

# Case Analysis Report: Differential Gene Expression and Enrichment Analysis of GSE10072 (Lung Adenocarcinoma vs Normal Lung Tissue)

## 1. Pendahuluan

Kanker paru merupakan penyebab utama kematian akibat kanker secara global, dengan lung adenocarcinoma sebagai subtype paling umum dari non-small cell lung cancer (NSCLC). Perkembangan kanker melibatkan perubahan molekuler kompleks yang mencakup disregulasi proliferasi, angiogenesis, resistensi terhadap apoptosis, serta remodeling mikroenvironment tumor (Hanahan & Weinberg, 2011). Oleh karena itu, pendekatan transcriptomics menjadi penting untuk mengidentifikasi gen yang mengalami perubahan ekspresi signifikan antara jaringan kanker dan jaringan normal.

Dataset GSE10072 yang tersedia pada Gene Expression Omnibus (GEO) menyediakan profil ekspresi gen dari jaringan adenokarsinoma paru dan jaringan paru normal. Analisis Differentially Expressed Genes (DEGs) dilakukan menggunakan pendekatan linear modeling melalui paket limma, yang dikenal memiliki kekuatan statistik tinggi untuk analisis microarray maupun RNA-seq (Ritchie et al., 2015). Selanjutnya, interpretasi biologis dilakukan melalui analisis enrichment Gene Ontology (GO) dan KEGG pathway menggunakan clusterProfiler (Yu et al., 2012).

## 2. Metode

### 2.1 Sumber Data

Dataset GSE10072 diunduh dari GEO (NCBI) dan terdiri dari 107 sampel dengan total 22.283 probe berbasis platform Affymetrix Human Genome U133A Array. Data ekspresi diekstraksi menggunakan paket GEOquery.

### 2.2 Analisis Differential Expression

Analisis diferensial dilakukan menggunakan paket limma dengan model desain untuk membandingkan kelompok “Adenocarcinoma of the Lung” dan “Normal Lung Tissue”. Pendekatan empirical Bayes yang digunakan oleh limma meningkatkan stabilitas estimasi varians pada data microarray (Ritchie et al., 2015).

Gen dianggap signifikan jika memenuhi kriteria:

- adjusted p-value (FDR)  $< 0.01$
- $|\log_2 \text{Fold Change}| > 1$

Kriteria ini memastikan bahwa gen yang terpilih memiliki perubahan biologis yang bermakna sekaligus terkontrol terhadap kesalahan multipel.

### 2.3 Visualisasi

Visualisasi dilakukan untuk mengevaluasi distribusi data dan pola pemisahan kelompok:

- Boxplot dan density plot untuk distribusi global ekspresi

- UMAP untuk reduksi dimensi dan pemisahan sampel
- Volcano plot untuk distribusi DEG
- Heatmap 50 gen dengan adjusted p-value terendah

## 2.4 Analisis Enrichment

Gen signifikan dikonversi menjadi Entrez ID menggunakan org.Hs.eg.db dan dianalisis menggunakan:

- Gene Ontology (Biological Process)
- KEGG pathway enrichment

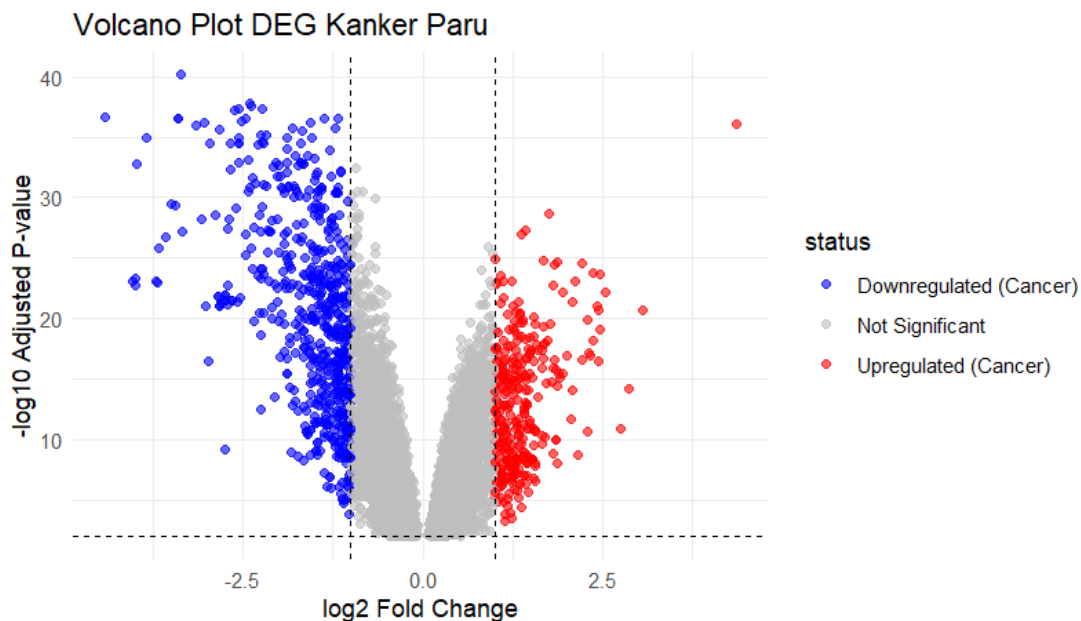
Analisis dilakukan menggunakan clusterProfiler yang memungkinkan interpretasi sistematis dari kumpulan gen diferensial.

## 3. Hasil dan Interpretasi

### 3.1 Differentially Expressed Genes

Ditemukan total 920 gen signifikan, terdiri dari:

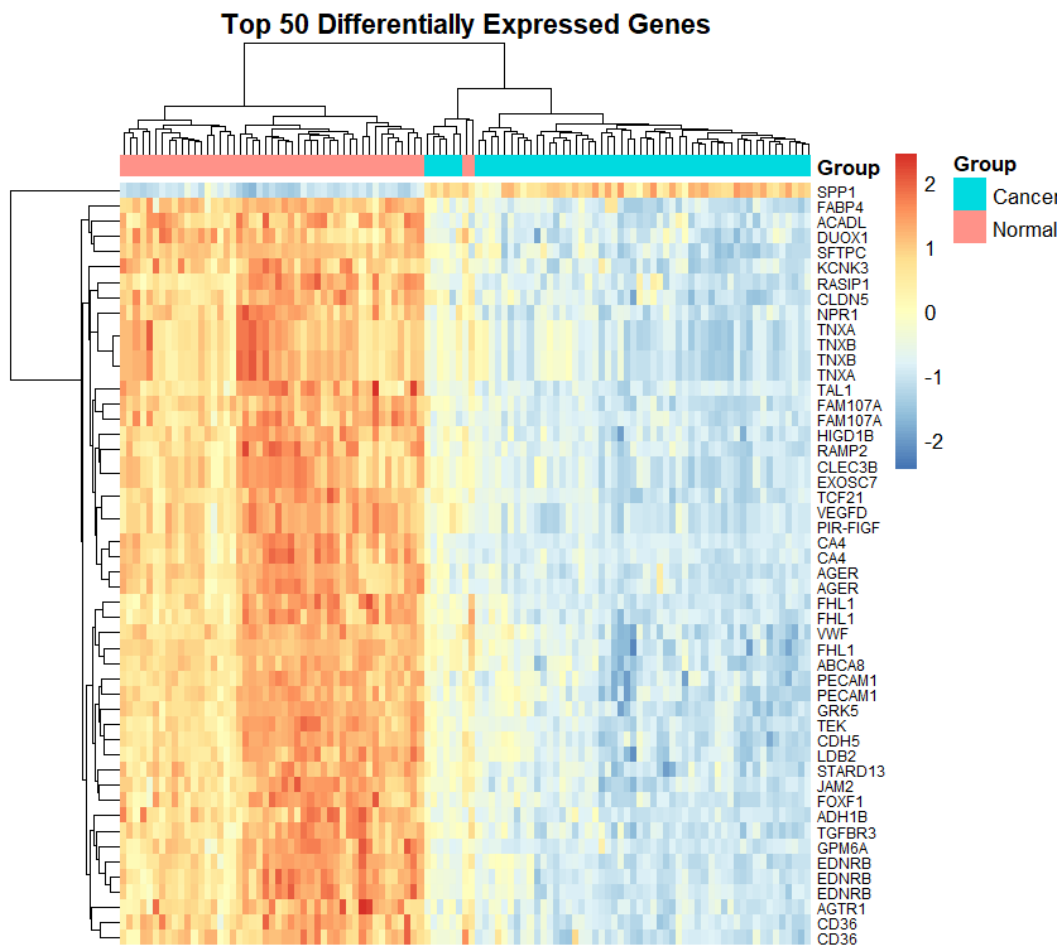
- 336 gen upregulated pada kanker
- 584 gen downregulated pada kanker



Distribusi ini menunjukkan adanya pergeseran ekspresi gen yang substansial antara jaringan kanker dan jaringan normal. Volcano plot memperlihatkan pemisahan yang jelas antara gen upregulated dan downregulated, mengindikasikan perubahan regulasi transkripsional yang luas pada adenokarsinoma paru.

### 3.2 Heatmap Top 50 DEG

Heatmap 50 gen dengan adjusted p-value terendah menunjukkan klusterisasi yang jelas antara sampel kanker dan normal. Pola ini mengindikasikan bahwa perubahan ekspresi gen cukup konsisten untuk membedakan kedua kelompok secara molekuler.

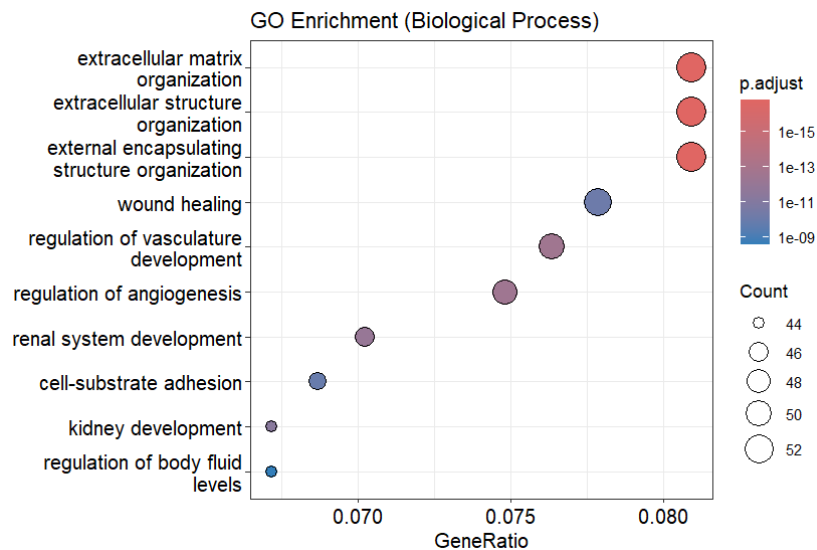


Beberapa gen yang mengalami perubahan signifikan berkaitan dengan proses adhesi sel, interaksi matriks, dan regulasi vaskular, yang dikenal berperan dalam progresi kanker.

### 3.3 Gene Ontology (GO)

Analisis GO menunjukkan enrichment signifikan pada proses biologis berikut:

- Extracellular matrix organization
- Regulation of angiogenesis
- Wound healing



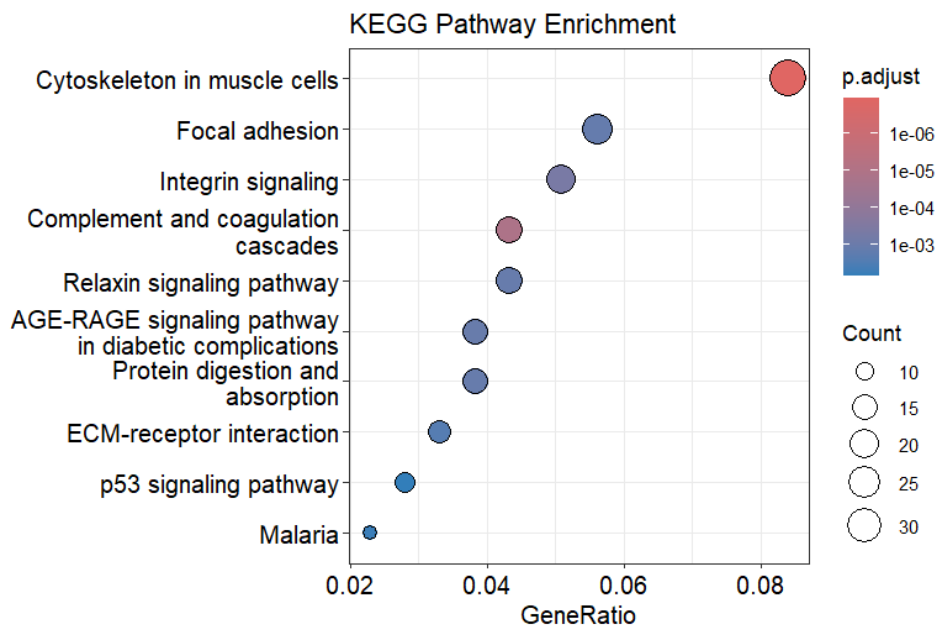
Remodeling matriks ekstraseluler (ECM) merupakan salah satu karakteristik penting progresi tumor. Perubahan komposisi dan struktur ECM memfasilitasi invasi dan metastasis sel kanker melalui modifikasi adhesi dan migrasi sel (Hanahan & Weinberg, 2011).

Angiogenesis juga muncul sebagai proses yang diperkaya secara signifikan. Pembentukan pembuluh darah baru diperlukan untuk mendukung pertumbuhan tumor dan suplai nutrisi, serta merupakan salah satu “hallmarks of cancer” (Hanahan & Weinberg, 2011). Enrichment pada proses wound healing mencerminkan kesamaan molekuler antara respons perbaikan jaringan dan mekanisme invasi tumor.

### 3.4 KEGG Pathway

Pathway yang diperkaya secara signifikan meliputi:

- Focal adhesion
- Integrin signaling
- ECM-receptor interaction
- p53 signaling pathway



Focal adhesion dan integrin signaling berperan dalam interaksi sel dengan matriks ekstraseluler. Aktivasi jalur ini dapat meningkatkan migrasi dan invasi sel kanker melalui reorganisasi sitoskeleton dan aktivasi sinyal pro-survival. Enrichment pada jalur ECM-receptor interaction konsisten dengan hasil GO yang menunjukkan peran sentral remodeling matriks dalam adenokarsinoma paru. Jalur p53 signaling pathway juga teridentifikasi signifikan. p53 merupakan tumor suppressor utama yang mengatur siklus sel dan apoptosis. Disregulasi jalur ini merupakan mekanisme umum pada berbagai jenis kanker, termasuk kanker paru (Hanahan & Weinberg, 2011).

Secara keseluruhan, hasil enrichment menunjukkan bahwa perubahan ekspresi gen pada dataset GSE10072 tidak bersifat acak, melainkan terorganisir dalam jaringan proses biologis yang secara konsisten terkait dengan progresi kanker paru.

## 4. Kesimpulan

Analisis transcriptomics pada dataset GSE10072 mengidentifikasi 920 gen yang mengalami perubahan ekspresi signifikan antara jaringan adenokarsinoma paru dan jaringan normal. Enrichment analysis menunjukkan keterlibatan kuat proses remodeling matriks ekstraseluler, angiogenesis, adhesi sel, serta jalur regulasi siklus sel.

Temuan ini konsisten dengan konsep molekuler kanker modern yang menekankan peran mikroenvironment tumor dan sinyal adhesi dalam perkembangan dan invasi kanker (Hanahan & Weinberg, 2011). Dengan demikian, pendekatan integratif menggunakan limma dan clusterProfiler memberikan gambaran sistemik mengenai perubahan biologis pada adenokarsinoma paru.

## Daftar Pustaka

- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., ... & Sherlock, G. (2000). Gene ontology: Tool for the unification of biology. *Nature Genetics*, 25(1), 25–29. <https://doi.org/10.1038/75556>
- Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell*, 144(5), 646–674. <https://doi.org/10.1016/j.cell.2011.02.013>
- Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, 28(1), 27–30. <https://doi.org/10.1093/nar/28.1.27>
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, 43(7), e47. <https://doi.org/10.1093/nar/gkv007>
- Yu, G., Wang, L.-G., Han, Y., & He, Q.-Y. (2012). clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS: A Journal of Integrative Biology*, 16(5), 284–287. <https://doi.org/10.1089/omi.2011.0118>