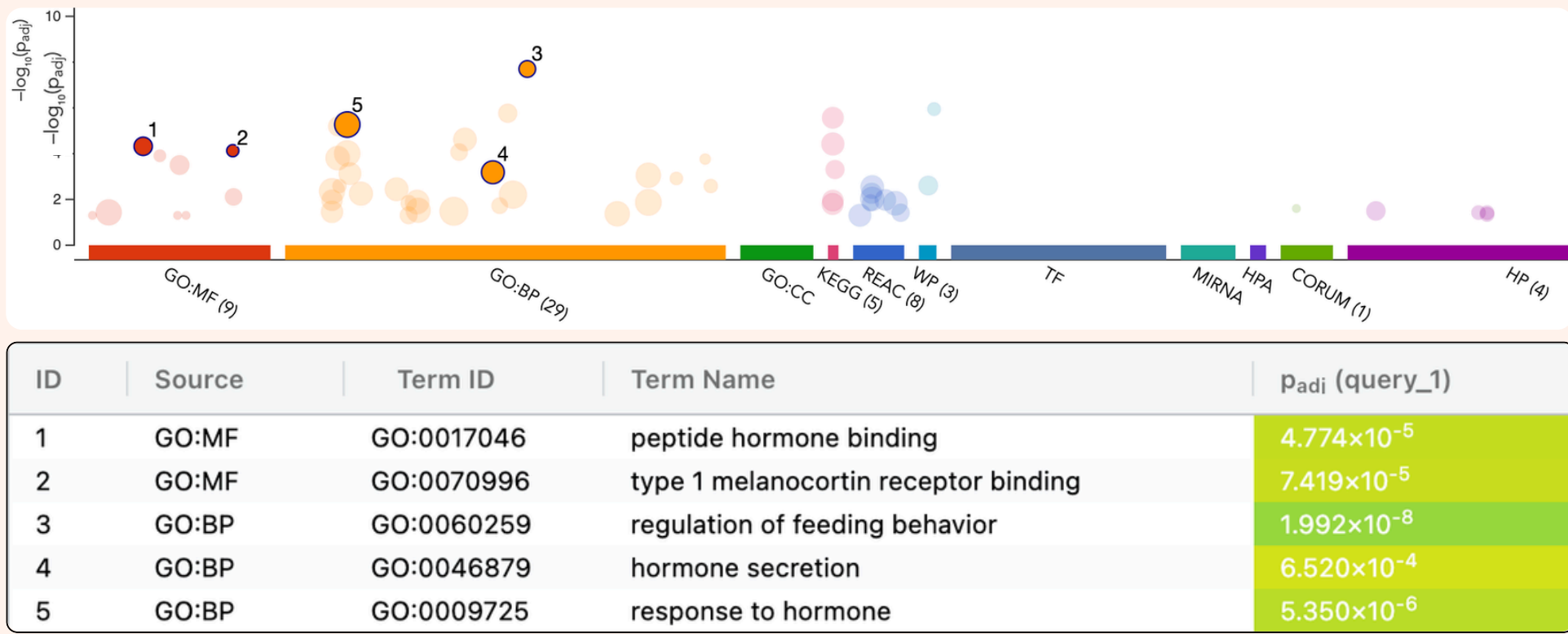


Figure 6. Functional enrichment analysis of candidate genes. Overrepresentation analysis (g:Profiler) of the six candidate genes across biological categories (x-axis) depicted by statistical significance (y-axis, $-\log_{10}(p\text{-adj})$). The plot depicts the most overrepresented biological themes. Unselected terms are transparently depicted.

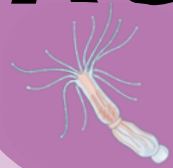
Conclusion & Discussion

- **Correlation:** PCA shows that the hypothalamic transcriptome aligns more closely with the gut than the cerebellum. Similarly, the association analyses of average gene expression display a more similar transcriptome between the hypothalamus and gut ($R^2 = 0.70$) than cerebellum and gut ($R^2 = 0.66$). However, the candidate genes do not show a closer gut-hypothalamus transcriptome compared to that of the gut-cerebellum, at either the cohort-wide or four-donor levels.
- **Function:** The functional validation points to peptide hormone binding and regulation of feeding behaviour, highlighting a specific functional module within a broader systemic metabolic homeostasis network. This confirms that the candidate genes are exclusively relevant to the gut-brain axis.
- **Future perspectives:** Since the answer to the research question is inconclusive, more research is needed in different genes and pathways with more individual patient information in all 3 tissues.

References



Regeneration Genes in Sea Anemones

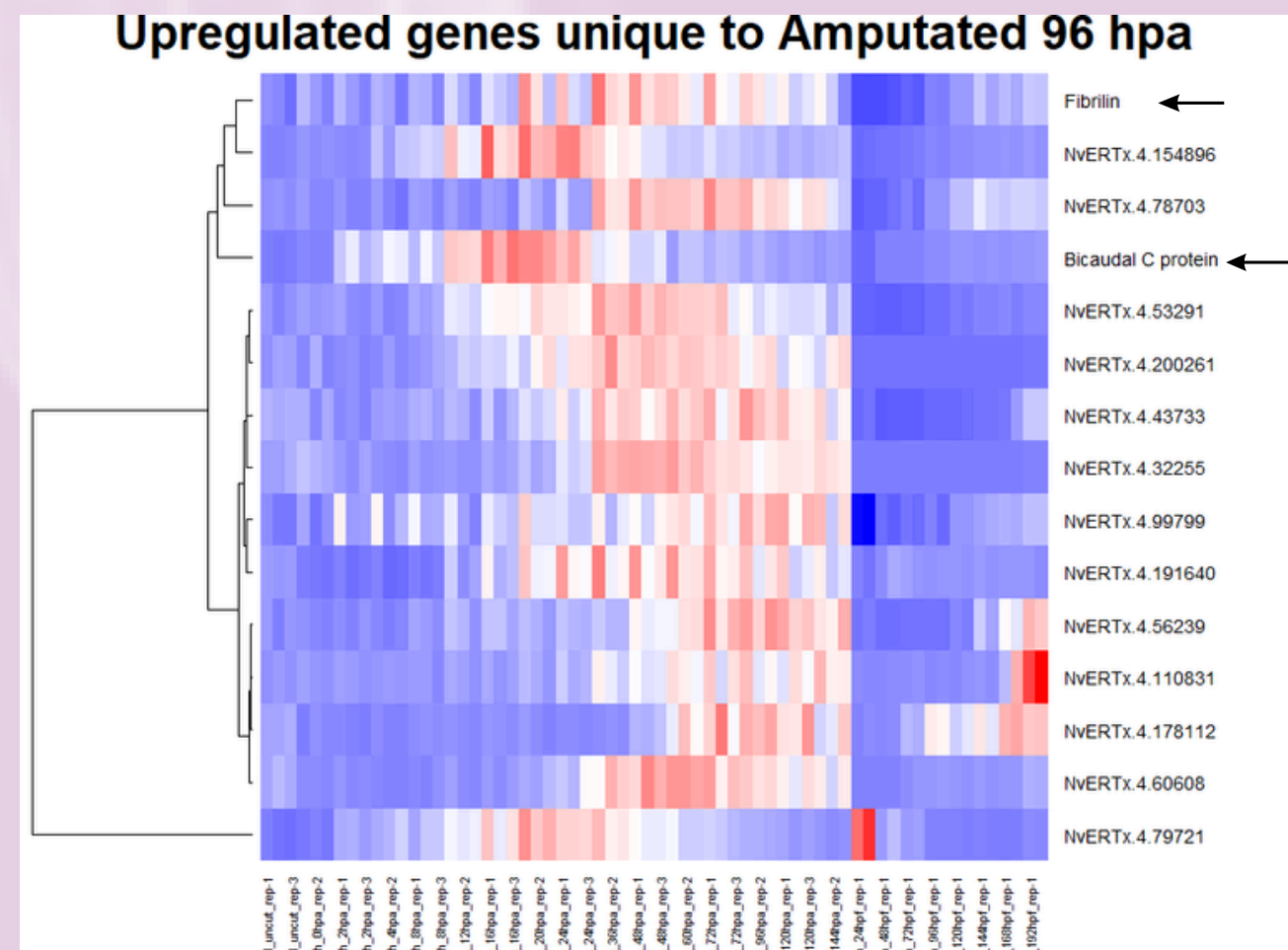
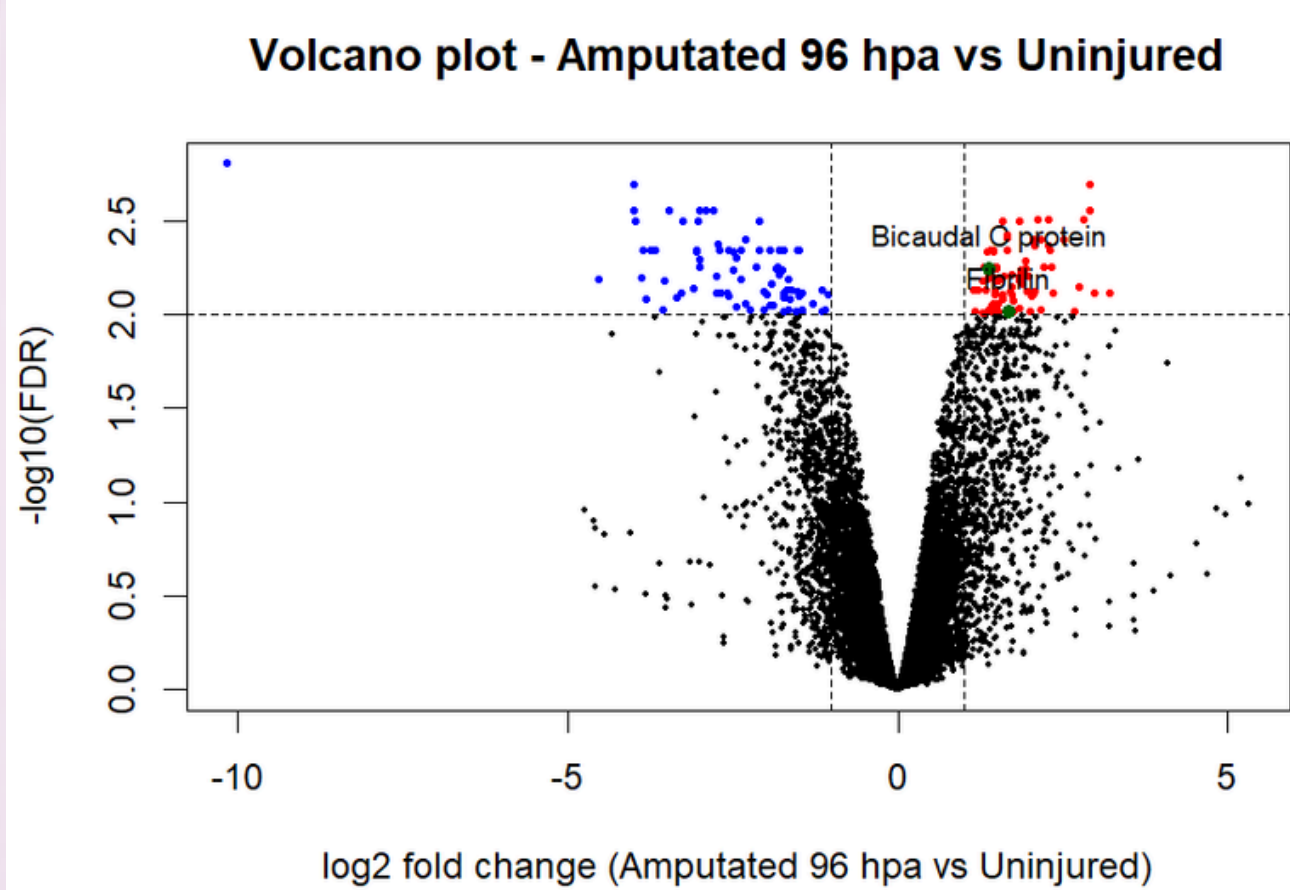
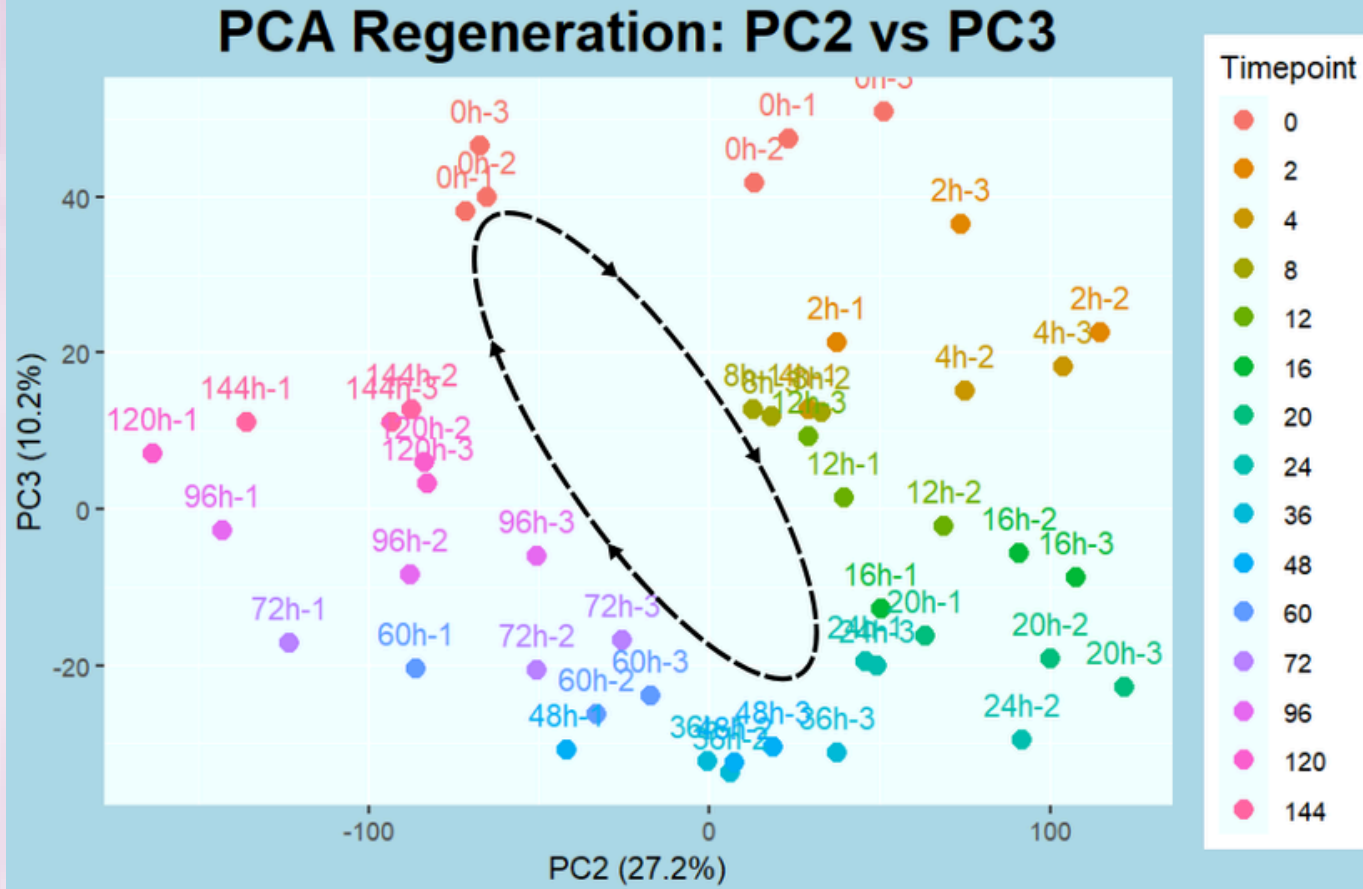
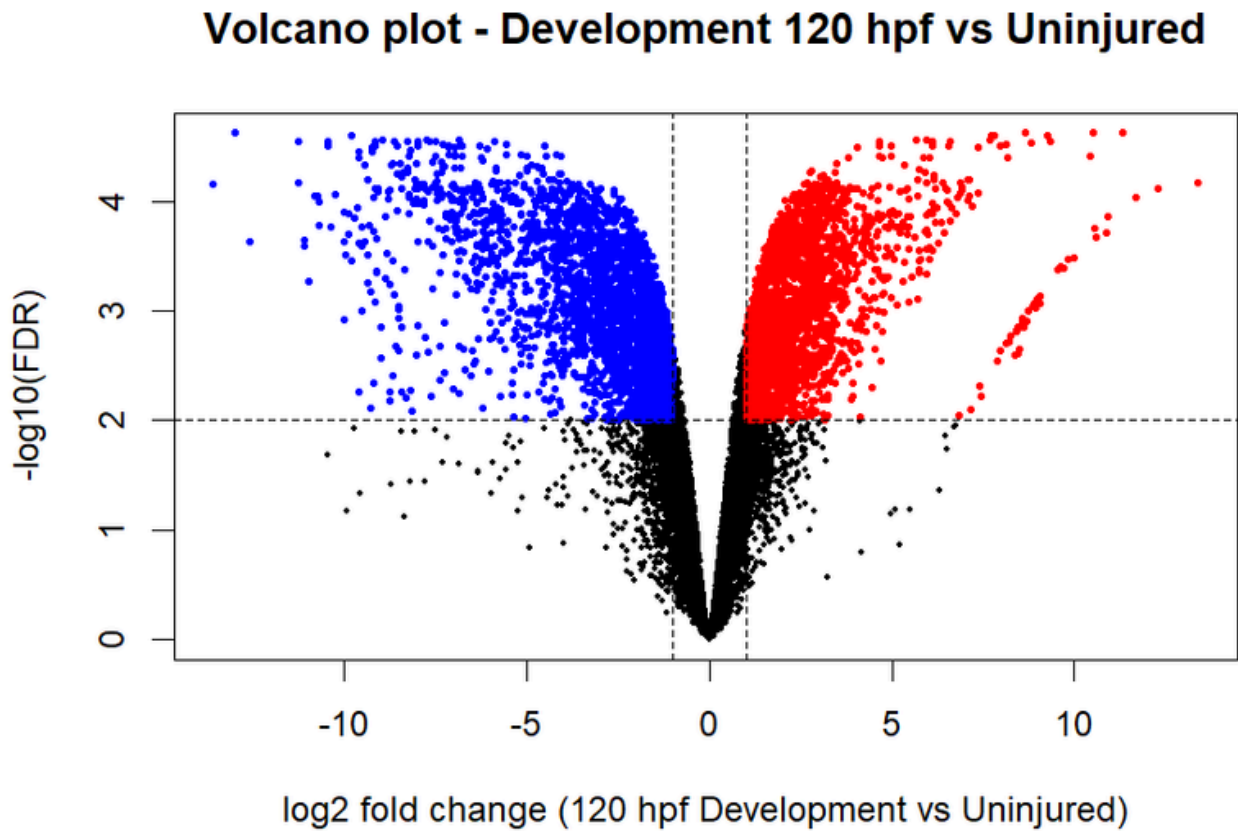
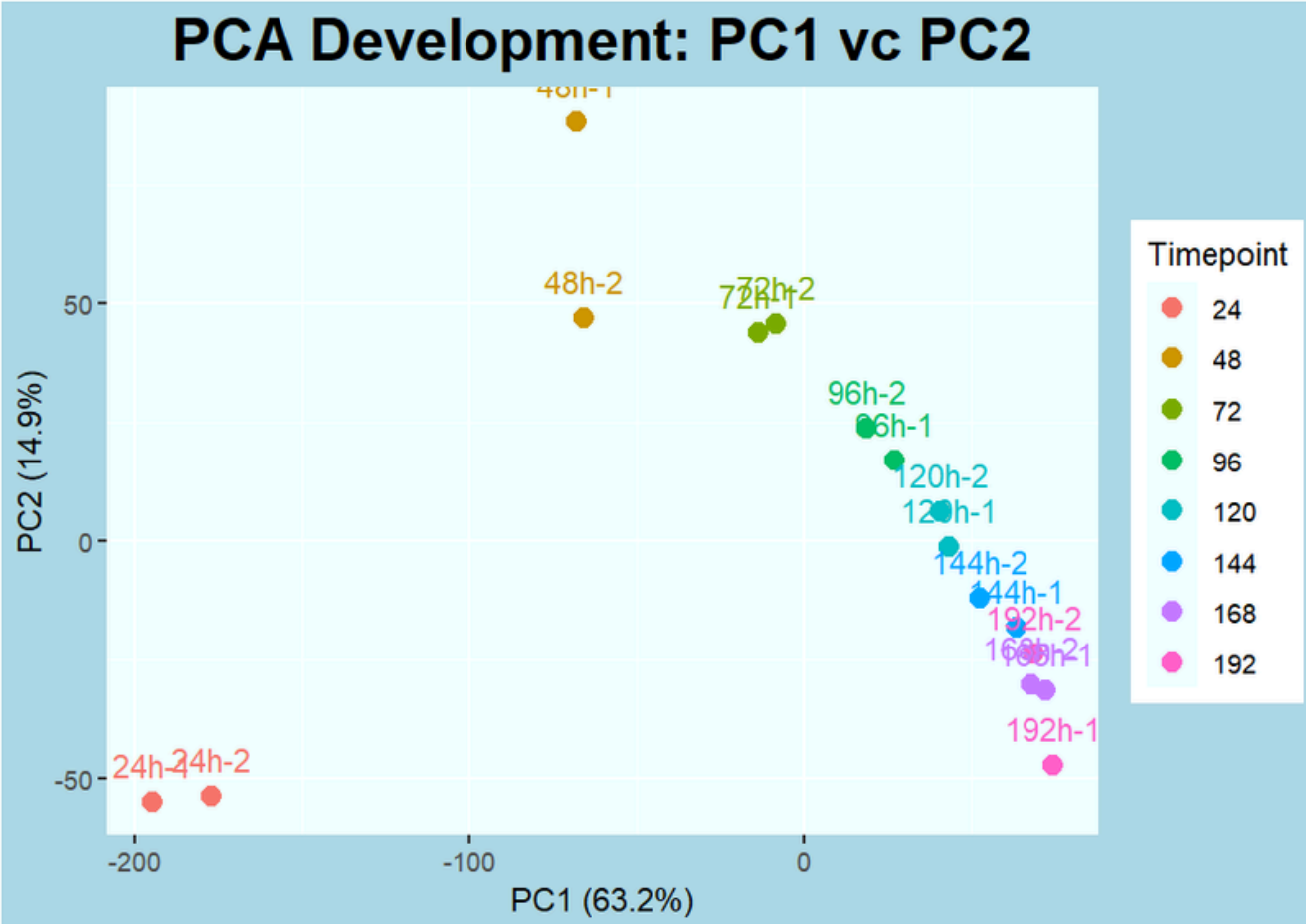
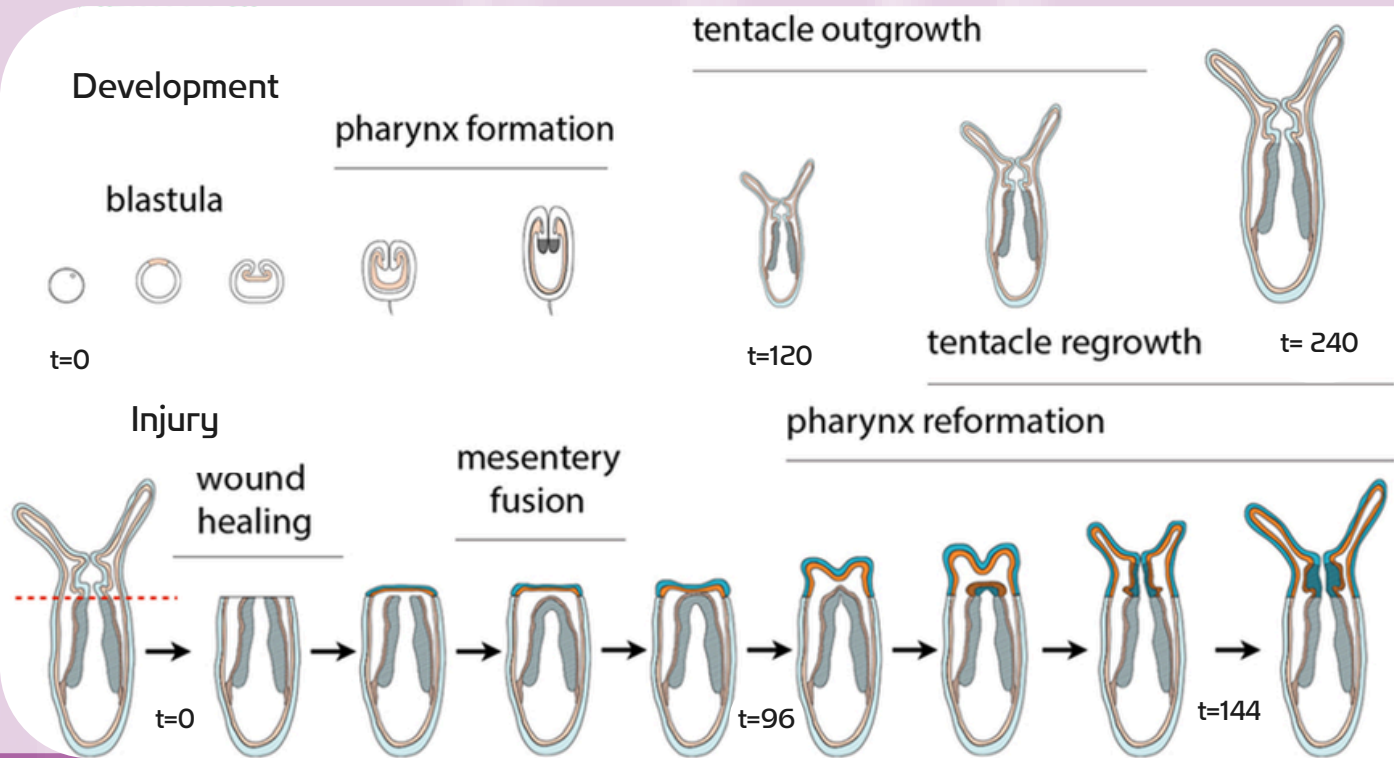


Femke Yska, Rosalie Molenaar, Jorg de Wit, Laurens de Beer

Nematostella vectensis → sea anemone

- Rapid and efficient **regeneration** after amputation
- Regeneration involves changes in **gene expression**
- Identify **regeneration-specific genes** involved in tissue repair
- Identify corresponding **human orthologs**
- Reveal **conserved regenerative mechanisms**
- Potential relevance for **human regenerative medicine**

→ Identify regeneration-specific genes in *Nematostella* and determine their human orthologs.



Fibrillins are structural proteins that form **microfibrils** in the **extracellular matrix**. They provide **strength, elasticity, and support** to connective tissues and help **regulate tissue development**.

Bicaudal C protein is an **RNA-binding protein** that regulates the stability and translation of specific **mRNAs**. It plays roles in **development and cell differentiation**.

In conclusion

We found genes specifically associated with regeneration. These code for proteins also present in humans. These insights could potentially lead to advances in regenerative medicine.



Differential Gene Expression in Male & Female Human Breast Mammary Tissue

TEAM 9

Couple 20: Sofia Teixeira, s5096081; Kiana Khalili, s3747204 & Alessandra Mollica, s6824838
Couple 29: Ahmadreza Masoumi, s4240324 & Amir Almarzooqi, s6780210

INTRODUCTION

Sexual dimorphism-oriented research has gained increasing relevance due to growing awareness of sex-related differences in certain human diseases (Tung & Lin, 2026). Breast mammary tissue is present in both males and females, but undergoes distinct developmental and functional changes in response to sex-specific hormonal environments. These differences are reflected in the interactions among epithelial and immune compartments that contribute to tissue development and function (Ibañez-Rios et al., 2026).

Comparative transcriptomic studies of breast mammary tissue have identified:

- Conserved secretory and epithelial gene programs (Saitou, 2026);
- Breast-associated expression patterns;
- Sex-associated differences, including female-biased immune-related programs (Saitou, 2026).

The Genotype-Tissue Expression (GTEx) project provides a large-scale resource for studying human gene expression across tissues using RNA-sequencing data from human donors. In this study, GTEx V11 breast mammary tissue data was used to characterize differences in gene expression between male and female tissue.

RESEARCH QUESTION

To what extent does gene expression differ between male and female human breast mammary tissue, and what biological processes underlie these sex-associated differences beyond direct sex-chromosome effects?

MATERIALS AND METHODS

Data and Quality Control

GTEx V11 breast mammary tissue RNA-seq data from 514 donors (336 males, 178 females; ages 20–79). Genes with ≥ 1 CPM in ≥ 10 samples were retained (22,016 genes). Library size, RNA integrity number (RIN) and PCA were assessed. Associations involving RIN were tested using Spearman correlation, Wilcoxon rank-sum and Kruskal–Wallis tests.

Differential-expression analysis

DESeq2 model (Love et al., 2014): $\sim \text{AGE} + \text{scaled RIN} + \text{SEX}$. Female vs. male differences were tested using the Wald test with Benjamini–Hochberg correction. Significance: adjusted $p < 0.05$ and $|\log_2\text{FC}| \geq 1$.

Annotation and visualization

Genes were annotated using GENCODE v47. PCA assessed variation by sex, age and RIN. A volcano plot visualized differential expression, while a heatmap showed expression patterns of the strongest sex-associated genes. A barplot summarized the proportion of differentially expressed genes within chromosome types, while chromosome enrichment was tested using Fisher’s exact test. Autosomal differentially expressed genes were analyzed with g:Profiler GO Biological Process enrichment analyses.

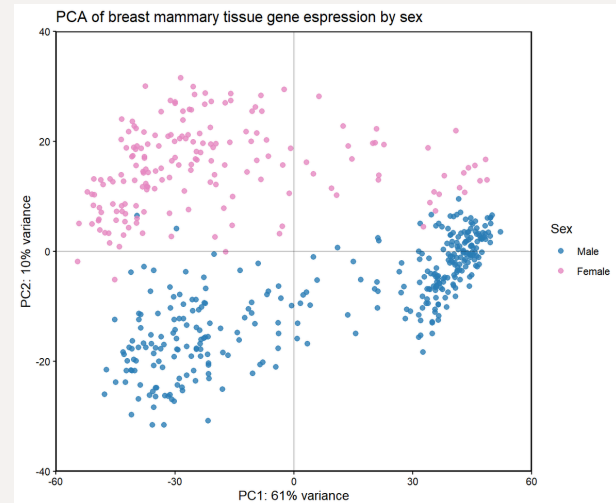


Figure 1. PCA of normalized gene expression profiles in breast mammary tissue samples, colored by biological sex. PC1 and PC2 explained 61% and 10% of the total variance and. Female and male samples showed substantial separation along PC2.

RESULTS

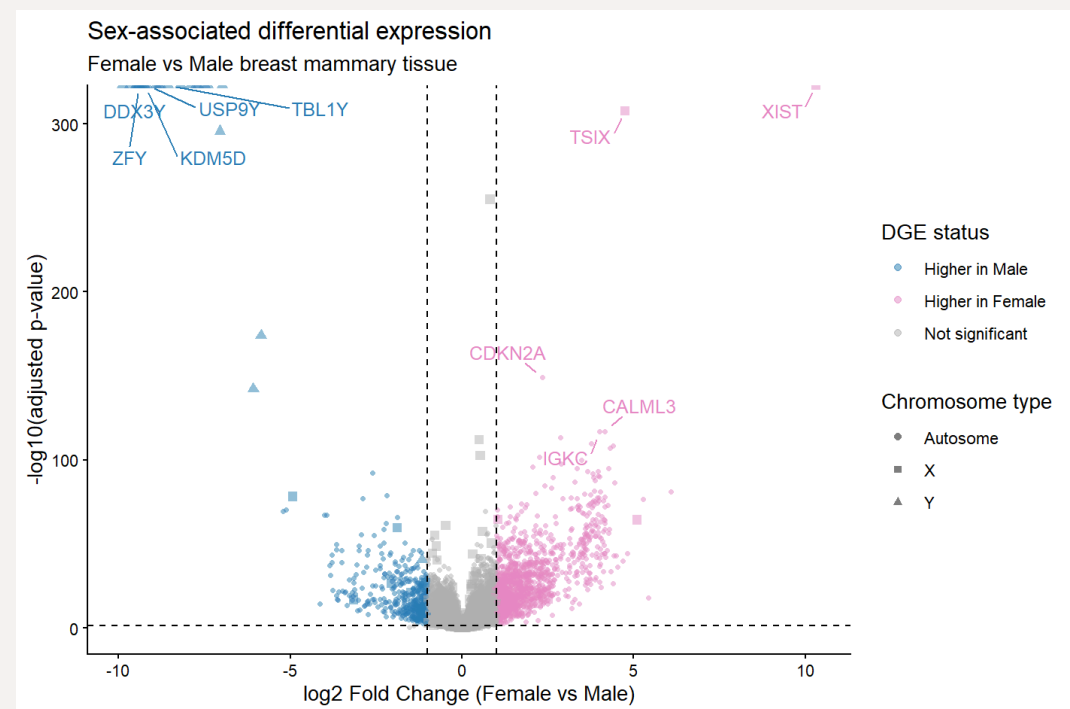


Figure 2. Sex-associated differential expression in breast tissue. Genes are colored according to higher expression in males or females. Shapes distinguish autosomal, X- and Y-linked genes. The 5 most significant differentially expressed genes for each sex were labeled. Highly significant, predominantly Y-linked, genes presented adjusted p-values of approximately 0, extending beyond the displayed y-axis range.

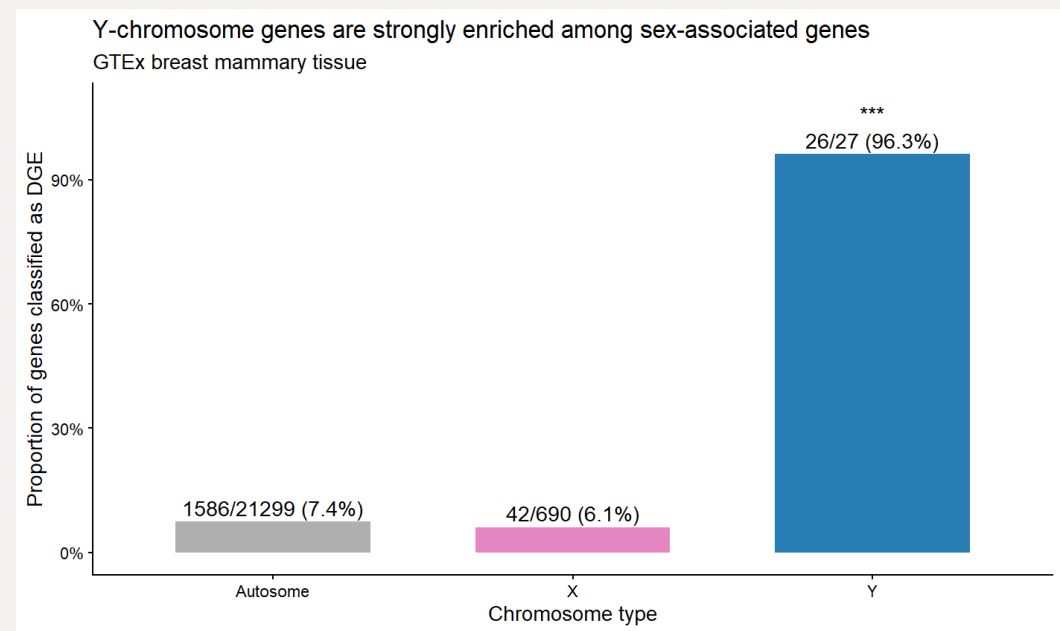


Figure 3. Proportion of genes classified as differentially expressed in autosomal, Y-chromosome and X-chromosome genes. Y-chromosome genes were significantly enriched (Fisher’s exact test, $p < 2.2 \cdot 10^{-16}$), likely due to strong expression of Y-linked genes. The majority of differentially expressed genes were autosomal.

Differential expression analysis identified **1,654 genes** differing between female and male breast mammary tissue (adjusted $p < 0.05$, $|\log_2\text{FC}| \geq 1$) (Fig. 2).

Y-linked genes were strongly enriched among sex-associated genes, contrary to **X-linked genes**. However most significant genes were **autosomal** (Fig. 3).

After excluding genes from sex chromosomes, **1,152 female-** and **434 male-associated** autosomal genes were identified.

Female-associated genes showed enrichment of immune, B-cell and immunoglobulin-related programs, whereas **male-associated genes** were enriched for cell-signaling, membrane excitability and neuronal/synaptic processes (Fig. 4 and 5).

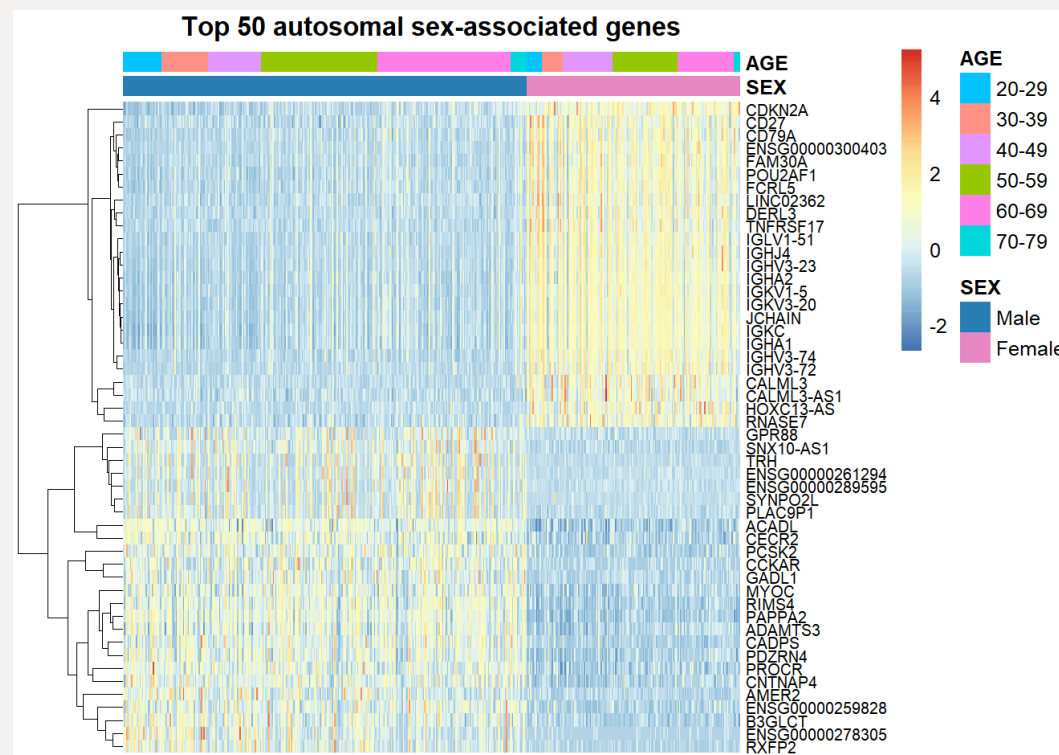


Figure 4. Top 50 autosomal sex-associated genes, consisting of the 25 genes with the strongest differential expression in each direction (higher in females and higher in males). Rows represent genes, while columns represent samples annotated by sex and age. A color gradient between red and blue indicates upregulation and downregulation of each gene compared to the other sex, highlighting sex-associated patterns beyond sex chromosomes.

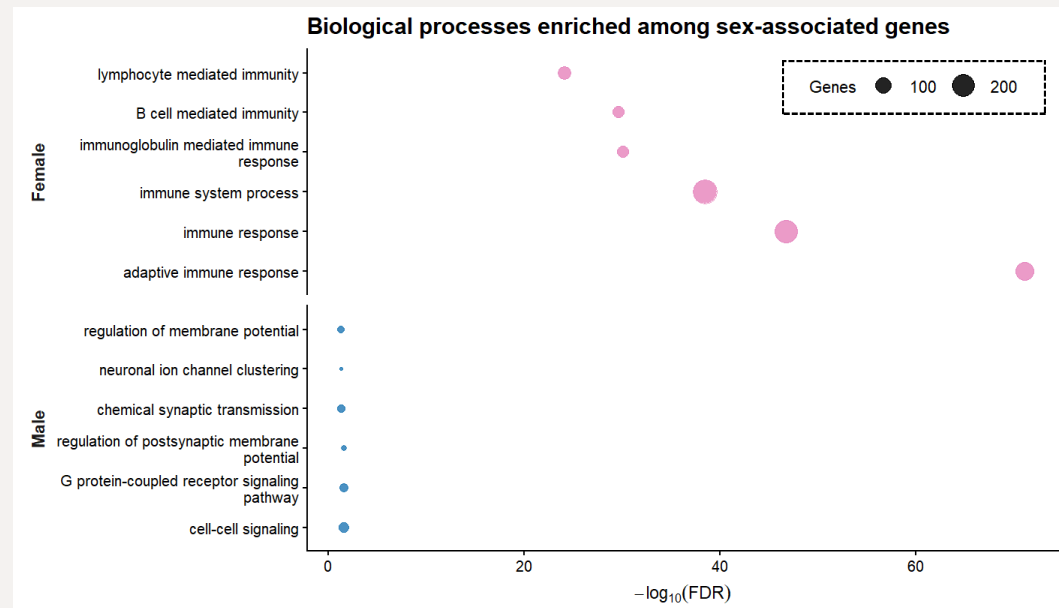


Figure 5. Sex-associated functional enrichment. Female-associated genes were enriched for immune processes, while male-associated genes were enriched for synaptic and cell signaling. Point size indicates number of genes contributing to each biological process and $-\log_{10}(\text{FDR})$ indicates enrichment significance.

CONCLUSIONS

- Breast mammary tissue shows clear **sex-associated transcriptional differences** that extend beyond sex chromosomes, with numerous **autosomal differences** coherent with existing GTEx research (Oliva et al., 2020).
- The strong enrichment of **immune-related processes** in **female-associated genes** is consistent with previous research (Saitou, 2026), supporting their relevance to the biology of the tissue. **Male-associated genes** showed distinct enrichment in **synaptic and cell-signaling processes**, further demonstrating that sex differences are reflected in broader autosomal transcriptional programs.
- Transcriptomic analysis therefore, provides a molecular perspective on sex-specific breast tissue biology, complementing known differences in tissue development and function, highlighting the importance of considering sex when studying breast mammary tissue. Further **cell-type and pathway-level analyses** are needed to determine how transcriptional differences relate to the developmental and pathological differences in breast mammary tissue.

REFERENCES & AI statement

- GENCODE. (n.d.). Human release 47 (GRCh38.p14). https://www.gencodegenes.org/human/release_47.html
- GTEx Consortium. (2020). The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science*, 369(6509), 1318–1330. <https://doi.org/10.1126/science.aba1278>
- GTEx Portal. (n.d.). GTEx portal. Retrieved October 8, 2026, from <https://www.gtexportal.org/>
- Ibañez-Rios, M.-I., Aalam, S. M. M., Ritting, M. L., Jore, A., Chaludis, K., Empenmal, C. P., Jakub, J. W., McLaughlin, S. A., Degnim, A. C., Boughhey, J. C., Couch, F. J., Yadav, S., Sadanandam, A., Sherman, M. E., Radisky, D., Knapp, D. J. H. F., & Kannan, N. (2026). The human male mammary gland has similar epithelial populations to female but distinct composition and transcriptional properties [Preprint]. *bioRxiv*. <https://doi.org/10.64898/2026.03.27.714915>
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15, Article 550. <https://doi.org/10.1186/s13059-014-0550-3>
- Oliva, M., et al. (2020). The impact of sex on gene expression across human tissues. *Science*, 369(6509), eaaba3066. <https://doi.org/10.1126/science.aba3066>
- Saitou, M. (2026a). Comparative transcriptomics of human breast tissue suggests conserved epithelial secretory programs and tissue-associated regulatory specialization [Preprint]. *bioRxiv*. <https://doi.org/10.64898/2026.12.19.695381>
- Saitou, M. (2026b). Comparative transcriptomics suggests that breast secretory epithelium reuses gene repertoires conserved across vertebrates beyond mammals. *Genome Biology and Evolution*, 18(8), ega150. <https://doi.org/10.1093/gbe/egaa150>
- Tung, K. F., & Lin, W. C. (2026). A visualization tool for individual gene expression profiles among males and females in GTEx tissues. *Biology of Sex Differences*, 17(1). <https://doi.org/10.1186/s13293-025-00796-3>

AI Statement:
In this poster, Artificial Intelligence (AI) was used in compliance with the University of Groningen’s guidelines on AI for students of the Biology cluster. These can be found by accessing <https://student.portal.rug.nl/infonet/studenten/tse/programmes/msc-bms/all-years/guidelines-on-ai-for-students-biology-cluster>. The model Chat GPT-5.6 Luna (<https://chat.openai.com/>) was used for brainstorming, creating an outline, improving R code, feedback on figure quality, re-structuring text and checking grammar.

Differences in Gene Expression Between COPD Patients and Healthy Controls



university of
 groningen

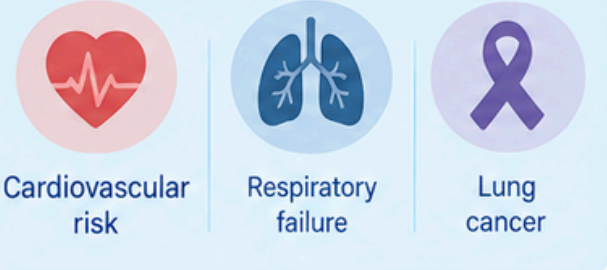
Eline Boiten (S3955001), Jillian Westerling (S4639774), Marisa Horstra (S6730000), Loes Niamut (S4765338) & Esmee Douma (S5170427)
October, 9 2026

INTRODUCTION

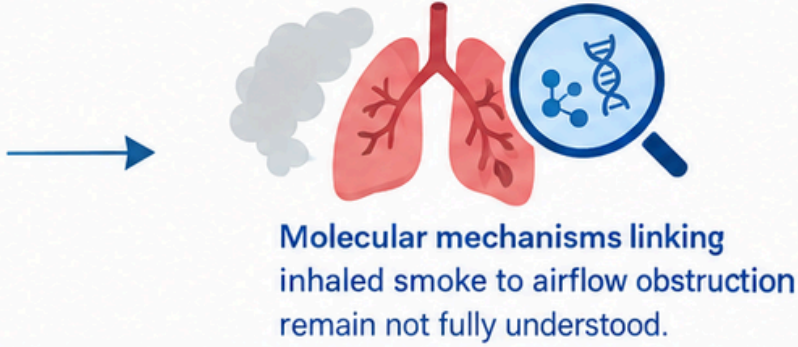
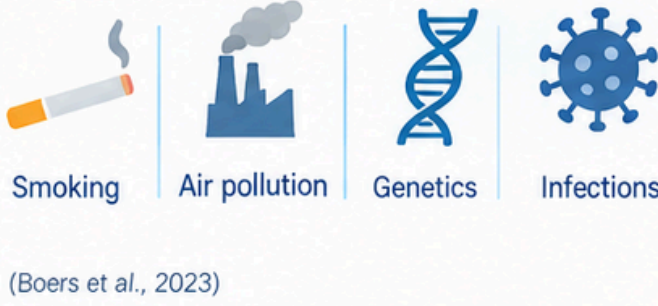


COPD
A global health challenge affecting millions of people.
3rd leading cause of death
Associated with socioeconomic and health consequences
(De Oca et al., 2025)

Major consequences



Risk factors



RESEARCH AIM
Identify the gene most strongly associated with COPD and investigate differences in its expression between COPD patients and healthy control subjects.



RESEARCH QUESTION
Which gene is most significantly associated with COPD, and how does the expression of this gene differ between COPD patients and healthy control subjects?

TOOLS

Dataset:
We used the RNA-seq dataset from Kim et al. (2015) (NCBI GEO accession GSE57148).

It consists of lung tissue samples from 98 COPD subjects and 91 control subjects, all being male smokers or ex-smokers.

- Tools:**
- R studio
 - BioRender
 - Gemini / ChatGPT: used for image editing and rewriting

References:



FIGURES

Volcano plot

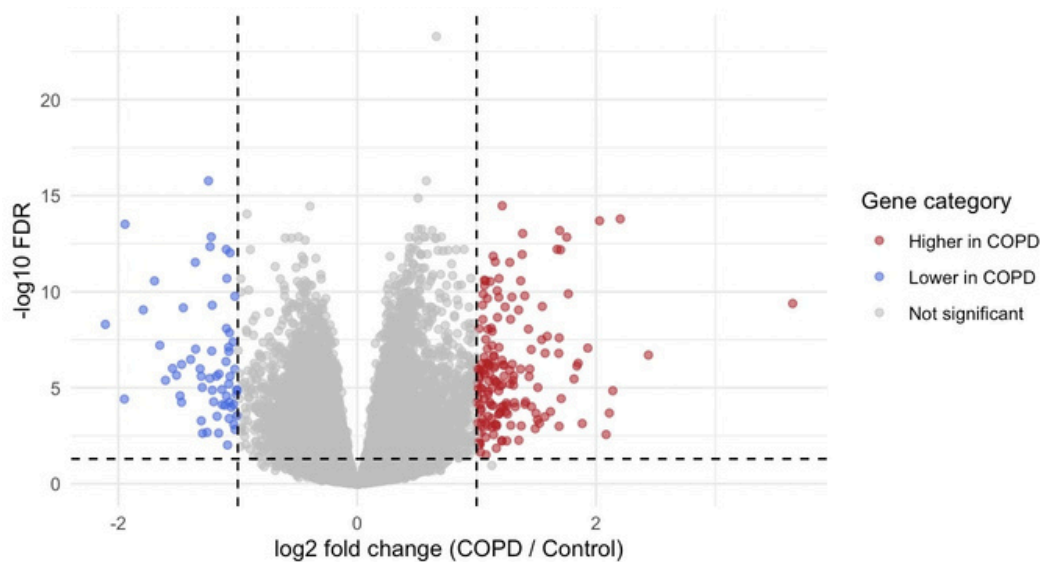


Figure 1. Differential gene expression in COPD lung tissue. RNA-seq analysis (GSE57148; 98 COPD vs. 91 controls) identified differentially expressed genes adjusted for age and pack-years. The x-axis shows log2 fold change (COPD / Control) and the y-axis shows log10 FDR. Genes with FDR < 0.05 and log2 FC ≥ 1 (dashed lines) were deemed significant. Red points indicate higher expression in COPD, blue points lower expression, and grey points non-significant genes.

MA Plot - COPD vs Control

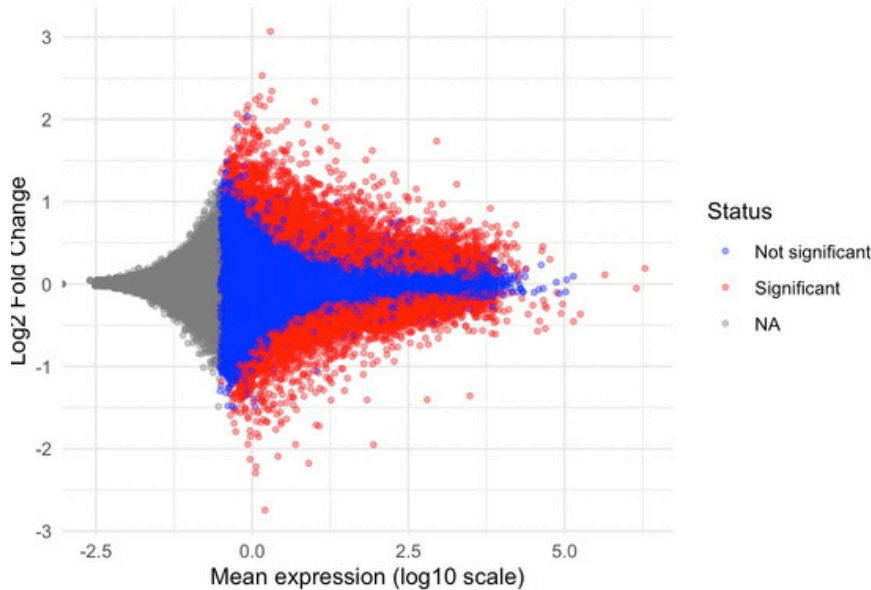


Figure 2. MA Plot - COPD vs Control. MA plot comparing gene expression between COPD patients and control subjects. The x-axis represents the mean expression level on a log10 scale, and the y-axis shows the log2 fold change (COPD vs. Control). Each dot represents an individual gene. Genes are colored based on their statistical significance: significantly differentially expressed genes are highlighted in red, non-significant genes are shown in blue, and genes with missing or excluded values (NA) are shown in grey.

Expression of most significant gene: ENSG0000023199
p-value: 1.51e-29

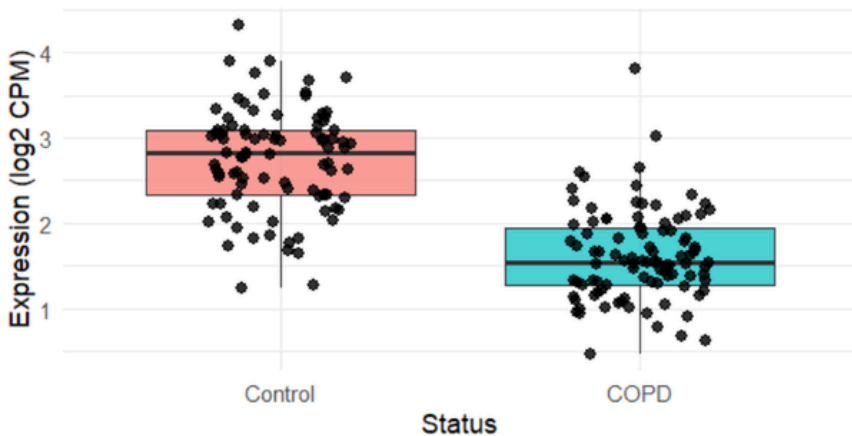


Figure 3. Box Plot. Expression Levels of ENSG00000231991 in Control vs. COPD Patients. Expression was measured by RNA sequencing as log2 counts per million (log2 CPM). The boxplots show the median and interquartile range (IQR), with individual black dots representing biological samples. ENSG00000231991 expression was significantly lower in COPD lung tissue compared to healthy controls. A Welch's t-test showed a significant difference ($p = 1.51 \times 10^{-29}$).

CONCLUSION



- **ENSG000002391** was identified as the most significant COPD-associated differentially expressed gene and was downregulated in COPD lung tissue.



Take-home:
ENSG0000023991 is a promising candidate for further functional and therapeutic investigation.

FUTURE STUDIES



- 1 Analyse loss-and-gain of function



- 2 Single-cell RNA sequencing



- 3 Therapeutic target validation