Convergent Transcriptional Regulation of Endothelial Dysfunction by Oxidative Stress and Inflammatory Stimulation: A Multi-Dataset Gene Expression Meta-Analysis

Dayakshin Arumugam
Emerald High School, Dublin, CA
** Corresponding Author Email: dayakshin.arumugam@gmail.com*

ABSTRACT

Endothelial cell activation is both an early and critical event in atherosclerosis and vascular inflammatory disease, which is caused by reactive oxygen species (ROS) and pro-inflammatory cytokines. Nevertheless, it is still unclear if these two different stimuli can activate the same program of gene expression related to the NF-κB signalling and leukocyte-adhesion pathways.

This study used 9 publicly available GEO transcriptomic datasets from human endothelial cells that were exposed to oxidative stress (H_2O_2 and other related redox issues) or inflammatory stimulation (TNF-α, IL-1β, LPS), across 154 samples (24 control, 130 stimulated). Following $\log_2$ transformation and per-dataset z-score standardisation to normalize the data, differences in gene expression were observed via an ordinary least squares (OLS) model with dummy variables to account for batch effects. Pathway enrichment was evaluated using over-representation analysis (ORA) and pre-ranked gene set enrichment analysis (GSEA).

A total of 1864 genes showed significant changes in gene expression, of which 1075 genes were upregulated, and 789 genes were downregulated. Of the four specified hub genes, NFKB1, ICAM1, and SELE were significantly upregulated with FDR values of 0.007, 0.011, and 0.018, respectively. For VCAM1, the effect size was consistently positive, but it did not reach significance (FDR value of 0.060). GSEA identified the most positively enriched pathways were TNF signalling (NES = 2.83, FDR $q < 0.001$), NIK/NF-κB signalling (NES = 2.41, FDR $q< 0.001$), and cytokine-cytokine receptor interaction (NES = 2.89, FDR $q< 0.001$).

Human endothelial cells that are subjected to oxidative and inflammatory stimulation produce a convergent gene expression program, where NF-κB activation and subsequent upregulation of leukocyte-adhesion molecules are observed. In both stimulus types, hub genes NFKB1, ICAM1, and SELE act as reproducible convergent hub genes, supporting that they represent possible therapeutic targets in vascular inflammatory disease.

1. INTRODUCTION

The vascular endothelium is responsible for the control of haemostasis, vascular tone, and the movement of immune cells. It is also critical in maintaining a non-thrombogenic but selectively permeable layer between blood and the vessel wall. When the human body is functioning properly, endothelial cells(ECs) are typically in a quiescent and anti-inflammatory state and do not express high levels of adhesion molecules for leukocytes. Instead, they have a more balanced redox phenotype that is more protective in normal body conditions [1]. However, when subjected to appropriate stimuli(i.e., haemodynamic, oxidative, or inflammatory stimuli), the endothelium undergoes change. This transformation, known as endothelial dysfunction, is a shift from the normally anti-inflammatory and anti-thrombotic endothelium towards a more pro-inflammatory and pro-thrombotic state. This dysfunction contributes to the pathogenesis of atherosclerosis, sepsis, and ischaemia/reperfusion injury [2,3].

Endothelial dysfunction can be triggered by two major stimuli. First, there are reactive oxygen species (ROS), such as the hydrogen peroxide (H_2O_2) produced by NADPH oxidase or mitochondrial uncoupling [4]. The second class of stimuli is pro-inflammatory cytokines, such as TNF-α, IL-1, or bacterial lipopolysaccharide (LPS) [5]. Oxidative and inflammatory stimuli can activate the transcription factor nuclear factor-κB (NF-κB), which induces inflammation and is responsible for the increase in expression of vascular cell adhesion molecule-1 (VCAM1), intercellular adhesion molecule-1 (ICAM1), and E-selectin (SELE) [6,7].

Although oxidative and inflammatory stimuli are known to induce transcription factor NF-κB, it is unclear whether these stimuli affect a shared set of genes genuinely or if these responses are similar due to chance. Most prior transcriptomic studies used a single cell line, a single stimulus, and a single measurement platform, which limits how widely their conclusions generalise [8].

We address this gap through an integrated meta-analysis of 9 publicly available GEO transcriptomic datasets [9] that span both oxidative and inflammatory EC stimulation models. We tested three hypotheses: first, that both stimulus types produce overlapping sets of differentially expressed genes (DEGs); second, that these DEGs are enriched for NF-κB signalling and leukocyte-adhesion pathways; and third, that a small set of shared hub genes (VCAM1, ICAM1, SELE, and NFKB1) act as convergent regulators of endothelial dysfunction [10,11]. To test these hypotheses we applied batch-corrected linear modelling, Over-representation analysis (ORA), gene set enrichment analysis (GSEA), and network community detection, deriving a cross-stimulus endothelial activation signature.

2. MATERIALS AND METHODS

2.1 Dataset Curation and Selection

Transcriptomic datasets were identified from the NCBI Gene Expression Omnibus (GEO) [9] through a systematic search for studies of human endothelial cells under oxidative stress or inflammatory stimulation. Twenty candidate datasets were screened and scored for relevance on the basis of cell type (endothelial), stimulus (oxidative or inflammatory), species (human), presence of a control group, and data availability. Thirteen datasets met the inclusion criteria and were downloaded programmatically from GEO together with their supplementary count or expression files. After gene-identifier harmonisation and quality filtering, 9 datasets with mappable expression matrices and consistent sample annotation were retained for the integrated analysis. These 9 datasets, their sources, and short descriptions are listed in Table 1.

GEO accession	Cell model	Stimulus / model	Samples (C/T)	Short description
GSE201466	HUVEC	TNF-α (0–10 h), CTRL vs RELA siRNA	18 (3C / 15T)	Inflammatory NF-κB signalling and its effect on endothelial cholesterol handling.
GSE131590	HUVEC	IL-1α, TNF-α, or LPS (6 h)	41 (0C / 41T)	Direct cytokine and endotoxin stimulation modelling leukocyte-endothelial interaction in infection.
GSE72991	HUVEC	H_2O_2 oxidative-stress time course	20 (2C / 18T)	Temporal oxidative-stress response with heme-oxygenase-1 as a regulator of endothelial cell fate.
GSE296602	HUVEC	TNF-α, acute (5 h) and chronic (9 d)	12 (6C / 6T)	Acute versus chronic TNF-α inflammation and its interplay with endothelial senescence.

GSE275227	HUVEC	TNF-α ± AdipoRon / SKI-I (12 h)	12 (2C / 10T)	TNF-α-induced activation and adiponectin-receptor-mediated endothelial protection.
GSE316445	HUVEC	TNF ± pro-resolving mediator (48 h)	9 (0C / 9T)	TNF-stimulated endothelium during neutrophil interaction and inflammation resolution.
GSE302821	HUVEC	H_2O_2-linked redox metabolite (4/24 h)	12 (6C / 6T)	Redox-sensitive glutamine metabolism under an oxidative challenge in endothelial cells.
GSE235049	HMEC-1	LPS injury ± exosome treatment	20 (0C / 20T)	Combined oxidative and inflammatory endothelial injury with exosome-based protection.
GSE196161	HUVEC	miR-125a-5p vs control miRNA	10 (5C / 5T)	miR-125a-5p regulation of the endothelial barrier under inflammatory conditions.

Table 1. GEO datasets used in the integrated analysis, with clickable accession links to the source records. Accession numbers link to the corresponding GEO series pages. HUVEC: human umbilical vein endothelial cells; HMEC-1: human microvascular endothelial cells; C: control samples; T: stimulated (inflamed or oxidised) samples. Sample labels were assigned from GEO sample titles and characteristics fields. Datasets without an internal untreated arm contributed only stimulated samples to the pooled control-versus-stimulated contrast.

2.2 Data Preprocessing and Normalisation

RNA-seq raw counts were transformed with $\log_2(x + 1)$ to stabilise variance and approximate normality [12]. Datasets already provided on a log scale (variance-stabilising transformation or $\log_2$-normalised counts) or as microarray signal intensities were retained without further transformation. Missing values within each dataset were imputed with the per-gene median. Each dataset was then standardised independently by gene using z-scores (scikit-learn StandardScaler), so that expression levels were comparable across studies before pooling. Genes missing in more than 50% of samples were removed, followed by removal of the least variable genes (bottom 20th percentile of median absolute deviation). The final expression matrix comprised 13,497 genes across 154 samples. Samples were assigned binary labels (0 = control, 1 = stimulated) based on inspection of the GEO sample titles and characteristics fields for each dataset.

2.3 Differential Gene Expression Analysis

Differential expression was estimated with a vectorised ordinary least squares (OLS) linear model of the form $\text{expr} = \beta_0 + \beta_1 \cdot \text{label} + \Sigma\, \beta_{\square} \cdot \text{dataset}_{\square} + \varepsilon$, in which β_1 is the batch-adjusted stimulated-versus-control effect size and the dataset dummy variables absorb between-study variance [13]. The design matrix combined the binary treatment label with 8 dataset indicator columns (drop-first encoding), giving a shared intercept.

For all 13,497 genes, coefficients were found using the normal equations ($B = (X^TX)^{-1}X^TY$). Simultaneously, after adjusting for experimental conditions and differences within data sets, residual variation was measured with 144 degrees of freedom, and label coefficients and standard errors were used to calculate the gene-level t-statistics. To deal with significance due to chance, p-values were corrected with the Benjamini-Hochberg false discovery rate procedure [14]. Genes with FDR below 0.05 and an absolute label coefficient more than 0.3 were statistically reliable and designated as significant DEGS.

2.4 Pathway Enrichment Analysis

Two complementary enrichment strategies were applied with the gseapy package [15]. Using the upregulated and downregulated DEG lists and four gene-set libraries, Enrichr API performed ORA [16,17]. The four gene-set libraries were GO Biological Process 2023 [18], GO Molecular Function 2023, KEGG 2021 Human [19], and Reactome 2022 [20]. The full set of tested genes was used as background statistics. Terms with an FDR-adjusted p-value less than 0.05 were considered significant. Pre-ranked gene set enrichment analysis (GSEA) [21] used the DGE t-statistic as the ranking metric across all 13,497 tested genes, with 100 permutations, a minimum gene-set size of 15, and a maximum of 500. GSEA results were statistically significant if lower than the standard FDR q-value threshold of 0.25.

2.5 Network Analysis

Three gene co-expression networks were constructed with NetworkX [22]. For the first network, a hub-gene co-expression network was created, where Pearson correlation was calculated between the 4 hub genes(VCAM1, ICAM1, SELE, NFKB1) and the significant DEGs. The top 12 correlated partners for each of the hub genes ($|r| \geq 0.40$) were retained as edges. The second was an NF-κB pathway sub-network built from all pairwise correlations among the curated NF-κB and leukocyte-adhesion genes present in the matrix ($|r| \geq 0.35$). The third was a top-DEG community network connecting the top 30 upregulated and 30 downregulated DEGs plus the hub genes at $|r| \geq 0.55$, to which Louvain community detection [23] was applied to identify co-regulated modules. In the pathway and community networks, node sizes were scaled by |t-statistic| and edge widths were proportional to |r|.

2.6 Statistical Software

Analyses were performed in Python 3.11 using pandas [13], NumPy, scikit-learn [12], SciPy, statsmodels, gseapy [15], NetworkX [22], matplotlib, and seaborn.

3. RESULTS

3.1 Dataset Integration and Sample Characteristics

9 GEO datasets containings 154 human endothelial cell samples (24 control, 130 stimulated) were integrated after preprocessing the data(Table 1). The GEO datasets covered both categories of stimulation. In oxidative stress experiments, endothelial cells where exposed to H_2O_2 and other chemicals disturbing redox balance, which are covered in datasets GSE72991, GSE302821, and the oxidative arm of GSE235049. In contrast, for inflammatory experiments, endothelial cells where subject to TNF-α, IL-1, and LPS, as shown in datasets GSE201466, GSE131590, GSE296602, GSE275227, GSE316445, GSE196161, and the inflammatory arm of GSE235049. One study used HMEC-1, while the rest of the studies used HUVEC. After $\log_2$ transformation, per-dataset z-scoring, and gene-level quality filtering, the final expression matrix consisted of 13,497 genes across 154 samples.

3.2 Differential Gene Expression

In total 1,864 genes (of 13,497 tested, 13.8% of genes tested) were found to be significantly differentially expressed (FDR < 0.05, $|\beta| > 0.3$) in simulated compared with control group samples, using the batch-corrected OLS model for differential gene expression. Among the 1,864 genes, 1075 where upregulated and 789 were downregulated, 58% and 42% respectively.(Figure 1).

RRAD, a Ras-related GTPase that is involved in dealing with oxidative and metabolic stress, is the most upregulated gene ($\beta = 1.25$, $t = 7.35$, FDR = 1.8×10^{-7}). It is followed by a peak of expression for ADM2 ($\beta = 1.21$), an adrenomedullin-2 peptide that is induced under hypoxic and inflammatory conditions, as well as several TNF-receptor and NF-κB-regulated genes, including

TNFRSF9 and TNIP3. Among the genes downregulated were MCM2 (β = −1.04, FDR = 0.0017), a DNA replication licensing factor and SUV39H1 (β = −1.01), a histone H3K9 methyltransferase. This pattern matches G1-phase cell-cycle arrest during endothelial activation [1].

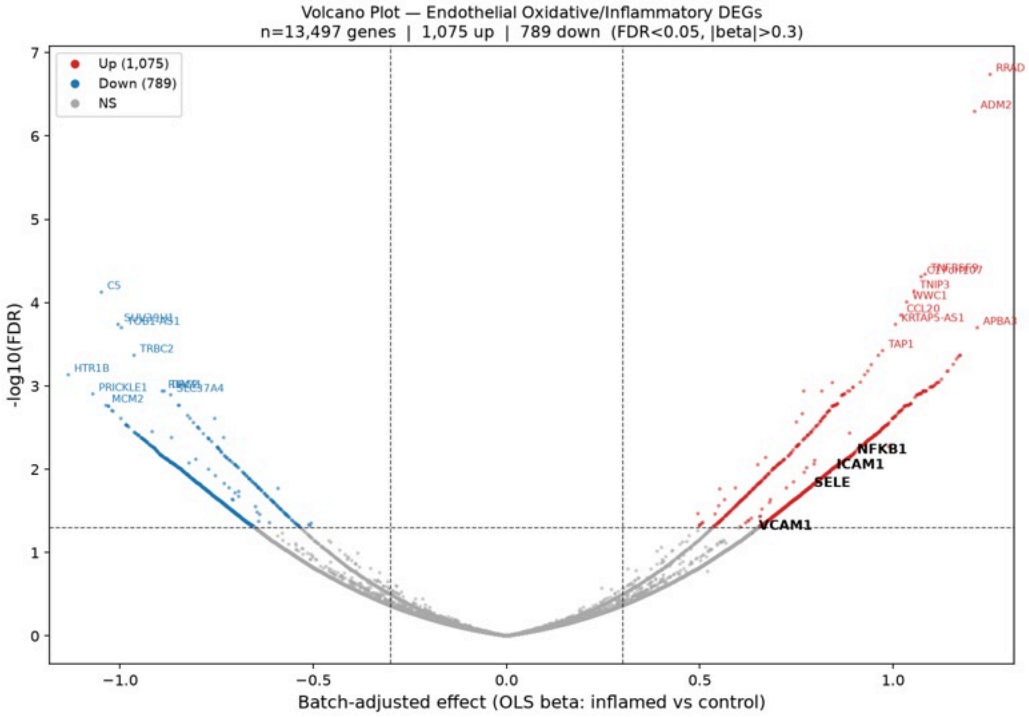


Figure 1. Volcano plot of differential gene expression in stimulated vs. control human endothelial cells. The x-axis shows the batch-adjusted OLS β-coefficients for stimulated vs. control conditions. The y-axis represents $-\log_{10}$(FDR). The red dots on the volcano plot show significantly upregulated genes (FDR < 0.05, β > 0.3, n = 1,075), blue dots represent significantly downregulated genes (n = 789). Non-significant points are denoted by grey points. The hub genes VCAM1, ICAM1, SELE, NFKB1 are in bold, and significance thresholds are represented with dashed lines. This plot includes 13,497 genes(1075 upregulated and 789 downregulated) taken from 154 samples across 9 GEO datasets.

Hierarchical clustering of the top 80 DEGs, where 40 genes are upregulated and 40 are downregulated, nearly separated control from stimulated samples (Figure 2). The cluster that is upregulated showed similar expression in both oxidative-stress and inflammatory datasets, highlighting the convergent path of transcriptional response.

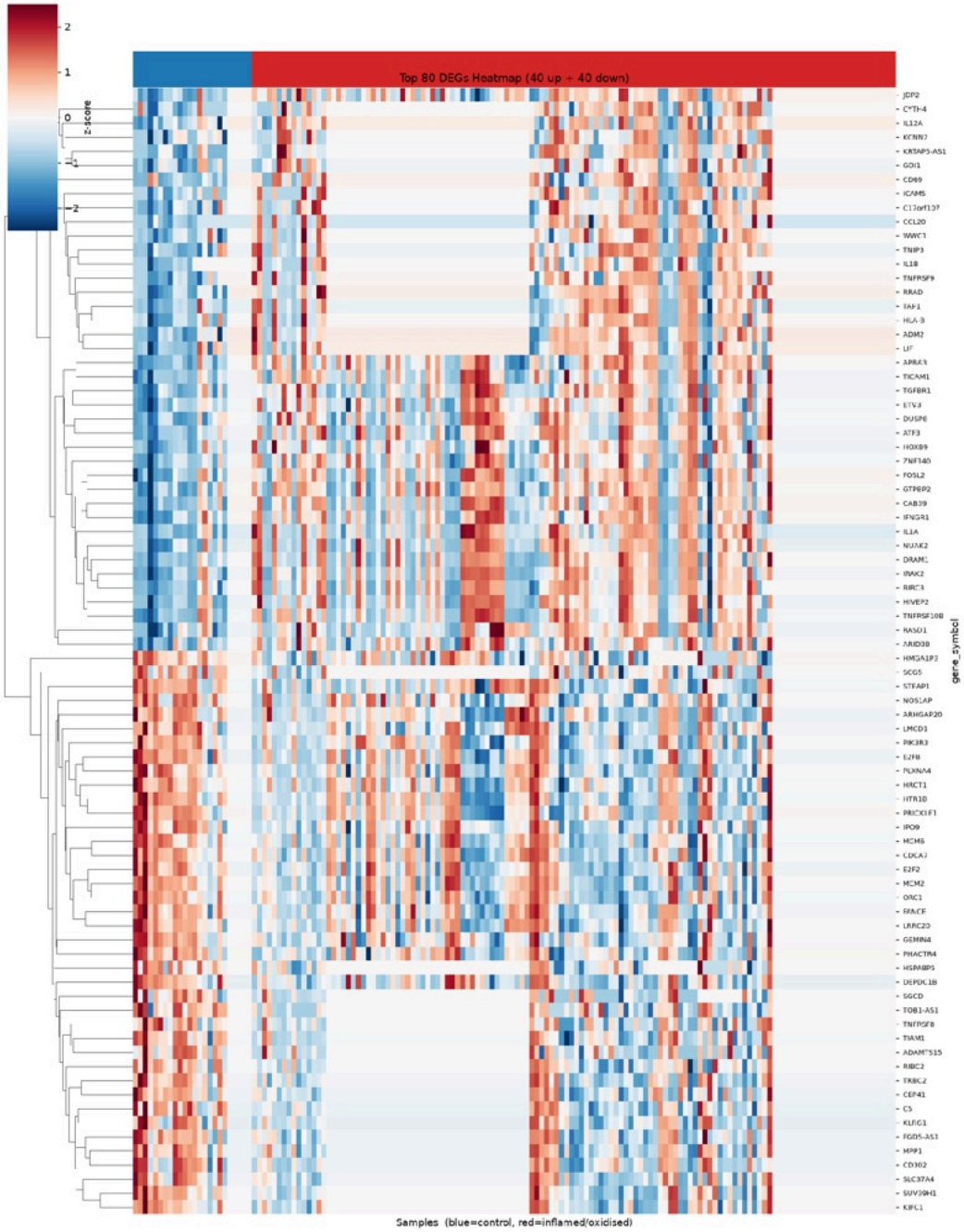


Figure 2. Heatmap of the top 80 DEGs, where 40 genes are upregulated and 40 are downregulated. Rows are genes which are clustered hierarchically. Columns are samples where control is on the left and stimulated is on the right. The color scale represents z-score expression values; the blue column bar marks control samples, and the red bar marks stimulated samples. The upregulated cluster (top) and downregulated cluster (bottom) demarcate the two phenotypic states across all datasets.

3.3 Hub Gene Analysis

Gene	β (effect size)	t-statistic	FDR	Significance	Endothelial function
VCAM1	0.630	2.64	0.0601	(*)	Mediates firm leukocyte adhesion via α4-integrin; NF-κB target; early marker of endothelial activation
ICAM1	0.831	3.55	0.0112	**	Binds LFA-1 and Mac-1; required for trans-endothelial migration; induced by cytokines and ROS
SELE	0.772	3.28	0.0184	**	E-selectin; mediates initial leukocyte tethering; rapidly induced by NF-κB
NFKB1	0.884	3.80	0.0073	**	NF-κB p50 subunit; transcriptional regulator of VCAM1, ICAM1, SELE, and many inflammatory genes

Table 2. Hub gene differential expression results. β is the OLS label coefficient (stimulated versus control); ** denotes FDR < 0.05, (*) denotes FDR < 0.10, and ns denotes not significant. All four genes show a positive β, consistent with upregulation in stimulated endothelium.

NFKB1 was the most significantly upregulated hub gene (β = 0.884, t = 3.80, FDR = 0.007), followed by ICAM1 (β = 0.831, FDR = 0.011) and SELE (β = 0.772, FDR = 0.018). VCAM1 showed a consistent positive effect size (β = 0.630) at borderline significance (FDR = 0.060). This most likely reflects the greater stimulus-specificity of VCAM1 induction, since cytokines induce VCAM1 more potently than H_2O_2 alone [10,5], which increases between-dataset variance in the pooled model. Boxplots of z-score expression for the four hub genes are shown in Figure 3.

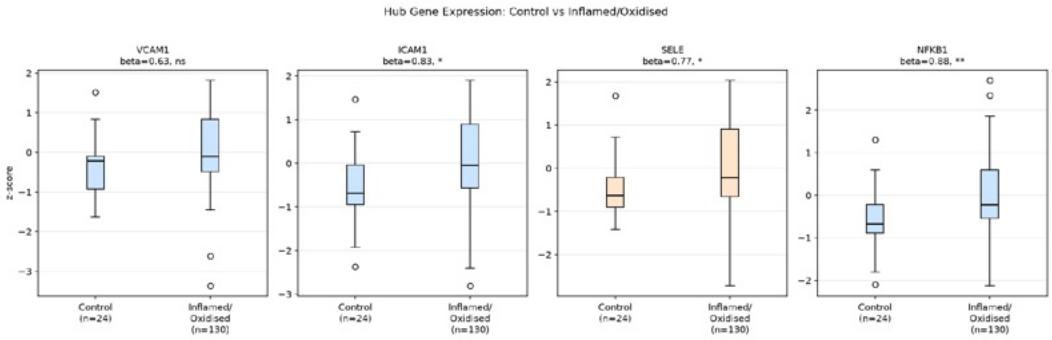


Figure 3. Boxplots of the expression of hub genes VCAM1, ICAM1, SELE, and NFKB1 in control vs. stimulated endothelial cells. The z-score expression values for the four genes in stimulated cells are plotted as boxplots, along with corresponding values for significance from the OLS model FDR (** FDR < 0.01; * FDR < 0.05; ns: not significant). All four genes trend positive in stimulated cells, and NFKB1, ICAM1, and SELE reach statistical significance.

3.4 NF-κB Signalling Cascade Activation

Of the 29 genes from the NF-κB signalling and leukocyte-adhesion pathways analyzed, 16 (55%) were significantly differentially expressed (FDR < 0.05, |β| > 0.3). In addition to NF-κB components being upregulated, the NF-κB subunits NFKB1, NFKB2, RELB and REL, the IKK complex member IKBKB, the negative feedback components NFKBIA, NFKBIB and TNFAIP3/A20, and the adaptor protein TRAF1 increased. The anti-apoptotic component BIRC3, the adhesion components ICAM1, SELE and SELP, and the chemokines and inflammatory components CCL2, CXCL10, IL1B, PTGS2 were also genes that were overexpressed. As described by Hayden and Ghosh, the upregulation of NF-κB activators and their inhibitors (NFKBIA, NFKBIB, TNFAIP3) is indicative of the self-limiting NF-κB oscillation. Theese oscillations are where the cell induces the inflammatory transcriptome while, at the same time, co-inducing molecular brakes that prevent sustained NF-κB nuclear localisation.

3.5 Pathway Enrichment Analysis

3.5.1 Over-Representation Analysis

ORA of the 1,075 upregulated DEGs found significant enrichment for all four gene-set libraries (Table 3). The top enriched KEGG pathways for upregulated DEGs were cytokine-cytokine receptor interactions (FDR = 9.8×10^{-18}) and TNF signalling (FDR = 6.6×10^{-17}); both of theese pathways are important for NF-κB activation [6]. Reactome analysis confirmed enrichment for the Immune System (FDR = 2.6×10^{-19}) and Cytokine Signalling in the Immune System (FDR = 2.6×10^{-16}) [20]. ORA bar charts for the upregulated DEGs are shown in Figure 4.

Pathway	Library	Overlap	FDR
Cytokine-cytokine receptor interaction	KEGG	86/539	9.78e-18
TNF signalling pathway	KEGG	47/108	6.56e-17
Necroptosis	KEGG	39/159	1.67e-09
Immune System	Reactome	245/2682	2.57e-19
Cytokine Signalling in Immune System	Reactome	149/1092	2.56e-16
Signalling by Interleukins	Reactome	90/472	3.57e-08
Cytokine Receptor Binding (GO:0005126)	GO MF	54/275	1.02e-05
NIK/NF-κB Signalling (GO:0038061)	GO BP	32/64	<0.001

Table 3. Selected significantly enriched pathways among the upregulated DEGs by ORA. Overlap gives the number of DEGs in the pathway relative to the total pathway size. FDR: Benjamini-Hochberg adjusted p-value.

The downregulated DEG set contains genes involved in cell-cycle and DNA-metabolism. These genes form many significant sets, including cell cycle (FDR = 6.8×10^{-18}), DNA replication (FDR = 6.9×10^{-13}), and DNA repair sets, most notably homologous recombination (FDR = 1.1×10^{-7}). This G1-phase arrest signature is a shared feature of both cytokine- and ROS-induced endothelial activation, reflecting a switch from a proliferating to a secretory, pro-inflammatory state [1]. The reciprocal enrichment of immune and cytokine processes among upregulated

genes and of cell-cycle and DNA-metabolism processes among downregulated genes is shown in Figure 5.

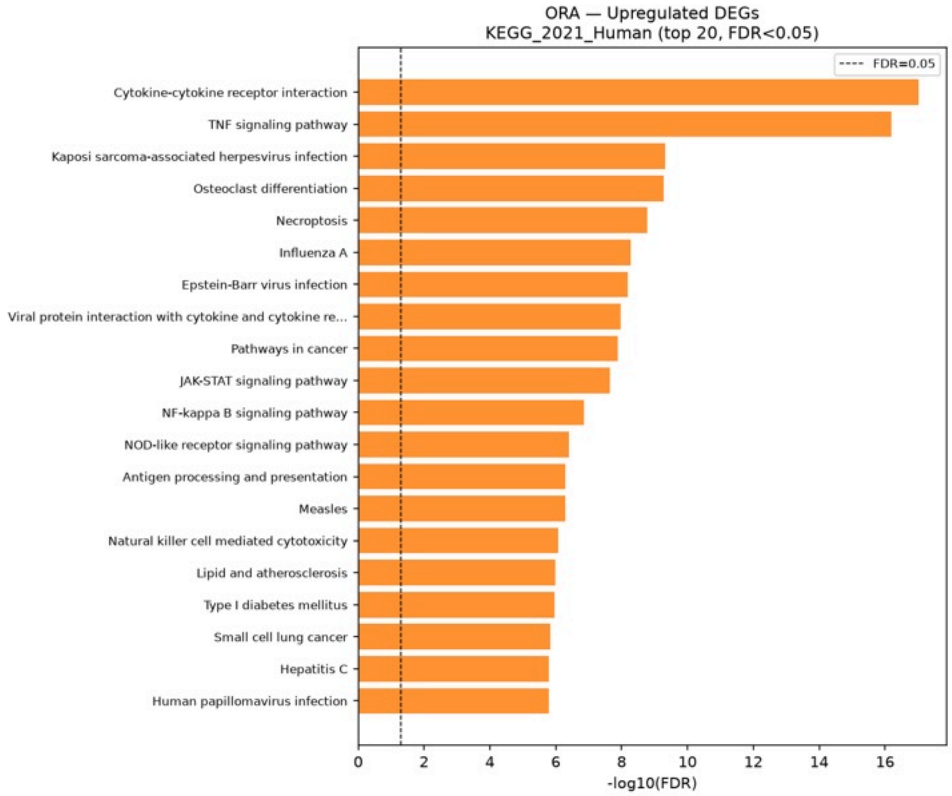


Figure 4. ORA bar chart of significantly enriched KEGG pathways (FDR < 0.05) among the upregulated DEGs (n = 1,075). The top 20 terms are shown in descending order. Note that the dashed line at −log10(0.05) separates terms with FDR values greater than 0.05 from those with FDR values less than 0.05. There is a visible difference where cytokine and TNF signalling pathways are highly ranked, which matches the NF-κB-centred transcriptional response.

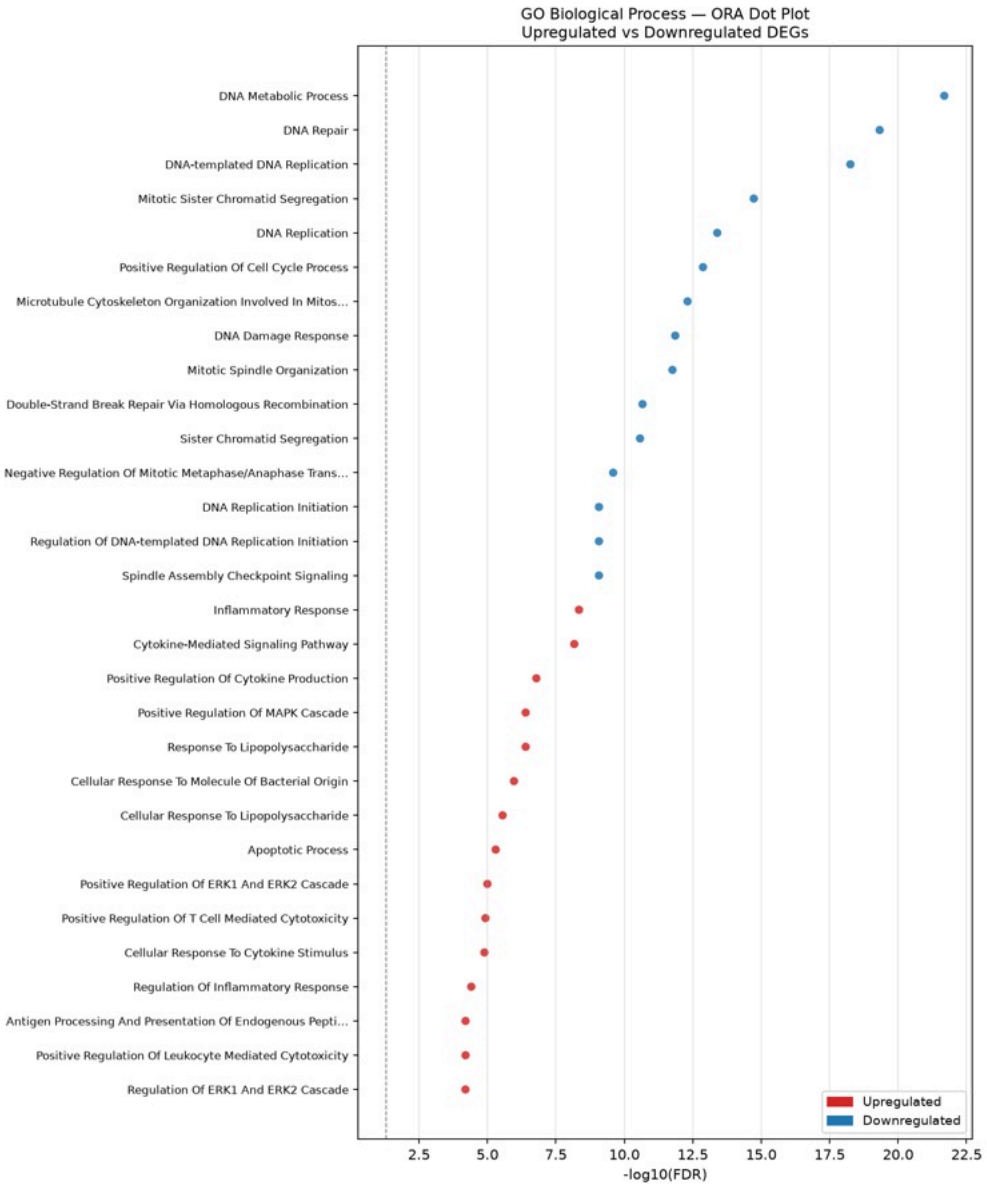


Figure 5. Combined GO Biological Process ORA dot plot for upregulated (red) and downregulated (blue) DEGs. Point size of dots corresponds to gene overlap count. The x-axis represents $-\log_{10}$(FDR). For both upregulated and downregulated genes, the reciprocal enrichment between immune and cytokine and between cell-cycle/DNA-metabolism genes matches with endothelial phenotypic switching.

3.5.2 Pre-ranked GSEA

GSEA using the DGE t-statistic ranking across all 13,497 genes identified 718 positively enriched and 204 negatively enriched GO Biological Process gene sets (FDR < 0.25), and 147 positively and 41 negatively enriched KEGG gene sets. Normalised enrichment scores (NES) for gene sets reached a high for cytokine-cytokine receptor interaction (NES = 2.89, FDR = 0),

for TNF signalling (NES = 2.83, FDR = 0) and for NIK/NF-κB signalling (NES = 2.41, FDR = 0), thus supporting the hypothesis also at the gene-set level. Negatively enriched gene sets were dominated by cell-cycle and DNA-repair pathways (DNA replication NES = −2.98; Fanconi anaemia pathway NES = −2.63), consistent with the ORA findings.

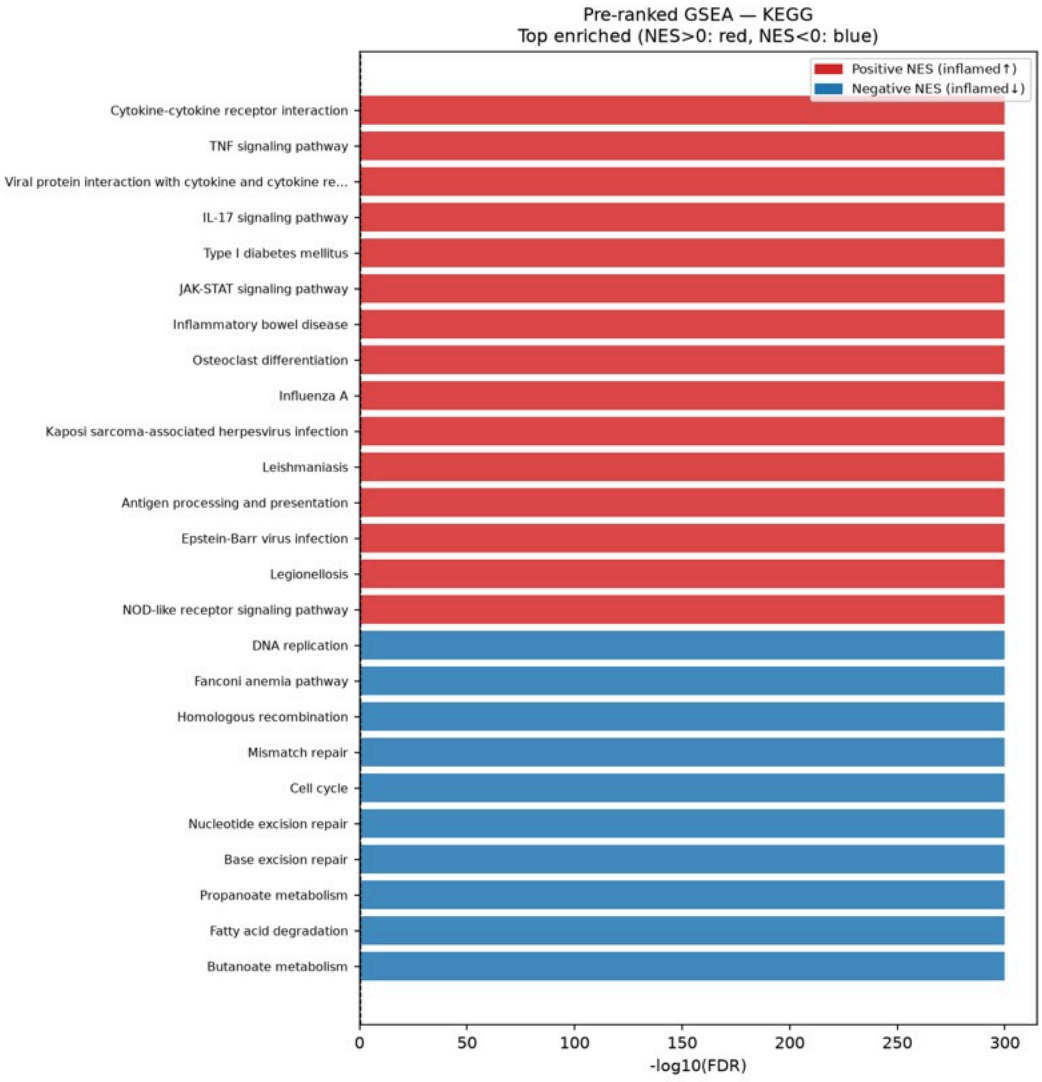
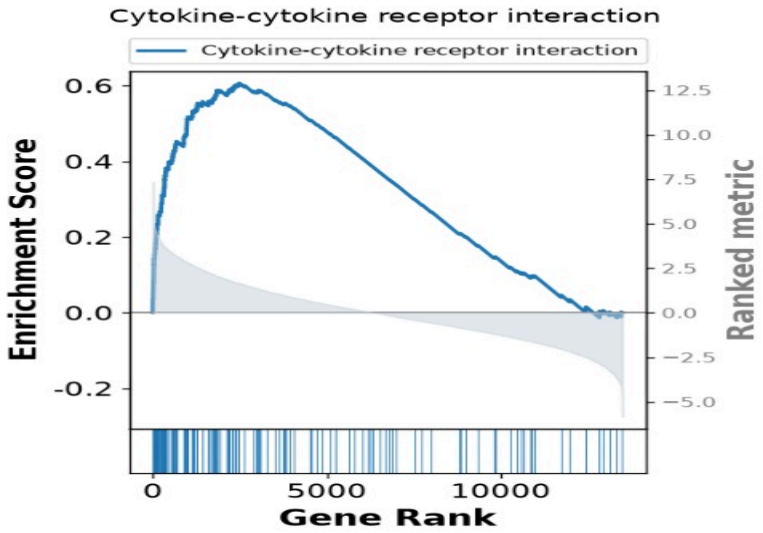
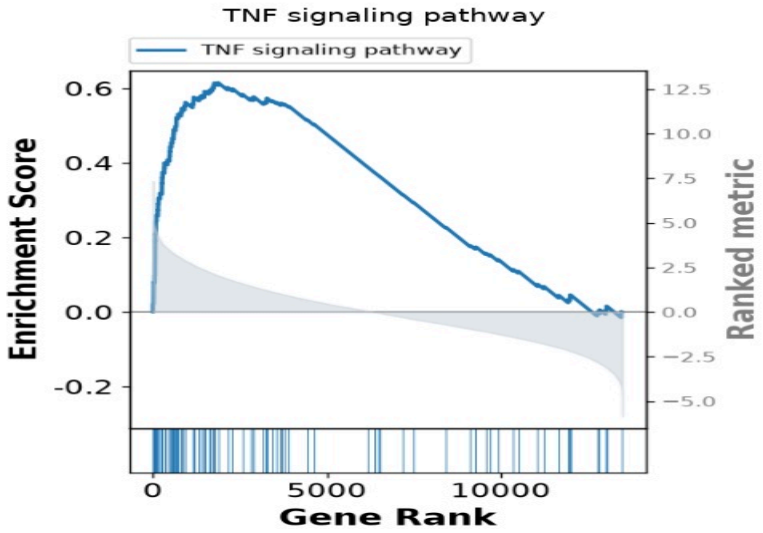


Figure 6. Pre-ranked GSEA bar chart for KEGG gene sets. The red bars show which pathways were expressed more strongly (NES > 0, higher in stimulated endothelial cells). The blue bars show each pathway where expressed less strongly (NES < 0, lower in stimulated endothelial cells). The dashed mark line at FDR = 0.25 shows the threshold. In this figure, TNF signalling cytokine pathways are the most enriched. On the other hand, DNA replication and cell cycles were the most suppressed.

GSEA Enrichment Curves — KEGG
NF-κB / Inflammatory terms





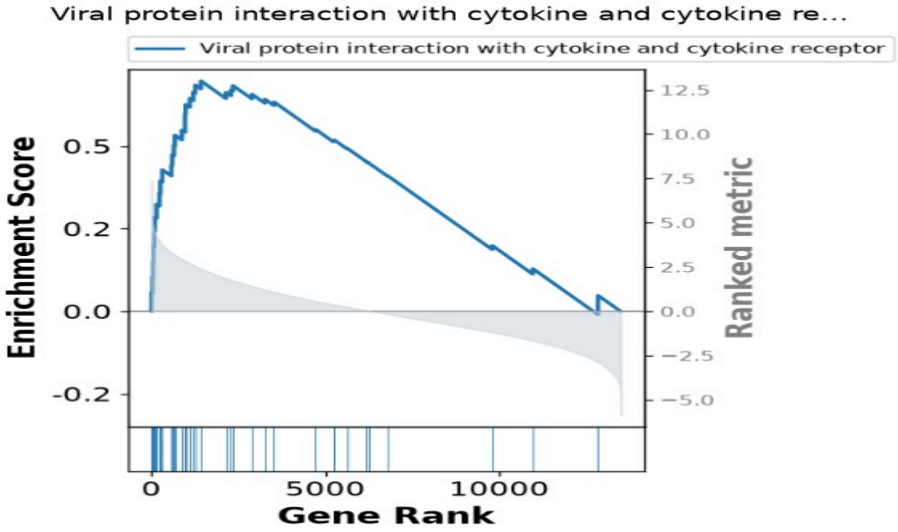


Figure 7. The GSEA enrichment-score curves of the top three NF-κB- and cytokine-related KEGG gene sets, which are cytokine-cytokine receptor interaction, TNF signalling, and IL-17 signalling. The figure shows the running enrichment score on the top. The positions of the pathway genes on the list are in the middle. The ranking metric with t-stastic is on the bottom. On the figure the leftward peak of each curve shows that the pathway genes are concentrated with the most upregulated genes in the stumlated endothelial cells.

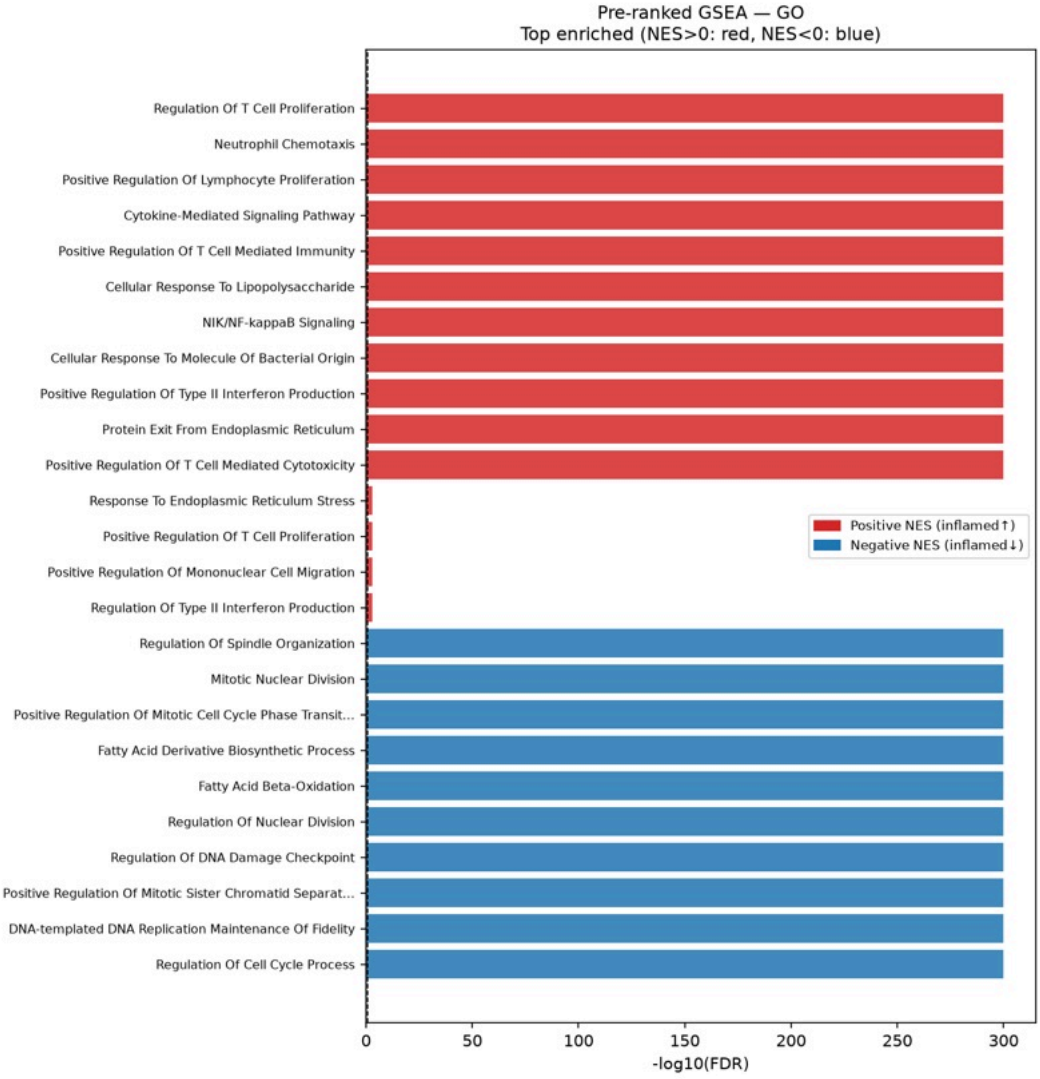


Figure 8. Pre-ranked GSEA bar chart for GO Biological Process gene sets (top 15 positive and top 10 negative NES). Inflammatory immune processes dominate the positively enriched sets, while mitotic division and fatty acid metabolism are negatively enriched.

3.6 Gene Co-Expression Network Analysis

Network analysis of hub-gene co-expression (Figure 9) showed that NFKB1, ICAM1, SELE, and VCAM1 each share high-correlation connections ($|r| \geq 0.40$) with distinct but partially overlapping

sets of inflammatory DEGs. Edges between the hub genes themselves were also detected, indicating co-regulation of all four hubs in stimulated endothelial cells. The network comprised 43 nodes and 54 edges, with the hub genes occupying higher-degree positions than peripheral DEGs.

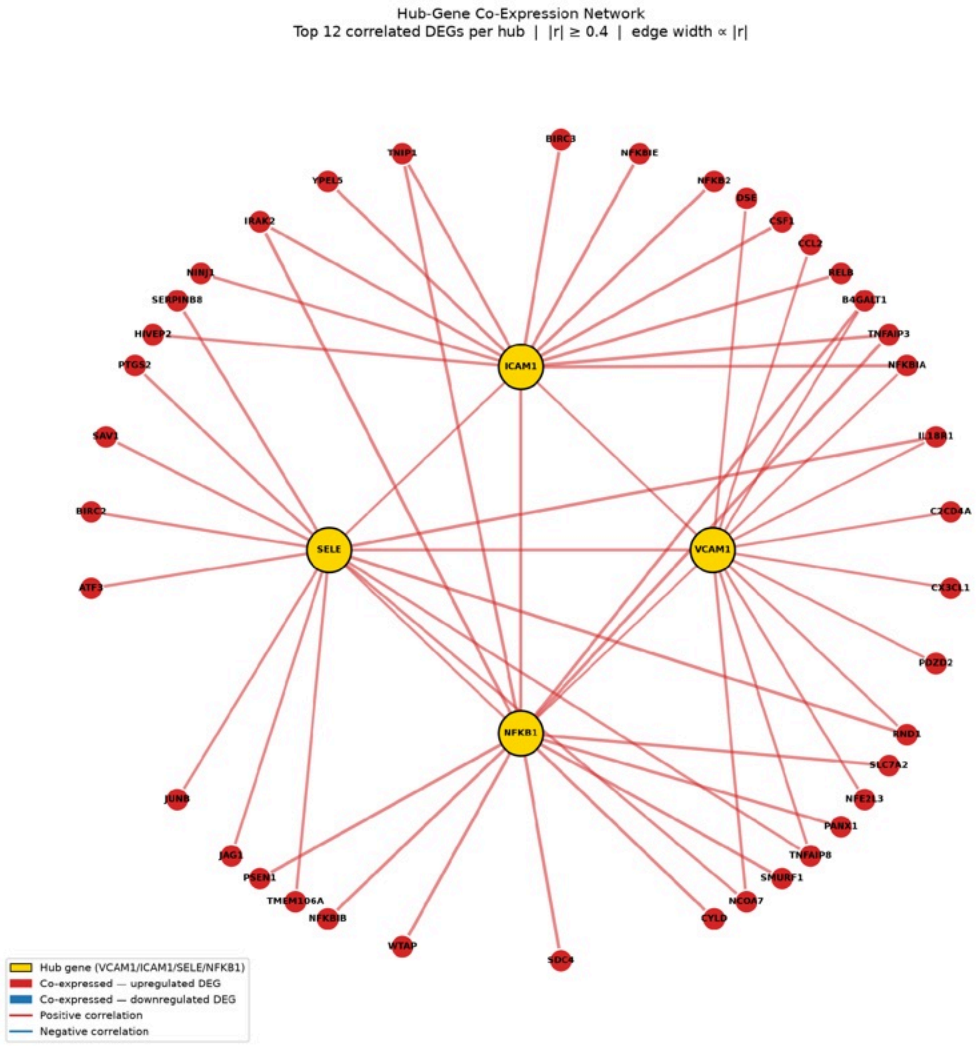


Figure 9. Hub-gene co-expression network. The gold nodes represent VCAM1, ICAM1, SELE, NFKB1 which are the hub genes. Red nodes represent co-expressed upregulated DEGs, while blue nodes show co-expressed downregulated DEGs. The red edges show positive correlation and blue edges indicate negative correlation, the width of theese edges are also proportional to |r|. In the figure, only connections with |r| ≥ 0.40 are visible. The network contains 43 nodes and 54 edges.

The NF-κB pathway sub-network (Figure 10) among 25 curated pathway genes present in the expression matrix showed dense connectivity (222 edges at |r| ≥ 0.35), indicating that NF-κB

activators, inhibitors, and effector molecules form a tightly co-regulated module in stimulated endothelium [7].

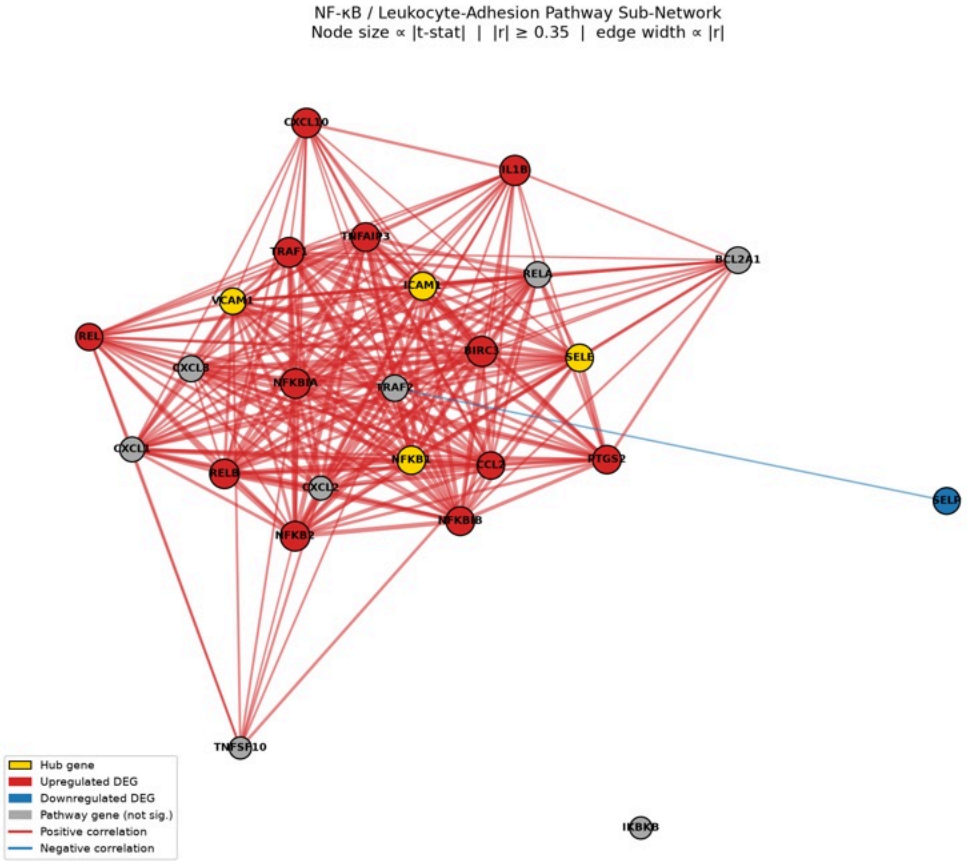


Figure 10. NF-κB and leukocyte-adhesion pathway sub-network. Nodes have different colors based on their expression. Gold nodes represent hub genes. Red nodes represent upregulated DEGs, while blue nodes indicate downregulated DEGs. Grew nodes are not significant. The node size of each node is proportional to |t-statistic|. Edges show Pearson correlations that have |r| ≥ 0.35. The network holds 25 nodes and 222 edges, and is densely connected, indicating a tightly co-regulated inflammatory module.

Louvain community detection in the top-DEG co-expression network (Figure 11) identified five distinct gene modules, which suggests that the endothelial activation programme is organised into semi-independent regulatory units (for example, cytokine-signalling, cell-cycle arrest, adhesion-molecule, antigen-presentation, and redox-response modules) rather than a single undifferentiated response [23]. The hub genes occupied bridging positions between multiple communities, consistent with their proposed role as convergent regulators.

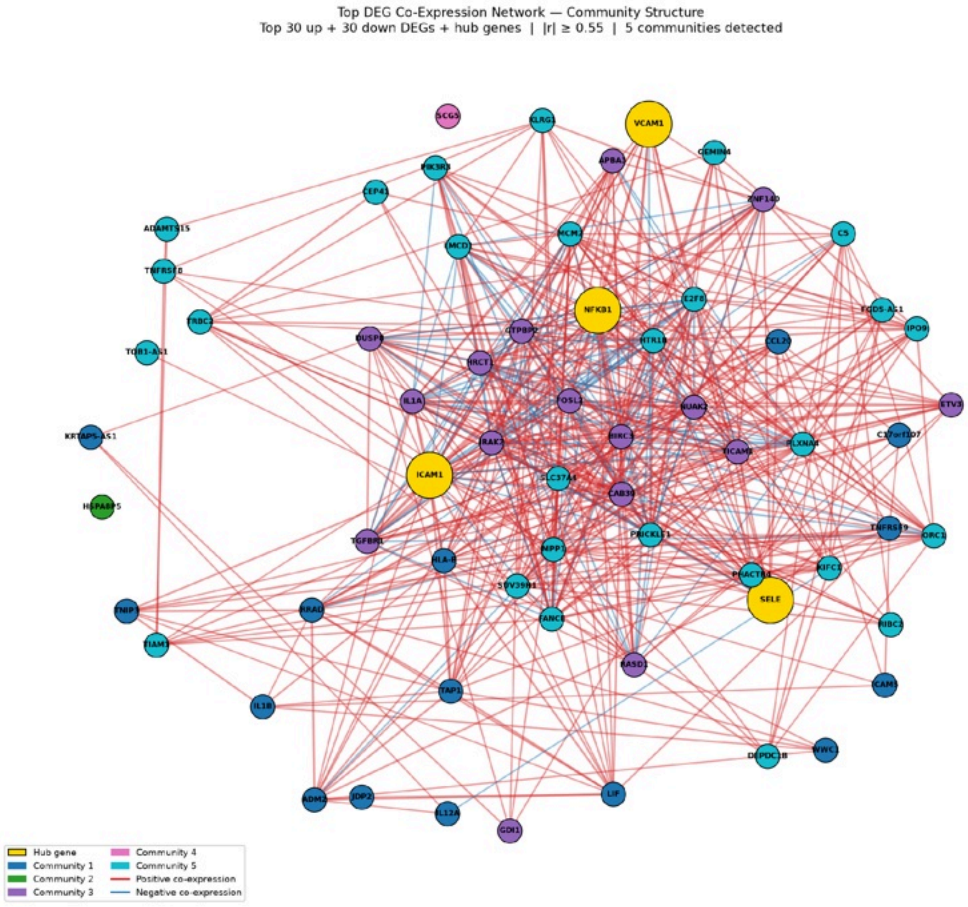


Figure 11. Top-DEG co-expression network with Louvain community structure. Nodes represent the top 30 upregulated and 30 downregulated DEGs as well as the hub genes, which are in gold. Colors indicate a different community, where five communities are present. The width of the edges is proportional to |r|, and only edges that have |r| ≥ 0.55 are visible. The network consists of 64 nodes and 495 edges, and the hub genes link together many community boundaries, similar to convergent transcriptional regulators.

4. DISCUSSION

This multi-dataset transcriptomic meta-analysis provided systematic evidence for the shared NF-κB-centred transcriptional program that human endothelial cells use in response to oxidative stress and inflammatory stimulation. The core findings were the significant increase in the expression of NFKB1, ICAM1, and SELE, a strong enrichment of TNF and NF-κB gene sets, and a suppression of the cell-cycle and pathways that repair DNA. These core findings were consistent throughout the analytical layers (DGE, ORA, GSEA, and network analysis) of this study. The findings also align with established endothelial biology [1,6].

The convergence of responses to H_2O_2, TNF-α, or LPS on a central pathway, NF-κB, is a known process that can be explained biologically. With H_2O_2, oxidative stress leads to NADPH oxidase-generated ROS. This ROS can oxidise cystine residues on IKKβ, which triggers kinase activity, leading to the phosphorylation and degradation of IκBα [4]. This is in line with TRAF2/RIP1-mediated activation of IKKβ through TNF-α [6]. These biological processes show that IKKβ is the point where both reactive oxygen species and inflammatory pathways converge. This shared translocation of NF-κB leads to the same process downstream, activating the same targets(VCAM1, ICAM1, SELE, CCL2). Due to these similar results, the cross-stimulus meta-analysis recovers a logical and biologically accurate expression activity.

The upregulation of negative feedback regulators, including NFKBIA, NFKBIB, TNFAIP3/A20, with NF-κB subunits and effectors, is a well-known characteristic of oscillatory NF-κB dynamics, having been examined at a cellular level and confirmed via endothelial cell models [6,7]. This process, which limits itself, prevents excessive inflammation while still permitting meaningful increases in adhesion molecules and chemokines. Its detection in a cross-dataset model shows that these patterns may be a set feature of human endothelial cells instead of an artefact for a specific cell line.

The downregulation of cell-cycle genes, including MCM2, SUV39H1, and DNA-repair gene sets, specifically homologous recombination, the Fanconi anaemia pathway, and mismatch repair, is biologically important. It demonstrates that ROS and cytokine stimulation both divert cellular resources from genome maintenance to the inflammatory secretome. Suppressing DNA repair in oxidative stress, which creates breaks in strands and modifies oxidative bases, creates a loop that may end up in genomic instability in chronically activated endothelium, as noted in the development of plaques in atherosclerosis.

VCAM1's borderline significance (FDR = 0.060) warrants an explanation. The expression of VCAM1 is highly specific to the cytokine that triggers it. TNF-α and IL-1 can strongly increase VCAM1's expression, but H_2O_2 only produces a weak and variable result for it[10,5]. In a model that mixes both stimuli, VCAM1 expression showed high variance, reducing statistical significance. Even so, the positive effect size of 0.630 and the position of VCAM1 in the co-expression network and the NF-κB pathway network strengthen its roles as a hub gene in endothelial dysfunction. To fix its statistical significance, larger and balanced data sets or analyses that are specific to each stimulus can be used.

Network analysis revealed five unique co-expression communities within the top-DEG network. This difference indicates that endothelial response has many different specialized groups working together, not being uniform. This distinct organization has therapeutic implications. An intervention that is targeting one hub gene, for example NFKB1 via IKKβ inhibition, may suppress one community but not have a significant effect on the others. Although hub genes connect different communities in comparison to peripheral pathway members, their inhibition

may disrupt communications between communities, therefore having a stronger anti-inflammatory effect [24].

This study does have some notable limitations. First, there was an imbalance in control and stimulated samples, with 24 control samples and 130 stimulated samples. Second, different stimuli, like H_2O_2, TNF-α, IL-1, and LPS, are pooled together into one stimulated label. This label helps identify responses that are shared by all the stimuli, but it obscures responses that are unique to each stimulus. Some datasets only gave stimulated samples, so the baseline for control was drawn from other studies. Third, most of the datasets use immortalised cell lines, like HUVEC and HMEC-1, and not primary human endothelial cells directly from vascular beds. Lastly, this analysis is correlational, so validation of the hub genes in the study will require experimentation in primary endothelial cultures.

5. CONCLUSIONS

This study uses 9 different human endothelial transcriptomic datasets to show that oxidative stress and inflammatory stimulation lead to similar gene expression programs. These programs primarily consist of genes involved in the NF-κB pathway, leukocyte-adhesion molecules, and suppressing the cell cycle and DNA repair. In both types of stimulus, genes NFKB1, ICAM1, and SELE emerge as convergent hub regulators of endothelial dysfunction. VCAM1 also shows directionality but at borderline significance. The study's findings support therapeutic interventions that target IKKβ, NF-κB, and the adhesion-molecule axis in situations of vascular inflammatory disease. Furthermore, providing a standard pattern for endothelial activation.

Data Availability: All of the GEO datasets integrated in this study are publicly available. Their accession numbers and source links are provided in Table 1. Analysis code, intermediate data files, and figure-generation scripts, processed expression matrices and enrichment results are available upon reasonable request.

REFERENCES

1. Pober JS, Sessa WC. Evolving functions of endothelial cells in inflammation. Nat Rev Immunol. 2007;7(10):803–815. doi:10.1038/nri2171
2. Libby P. Inflammation in atherosclerosis. Nature. 2002;420(6917):868–874. doi:10.1038/nature01323
3. Ross R. Atherosclerosis, an inflammatory disease. N Engl J Med. 1999;340(2):115–126. doi:10.1056/NEJM199901143400207
4. Taniyama Y, Griendling KK. Reactive oxygen species in the vasculature: molecular and cellular mechanisms. Hypertension. 2003;42(6):1075–1081. doi:10.1161/01.HYP.0000100443.09293.4F
5. Collins T, Read MA, Neish AS, Whitley MZ, Thanos D, Maniatis T. Transcriptional regulation of endothelial cell adhesion molecules: NF-κB and cytokine-inducible enhancers. FASEB J. 1995;9(10):899–909. doi:10.1096/fasebj.9.10.7542214
6. Hayden MS, Ghosh S. Shared principles in NF-κB signaling. Cell. 2008;132(3):344–362. doi:10.1016/j.cell.2008.01.020
7. Zhang Q, Lenardo MJ, Baltimore D. 30 years of NF-κB: a blossoming of relevance to human pathobiology. Cell. 2017;168(1–2):37–57. doi:10.1016/j.cell.2016.12.012
8. Rao J, Brown BN, Bhavnani SK, et al. Transcriptomic analysis of human endothelial cells under oxidative stress. J Proteome Res. 2014;13(1):40–48.
9. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 2002;30(1):207–210. doi:10.1093/nar/30.1.207
10. Iiyama K, Hajra L, Iiyama M, et al. Patterns of vascular cell adhesion molecule-1 and intercellular adhesion molecule-1 expression in rabbit and human atherosclerotic lesions and at sites predisposed to lesion formation. Circ Res. 1999;85(2):199–207. doi:10.1161/01.res.85.2.199
11. Granger DN, Senchenkova E. Inflammation and the Microcirculation. San Rafael (CA): Morgan & Claypool Life Sciences; 2010. Chapter 4: Leukocyte–endothelial cell adhesion.
12. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine Learning in Python. J Mach Learn Res. 2011;12:2825–2830.
13. The pandas development team. pandas-dev/pandas: Pandas. Zenodo. 2020. doi:10.5281/zenodo.3509134
14. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B. 1995;57(1):289–300. doi:10.1111/j.2517-6161.1995.tb02031.x

15. Fang Z, Liu X, Peltz G. GSEApy: a comprehensive package for performing gene set enrichment analysis in Python. Bioinformatics. 2023;39(1):btac757. doi:10.1093/bioinformatics/btac757
16. Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 2016;44(W1):W90–W97. doi:10.1093/nar/gkw377
17. Chen EY, Tan CM, Kou Y, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics. 2013;14:128. doi:10.1186/1471-2105-14-128
18. Ashburner M, Ball CA, Blake JA, et al. Gene Ontology: tool for the unification of biology. Nat Genet. 2000;25(1):25–29. doi:10.1038/75556
19. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 2000;28(1):27–30. doi:10.1093/nar/28.1.27
20. Jassal B, Matthews L, Viteri G, et al. The Reactome Pathway Knowledgebase. Nucleic Acids Res. 2020;48(D1):D498–D503. doi:10.1093/nar/gkz1031
21. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005;102(43):15545–15550. doi:10.1073/pnas.0506580102
22. Hagberg AA, Schult DA, Swart PJ. Exploring network structure, dynamics, and function using NetworkX. In: Proceedings of the 7th Python in Science Conference (SciPy 2008). 2008:11–15.
23. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. J Stat Mech. 2008;2008(10):P10008. doi:10.1088/1742-5468/2008/10/P10008
24. Blake GJ, Ridker PM. Inflammatory bio-markers and cardiovascular risk prediction. J Intern Med. 2002;252(4):283–294. doi:10.1046/j.1365-2796.2002.01019.x