

# **Analisis Differentially Expressed Genes (DEGs) pada Kanker Paru**

## **1. Pendahuluan**

Berdasarkan data GLOBOCAN 2022, kanker paru merupakan kanker dengan tingkat kematian tertinggi di dunia dengan 1,8 juta kematian per tahun dan salah satu kanker dengan jumlah diagnosis tertinggi, yaitu 2,48 juta kasus baru per tahun. Tingginya angka kematian dan kasus baru tersebut berkaitan dengan salah satu faktor risiko utama, yaitu merokok. Menurut WHO (2023), merokok tembakau merupakan faktor risiko utama kanker paru, yang mana komponen pada rokok tembakau mengandung zat karsinogenik dan dapat menyebabkan kerusakan DNA. Akan tetapi, perubahan molekuler pada tingkat ekspresi gen belum sepenuhnya terjadi. Oleh karena itu, dilakukan analisis Differentially Expressed Genes (DEGs) antara jaringan adenokarsinoma paru dan jaringan paru normal untuk mencari kandidat biomarker.

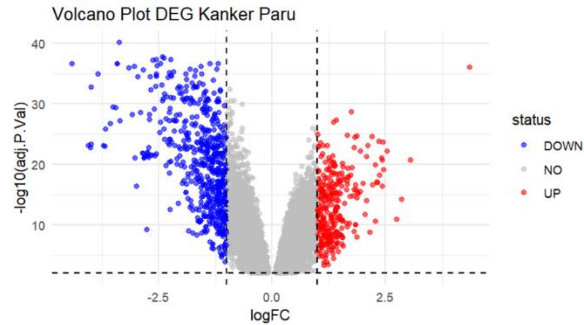
## **2. Metode**

Data ekspresi gen diperoleh dari dataset GSE10072 (Lung Adenocarcinoma vs jaringan paru normal) yang diunduh dari Gene Expression Omnibus (GEO) menggunakan paket GEOquery pada perangkat lunak R. Dataset ini menggunakan platform microarray Affymetrix Human Genome U133A (GPL96). Matriks ekspresi diekstraksi dan dilakukan evaluasi distribusi nilai menggunakan quantile dan diterapkan transformasi log2 untuk menstabilkan varians data. Quality control dilakukan melalui visualisasi boxplot, density plot, dan UMAP untuk mengevaluasi distribusi global dan pemisahan sampel antar kelompok.

Analisis Differentially Expressed Genes (DEGs) dilakukan menggunakan paket limma melalui pendekatan linear modeling dan empirical Bayes dengan penerapan koreksi False Discovery Rate (FDR) dan kriteria signifikansi adjusted p-value  $< 0,01$  serta  $|\log_2 \text{fold change}| > 1$ . Probe signifikan kemudian dianotasi menjadi Gene Symbol dan Gene Name menggunakan database Bioconductor hgu133a.db. Hasil analisis divisualisasikan menggunakan volcano plot dan heatmap untuk 50 gen paling signifikan, serta disimpan dalam format CSV untuk analisis lanjutan dan interpretasi biologis dalam konteks transcriptomics kanker paru.

## **3. Hasil dan Interpretasi**

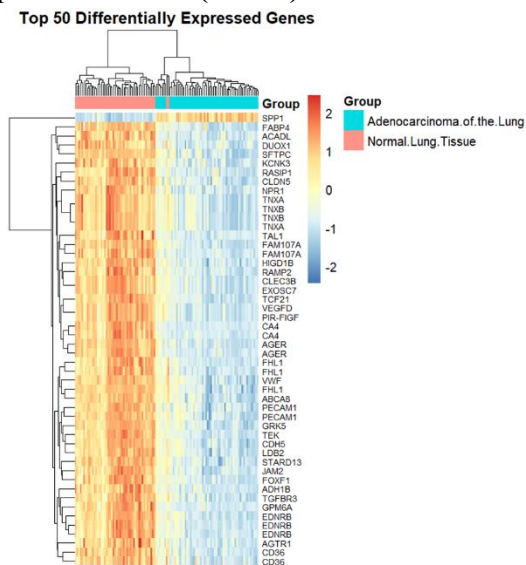
### **3.1 Up-regulation dan Down-regulation**



Gambar 1. Volcano Plot DEG Kanker Paru

Analisis DEGs pada dataset GSE10072 mengidentifikasi sebanyak 920 gen yang mengalami perubahan signifikan (DEGs) berdasarkan kriteria adjusted p-value  $< 0,01$  dan  $|\logFC| > 1$  dan  $|\logFC| < 1$ . Dari jumlah tersebut, 336 gen mengalami upregulation sedangkan 584 gen mengalami downregulation.

### 3.2 50 Differentially Expressed Genes (DEGs) teratas

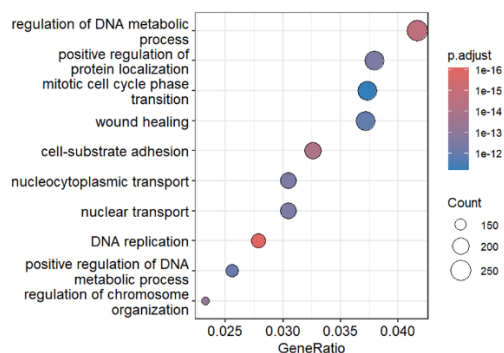


Gambar 2. Heatmap Top 50 Differentially Expressed Genes

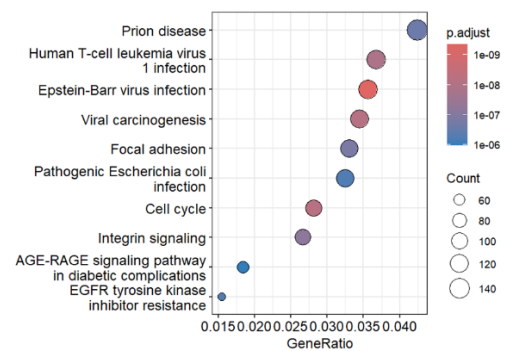
Heatmap yang menunjukkan 50 gen dengan adjusted p-value terendah menunjukkan pola ekspresi yang berbeda antara jaringan adenokarsinoma paru dan jaringan paru normal. Secara umum, sebagian besar gen pada jaringan kanker menunjukkan penurunan ekspresi dibandingkan jaringan normal. Salah satu gen yang menonjol adalah SPP1, yang terekspresi sangat tinggi pada jaringan adenokarsinoma paru (ditandai oleh warna merah) dibandingkan jaringan paru normal. Sementara itu, gen-gen lain seperti FABP4, ACADL, dan DUOX1, dll menunjukkan ekspresi yang tinggi pada jaringan paru normal namun mengalami penurunan yang tajam pada jaringan kanker (ditandai oleh warna biru).

### 3.3 Analisis enrichment Gene Ontology (GO) dan KEGG Pathway

Analisis enrichment dilakukan untuk menginterpretasikan fungsi biologis dan jalur molekuler yang terkait dengan gen-gen diferensial menggunakan web Enrichr. Hasil Gene Ontology (GO) menunjukkan bahwa gen-gen tersebut secara signifikan diperkaya dalam beberapa proses biologis penting, termasuk regulasi metabolisme DNA, transisi fase siklus sel mitosis, replikasi DNA, adhesi sel-substrat, serta transport nukleus dan nukleositoplasma. Selain itu, analisis KEGG menunjukkan bahwa gen-gen diferensial terhubung dengan sejumlah jalur biologis seperti focal adhesi, cell cycle, integrin signaling, serta jalur yang berhubungan dengan carcinogenesis virus (misalnya viral carcinogenesis, Human T-cell leukemia virus 1 infection, Epstein-Barr virus infection).



Gambar 3. Enrichment Gene Ontology (GO) DEGs adenokarsinoma paru dan jaringan paru normal



Gambar 4. Enrichment KEGG pathway DEGs adenokarsinoma paru dan jaringan paru normal

## 4. Kesimpulan

Analisis ekspresi gen pada dataset GSE10072 berhasil mengidentifikasi sejumlah gen yang berbeda secara signifikan antara jaringan adenokarsinoma paru dan jaringan paru normal dengan total 920 DEGs yang terdiri dari 336 gen up-regulated dan 584 gen down-regulated. Visualisasi melalui heatmap 50 gen paling signifikan memperlihatkan perbedaan pola ekspresi yang jelas antara kedua kelompok tersebut. Analisis enrichment menunjukkan bahwa gen-gen diferensial berperan dalam proses biologis penting seperti regulasi metabolisme DNA, transisi fase siklus sel, dan adhesi sel, serta terhubung dengan jalur molekuler seperti focal adhesi, cell cycle, dan jalur carcinogenesis. Secara keseluruhan, hasil ini mengindikasikan bahwa perubahan regulasi ekspresi gen yang signifikan berkontribusi pada mekanisme molekuler yang mendasari perkembangan adenokarsinoma paru serta menyediakan kandidat gen dan jalur potensial untuk penelitian lebih lanjut dalam pemahaman patogenesis dan pengembangan biomarker kanker paru.