

**NAMA : SANDRA PUTRI PERTIWI**

**TUGAS 3 : CASE ANALYSIS**

---

## **Analisis Kanker Paru (GSE10072)**

### **1. PENDAHULUAN**

Kanker didefinisikan sebagai proliferasi sel tubuh yang tidak terkendali dan bersifat ganas. Sebagai salah satu penyebab utama mortalitas secara global, penyakit ini tercatat menyumbang hampir 10 juta kematian pada tahun 2020, atau setara dengan satu dari setiap enam kasus kematian. Adapun jenis kanker dengan prevalensi tertinggi meliputi kanker payudara, paru-paru, kolorektal, serta prostat (Buana & Harahap, 2022). Di antara berbagai jenis kanker tersebut, Adenokarsinoma paru atau kanker paru menempati posisi yang signifikan sebagai tumor epitel primer pada saluran pernapasan, khususnya pada area bronkus. Penyakit ini memiliki karakteristik invasif terhadap struktur jaringan di sekitarnya serta memiliki potensi metastatik tinggi, di mana sel kanker dapat menyebar ke seluruh tubuh melalui sirkulasi darah maupun sistem limfatik (Ainun et al., 2021).

Permasalahan utama yang dihadapi dalam penanganan kanker paru adalah keterlambatan diagnosis akibat kemunculan gejala yang umumnya baru terjadi pada stadium lanjut. Dalam konteks penelitian modern, pendekatan transkriptomik digunakan sebagai instrumen untuk membandingkan profil genetik jaringan sehat dan tumor, yang bertujuan mengungkap mekanisme patofisiologis penyakit pada tingkat molekuler secara komprehensif. Tujuan dari analisis ini adalah untuk menganalisis perbedaan ekspresi gen antara jaringan Adenokarsinoma Paru dan jaringan paru normal untuk mencari kandidat biomarker.

### **2. METODE**

#### **2.1 Sumber Data dan Platform Penelitian**

Data ekspresi gen diperoleh dari basis data publik National Center for Biotechnology Information (NCBI) *Gene Expression Omnibus* (GEO) dengan nomor akses GSE10072. Dataset ini menggunakan platform microarray *Affymetrix Human Genome U133A* (GPL96) (Landi et al., 2008). Proses pengunduhan data dilakukan menggunakan paket GEOquery dalam lingkungan pemrograman R.

## 2.2 Pra-Pemrosesan Data

Sebelum dilakukan analisis statistik, data ekspresi mentah melalui tahap pra-pemrosesan sebagai berikut:

- **Transformasi Logaritma:** Dilakukan transformasi  $\log_2$  untuk menstabilkan varians dan memastikan sebaran data memenuhi asumsi model linear.
- **Normalisasi:** Distribusi nilai ekspresi antar sampel dievaluasi menggunakan *boxplot* dan *density plot* untuk memastikan konsistensi data sebelum dilakukan perbandingan antar kelompok.
- **Anotasi Gen:** Pemetaan ID *probe* Affymetrix ke *Gene Symbol* dan nama gen dilakukan secara otomatis melalui paket anotasi Bioconductor hgu133a.db dan AnnotationDbi.

## 2.3 Analisis Ekspresi Gen Diferensial (DEG)

Identifikasi gen yang terekspresi secara diferensial (DEG) antara jaringan adenokarsinoma paru dan jaringan normal dilakukan menggunakan pendekatan model linear melalui paket limma (*Linear Models for Microarray Data*).

- **Matriks Desain:** Statistik dihitung berdasarkan *design matrix* tanpa *intercept* untuk membandingkan grup kontrol dan perlakuan secara langsung.
- **Stabilitas Varians:** Algoritma *Empirical Bayes* (eBayes) digunakan untuk menstabilkan estimasi varians pada tingkat genetik.
- **Kriteria Signifikansi:** Penentuan DEG didasarkan pada nilai *Adjusted P-value* (menggunakan koreksi *False Discovery Rate / FDR*) dengan ambang batas **< 0,01** dan nilai mutlak *log2 Fold Change (|log2FC|) > 1*.

## 2.4 Visualisasi Data

Untuk menginterpretasikan pola ekspresi secara global dan spesifik, dilakukan visualisasi menggunakan beberapa instrumen:

- **UMAP (*Uniform Manifold Approximation and Projection*):** Digunakan untuk reduksi dimensi non-linear untuk melihat pengelompokan sampel secara global berdasarkan profil ekspresi gen.
- **Volcano Plot:** Digunakan menggunakan paket ggplot2 untuk memetakan distribusi signifikansi statistik terhadap besaran perubahan ekspresi gen.
- **Heatmap:** Menggunakan paket pheatmap untuk menampilkan pola ekspresi dari 50 DEG teratas (*top 50 DEGs*).

## 2.5 Analisis Lanjutan dan Jalur Biologis

Untuk memahami signifikansi biologis dari gen-gen yang teridentifikasi, dilakukan analisis lanjutan sebagai berikut:

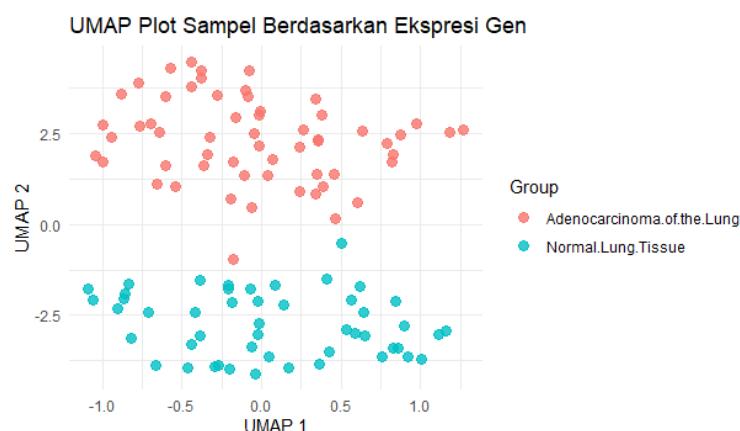
- **Gene Ontology (GO):** Analisis dilakukan menggunakan instrumen **g:Profiler** untuk memetakan DEG ke dalam tiga kategori utama: *Biological Process* (BP), *Molecular Function* (MF), dan *Cellular Component* (CC).
- **Jalur KEGG (Pathway Analysis):** Pemetaan jalur biokimia dilakukan melalui **KEGG Mapper** untuk mengidentifikasi jalur metabolisme atau pensinyalan seluler yang mengalami disregulasi signifikan pada kasus adenokarsinoma paru.

## 3. HASIL DAN INTERPRETASI

### 3.1 Validasi Kualitas Data dan Pengelompokan Sampel

Sebelum melakukan analisis *Differentially Expressed Genes* (DEG), normalisasi data dilakukan menggunakan *Boxplot* dan *Density Plot*. Hasil analisis menunjukkan bahwa distribusi nilai ekspresi gen telah ternormalisasi dengan baik dan konsisten antar sampel.

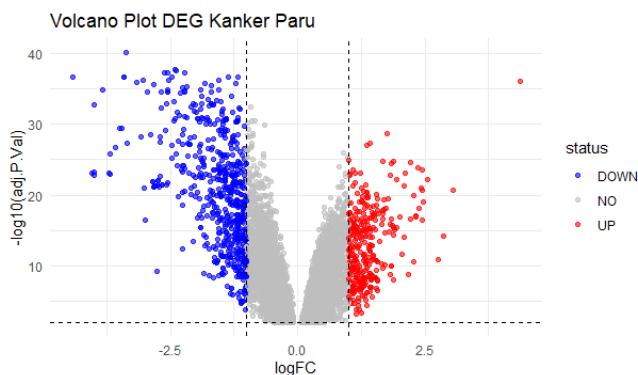
Selanjutnya, analisis reduksi dimensi menggunakan *Uniform Manifold Approximation and Projection* (UMAP) menunjukkan pemisahan kelompok yang sangat signifikan antara kelompok jaringan Normal (biru) dan jaringan Tumor (merah). Pemisahan ini membuktikan bahwa sel adenokarsinoma paru memiliki profil transkriptomik yang sangat berbeda dari jaringan asalnya, sehingga dataset ini divalidasi sangat representatif untuk pencarian kandidat biomarker.



Gambar 1. Visualisasi UMAP Plot

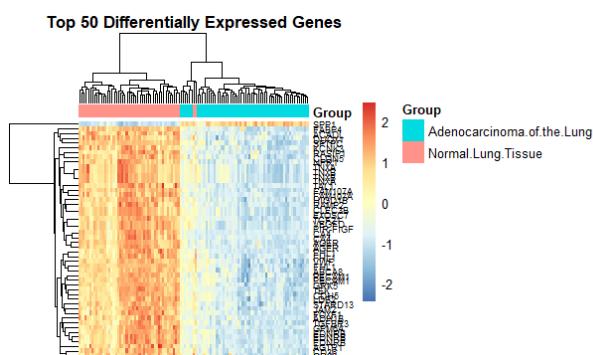
### 3.2 Identifikasi Differentially Expressed Genes (DEGs)

Berdasarkan filter signifikansi statistik ( $p\text{-value} < 0,01$  dan ambang batas ekspresi), hasil analisis menghasilkan ribuan gen yang mengalami disregulasi pada jaringan tumor. Pada *Volcano Plot* (Gambar 2), titik berwarna merah di kuadran kanan merepresentasikan gen dengan peningkatan ekspresi (*up-regulated*), sedangkan titik biru di kuadran kiri menunjukkan gen dengan penurunan ekspresi (*down-regulated*).



Gambar 2. Visualisasi Volcano Plot

Untuk mengevaluasi pola ekspresi gen secara spesifik, *Heatmap* (Gambar 3) dianalisis berdasarkan 50 DEG teratas (*top 50 DEGs*). Visualisasi ini menampilkan profil ekspresi gen yang sangat khas, di mana kelompok sampel Adenokarsinoma (kolom biru muda) menunjukkan pola regulasi yang signifikan dibandingkan sampel Normal (kolom merah muda). Mayoritas dari 50 gen teratas tersebut, seperti *SFTPC*, *AGER*, *FABP4*, dan *PECAM1* mengalami supresi masif atau *down-regulation* (ditandai dengan warna biru) pada jaringan tumor dibandingkan jaringan normal. Hal ini merepresentasikan hilangnya karakteristik fisiologis epitel paru yang sehat akibat dediferensiasi sel kanker. Sebaliknya, gen spesifik seperti *SPP1* terlihat mengalami *up-regulation* yang konsisten (warna merah) pada sampel tumor, menjadikannya kandidat biomarker yang sangat potensial untuk diagnostik atau target terapeutik adenokarsinoma paru.



Gambar 3. Visualisasi Heatmap

### **3.3 Analisis Fungsional (GO & KEGG)**

Analisis lanjutan ini bertujuan untuk menginterpretasikan makna biologis dari daftar gen yang mengalami disregulasi. Dengan memetakan gen-gen tersebut ke dalam *Gene Ontology* (GO) dan *KEGG Pathway*, akan dapat mengidentifikasi proses seluler apa yang terganggu pada adenokarsinoma paru.

#### **A. Gen Up-Regulated**

- Analisis Gene Ontology (GO) pada Gen Up-Regulated**

Berdasarkan hasil analisis menggunakan g:Profiler pada kelompok gen yang naik, ditemukan dominasi aktivitas pada tiga domain utama:

- 1) Biological Process (BP)**

Proses yang paling signifikan secara statistik adalah Developmental Process ( $P_{adj} = 3.429 \times 10^{-45}$ ) dan Cell Division ( $P_{adj} = 1.878 \times 10^{-3}$ ). Hal ini menunjukkan bahwa sel kanker telah memprogram ulang metabolisme seluler untuk mendukung pertumbuhan jaringan yang tidak terkendali.

- 2) Molecular Function (MF)**

Aktivitas Protein Binding ( $P_{adj} = 6.595 \times 10^{-78}$ ) dan DNA-binding transcription activator activity menunjukkan tingginya aktivitas regulasi ekspresi gen di tingkat transkripsi untuk mempertahankan fenotip keganasan sel kanker.

- 3) Cellular Component (CC)**

Lokalisasi gen-gen ini dominan berada di Nucleoplasm ( $P_{adj} = 2.665 \times 10^{-65}$ ), yang konsisten dengan peran mereka dalam replikasi DNA dan pembelahan inti sel.

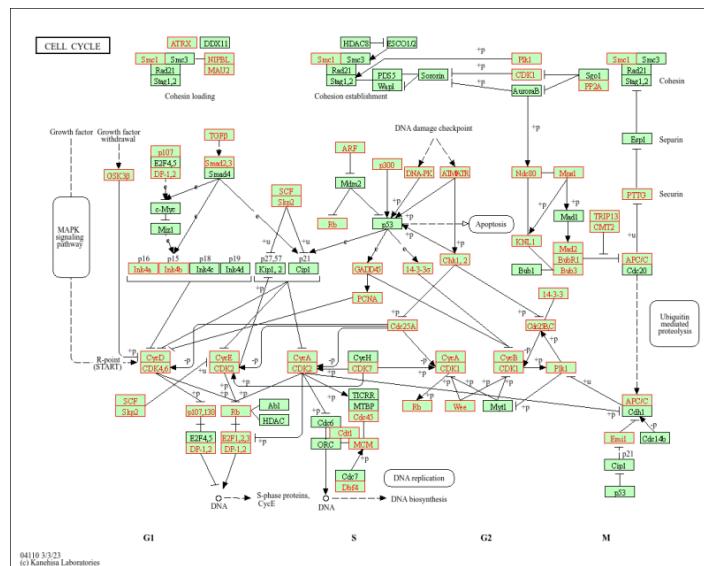
- Analisis Jalur KEGG (KEGG Pathway) - Cell Cycle**

Hasil pemetaan pada jalur Cell Cycle memberikan bukti visual yang sangat kuat mengenai disregulasi siklus sel. Berdasarkan gambar KEGG Mapper, terlihat hampir seluruh komponen kunci penggerak siklus sel mengalami peningkatan ekspresi (ditandai dengan kotak garis tepi merah)

Hal ini sejalan dengan teori yang dikemukakan oleh Cao et al (2023) bahwa tumor dapat dianggap sebagai "penyakit siklus sel" (*cell cycle disease*), di mana penyebab utama tumorigenesis adalah proliferasi sel yang tidak terbatas akibat gangguan pada siklus tersebut. Ketika mekanisme regulasi siklus sel rusak, terjadi pertumbuhan sel yang tidak terkendali yang memicu transformasi sel menjadi ganas.

Dalam data penelitian ini, bukti kerusakan regulasi tersebut terlihat pada beberapa fase kunci:

- 1) Fase G1 ke S: Terjadi peningkatan pada kompleks Cyclin D (CycD), CDK4,6, serta faktor transkripsi E2F1,2,3. Wu et al (2011) mengatakan bahwa CDK4 berperan vital sebagai subunit katalitik yang mengontrol transisi G1-S. Peningkatan ekspresi CDK4 memicu fosforilasi protein RB, yang kemudian membebaskan E2F untuk masuk ke dalam nukleus dan meregulasi ekspresi gen hilir guna memicu pembelahan sel. Kondisi ini menandakan bahwa "gerbang" utama pembelahan sel telah terbuka secara permanen.
  - 2) Replikasi DNA (Fase S): Gen-gen pengatur replikasi seperti PCNA, MCM, dan Cdc45 menunjukkan aktivitas tinggi, yang memfasilitasi penggandaan materi genetik secara cepat.
  - 3) Fase G2 ke M: Peningkatan pada Cyclin B (CycB) dan CDK1 memastikan sel terus melaju ke tahap mitosis tanpa hambatan.
  - 4) Kegagalan Checkpoint: Meskipun gen penahan siklus sel seperti p53 atau ATM/ATR terpetakan, aktivitas penggerak (Cyclin/CDK) yang jauh lebih masif menyebabkan mekanisme perbaikan DNA diabaikan, sehingga sel tumor dapat terus membelah meskipun membawa mutasi.



**Gambar 4.** KEGG Pathway – Cell Cycle

## B. Gen Down-Regulated

- **Analisis Gene Ontology (GO) pada Gen Down-Regulated**

Berbeda signifikan dengan gen yang naik, pengujian fungsional pada gen yang turun (*down-regulated*) menunjukkan hilangnya fungsi-fungsi spesifik jaringan dan mekanisme pertahanan tubuh:

- 1) **Biological Process (BP)**

Proses yang paling signifikan adalah Multicellular Organismal Process ( $P_{adj} = 3.228 \times 10^{-110}$ ) dan Sensory Perception of pain ( $P_{adj} = 8.024 \times 10^{-6}$ ). Hal ini menunjukkan adanya fenomena dediferensiasi, di mana sel adenokarsinoma kehilangan identitas fungsionalnya sebagai sel epitel paru-paru yang sehat. Selain itu, munculnya istilah Extracellular Structure Organization ( $P_{adj} = 1.376 \times 10^{-3}$ ) menandakan kerusakan pada matriks pendukung di sekitar sel.

- 2) **Molecular Function (MF)**

Dominasi pada Protein Binding ( $P_{adj} = 2.378 \times 10^{-84}$ ) menunjukkan gangguan pada interaksi antar protein yang menjaga stabilitas sel.

- 3) **Cellular Component (CC)**

Gen-gen yang turun banyak berlokasi di Cytoplasm ( $P_{adj} = 6.441 \times 10^{-81}$ ) dan Receptor Complex ( $P_{adj} = 1.443 \times 10^{-8}$ ).

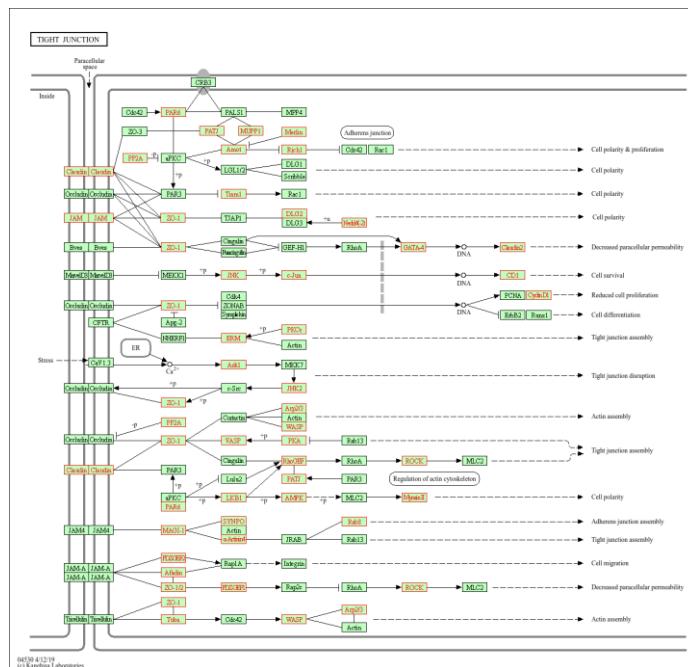
- **Analisis Jalur KEGG (KEGG Pathway) – Tight Junction**

Hasil pemetaan pada jalur Tight Junction memberikan penjelasan molekuler mengenai sel kanker paru yang sangat mudah menyebar (metastasis). Tight junction adalah "perekat" antar sel yang berfungsi menjaga kerapatan jaringan dan mencegah sel lepas dari lingkungan asalnya.

Berdasarkan visualisasi jalur, kita melihat disregulasi masif pada komponen struktural (ditandai dengan kotak garis tepi merah):

- 1) Komponen Perekat Utama: Gen Claudin, Occludin, dan JAM mengalami disregulasi. Menurut Soini (2012), JAM (Junctional Adhesion Molecules) adalah protein dari immunoglobulin superfamily yang berikatan dengan protein seperti ZO-1 untuk mengunci sel agar tetap rapat. Penurunan ekspresi JAMs secara langsung meningkatkan invasi dan potensi metastatik sel kanker. Hal ini merupakan indikasi terjadinya proses Epithelial-Mesenchymal Transition (EMT), di mana sel kehilangan sifat adhesifnya dan menjadi lebih agresif.

- 2) Sitoskeleton dan Adhesi: Disregulasi pada gen Actin, Myosin II, dan kompleks Arp2/3 menunjukkan bahwa sel kanker sedang merombak kerangka internalnya. Ini bukan sekadar kerusakan, melainkan perubahan dari sel yang "diam dan menempel" menjadi sel yang "fleksibel dan mampu bermigrasi".
  - 3) Signaling Jalur Survival: Jalur ini juga terhubung dengan JNK dan c-Jun, yang memicu sinyal ketahanan sel (Cell Survival) meskipun sel tersebut sudah terlepas dari jaringan asalnya.



**Gambar 5.** KEGG Pathway – Tight Junction

### **3.4 Identifikasi Kandidat Biomarker Spesifik**

Berdasarkan integrasi data antara nilai signifikansi statistik ( $P_{adj} < 0,01$ ) dan konsistensi pola ekspresi pada heatmap, penelitian ini berhasil mengidentifikasi beberapa gen kunci yang berpotensi menjadi biomarker untuk diagnostik maupun target terapi adenokarsinoma paru:

- **SPP1 (Osteopontin) sebagai Biomarker Diagnostik:** Gen ini merupakan satu-satunya yang menunjukkan pola *up-regulation* paling konsisten (warna merah menyala) di seluruh sampel tumor dibandingkan sampel normal pada *top 50 DEGs*. Secara biologis, *SPP1* berperan dalam mempromosikan transisi epitel-mesenkim (EMT), yang selaras dengan temuan disregulasi pada jalur *Tight Junction*.
  - **SFTPC dan AGER sebagai Penanda Diferensiasi:** Penurunan drastis (*down-regulation*) pada gen *SFTPC* (Surfactant Protein C) dan *AGER* di jaringan tumor dapat digunakan

sebagai indikator tingkat dediferensiasi sel. Semakin rendah ekspresi gen-gen ini, semakin jauh sel kanker telah menyimpang dari fungsi asli jaringan paru-paru.

## **4. KESIMPULAN**

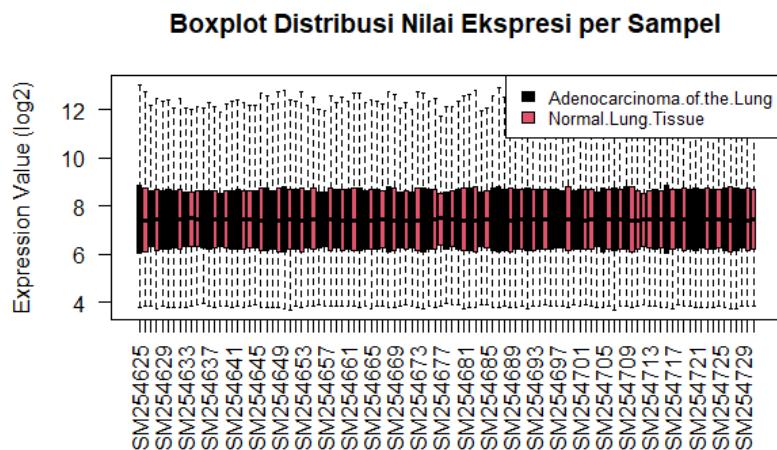
Analisis bioinformatika pada dataset adenokarsinoma paru menunjukkan bahwa proses normalisasi telah berhasil menyetarakan distribusi data secara teknis, sementara plot UMAP mengonfirmasi perbedaan profil transkriptomik yang sangat kontras antara kelompok tumor dan normal. Temuan utama penelitian ini menyoroti disregulasi masif pada jalur Siklus Sel melalui peningkatan ekspresi gen penggerak replikasi, yang memicu pembuatan sel kanker yang tidak terkendali. Secara bersamaan, terjadi penurunan ekspresi pada gen-gen penyusun Tight Junction dan penanda fungsional paru seperti SFTPC dan AGER, yang mengindikasikan hilangnya identitas jaringan sehat serta meningkatnya potensi metastasis. Gen SPP1 berhasil diidentifikasi sebagai kandidat biomarker diagnostik paling potensial karena menunjukkan pola peningkatan ekspresi yang sangat konsisten di seluruh sampel tumor.

## DAFTAR PUSTAKA

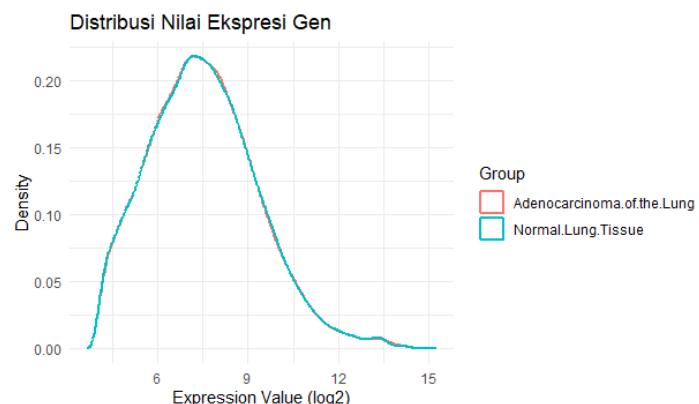
- Ainun, Wiriansya, E. P., Karim, M., Syahril, E., Kanang, I. L. D., & Ahmah, A. I. (2021). Profil Penderita Adenokarsinoma Paru di RS. Ibnu Sina Makassar Periode Januari 2016-Mei 2019. *Jurnal Mahasiswa Kedokteran*, 1(3), 196–202.
- Buana, I., & Harahap, D. A. (2022). Asbestos, Radon, dan Polusi Udara Sebagai Faktor Resiko Kanker Paru pada Perempuan Bukan Perokok. *Jurnal Kedokteran Dan Kesehatan Malikussaleh*, 8(1), 1–16.
- Cao, S., Xiao, S., Zhang, J., & Li, S. (2023). Identification of the cell cycle characteristics of non-small cell lung cancer and its relationship with tumor immune microenvironment, cell death pathways, and metabolic reprogramming. *Frontiers in Endocrinology*, Volume 14-2023. <https://doi.org/10.3389/fendo.2023.1147366>
- Landi, M. T., Dracheva, T., Rotunno, M., Figueroa, J. D., Liu, H., Dasgupta, A., Mann, F. E., Fukuoka, J., Hames, M., Bergen, A. W., Murphy, S. E., Yang, P., Pesatori, A. C., Consonni, D., Bertazzi, P. A., Wacholder, S., Shih, J. H., Caporaso, N. E., & Jen, J. (2008). Gene Expression Signature of Cigarette Smoking and Its Role in Lung Adenocarcinoma Development and Survival. *PLOS ONE*, 3(2), 1–8. <https://doi.org/10.1371/journal.pone.0001651>
- Soini, Y. (2012). Tight junctions in lung cancer and lung metastasis: a review. *International Journal of Clinical and Experimental Pathology*, 5(2), 126–136.
- Wu, A., Wu, B., Guo, J., Luo, W., Wu, D., Yang, H., Zhen, Y., Yu, X., Wang, H., Zhou, Y., Liu, Z., Fang, W., & Yang, Z. (2011). Elevated expression of CDK4 in lung cancer. *Journal of Translational Medicine*, 9(38), 1–9.

## LAMPIRAN

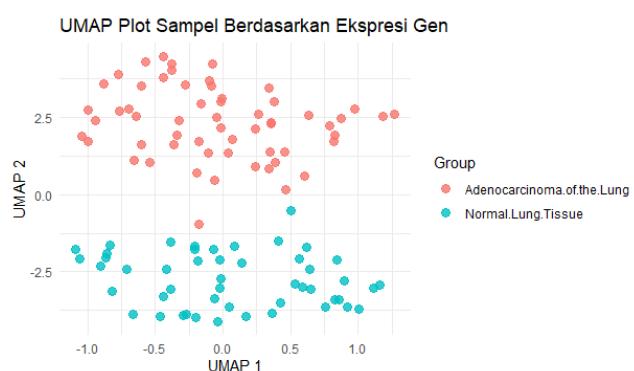
- **Lampiran 1:** Visualisasi BoxPlot



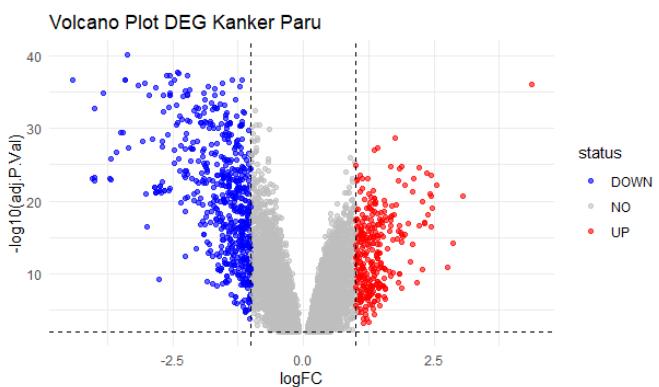
- **Lampiran 2:** Visualisasi Distribusi Nilai Ekspresi Gen



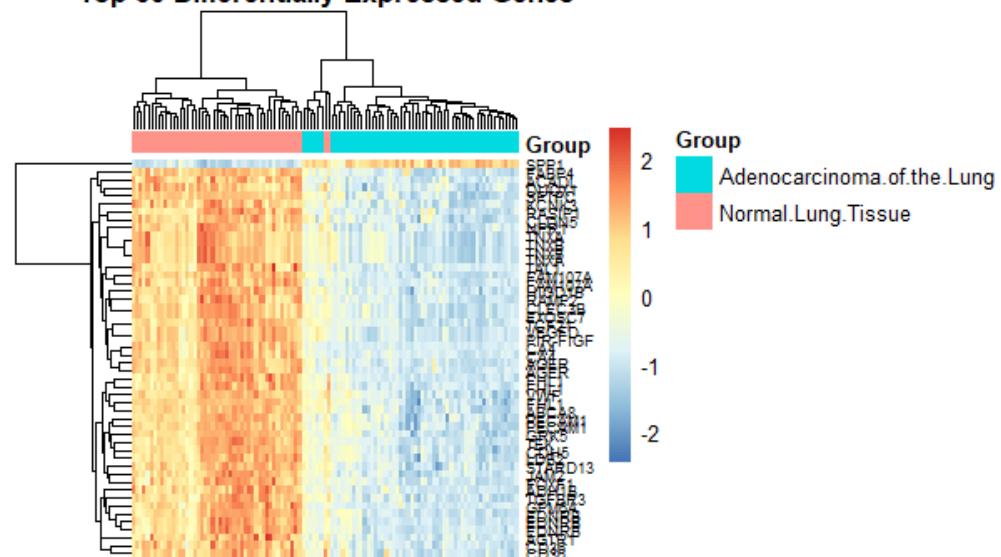
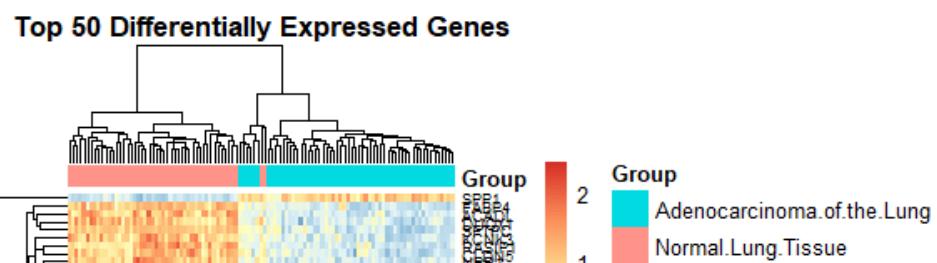
- **Lampiran 3:** Visualisasi UMAP



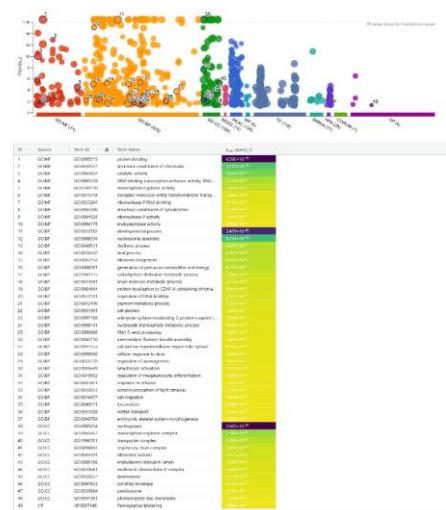
- **Lampiran 4:** Visualisasi Volcano Plot



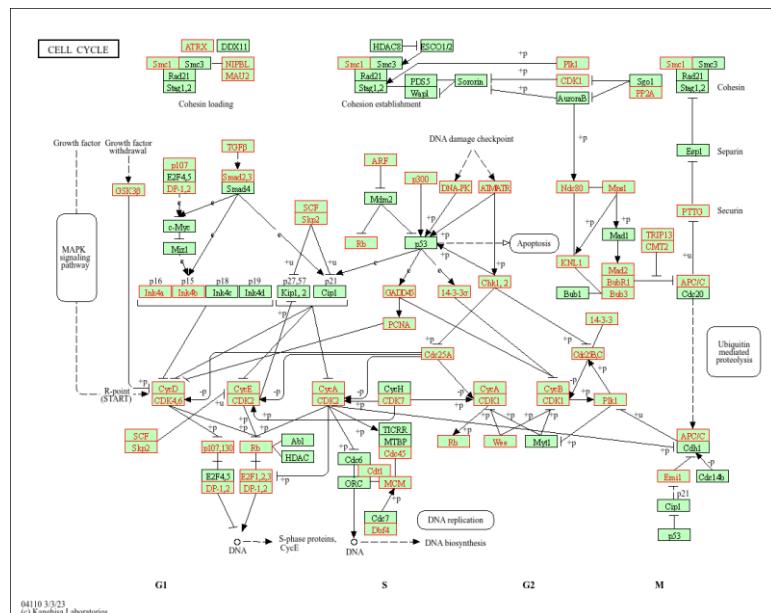
- **Lampiran 5:** Visualisasi Heatmap



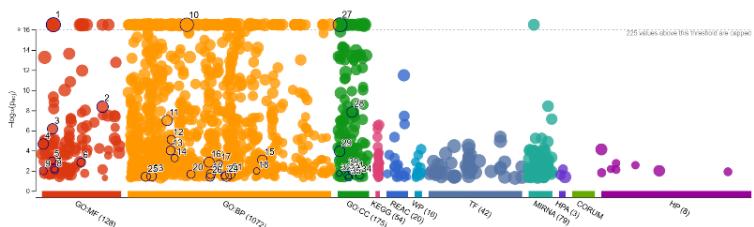
- **Lampiran 6:** Visualisasi analisis Gene Ontology (GO) pada Gen Up-Regulated



- Lampiran 7: Visualisasi pathway KEGG Mapper pada Gen Up-Regulated (Cell Cycle)**



- Lampiran 8: Visualisasi analisis Gene Ontology (GO) pada Gen Down-Regulated**



ID	Source	Term ID	Term Name	P-value (query 1)
1	GO:MF	GO:0005815	protein binding	5.378e-164
2	GO:MF	GO:0005099	molecular catalytic activity	3.465e-161
3	GO:MF	GO:0005111	nonenzymatic catalytic activity	3.16e-160
4	GO:MF	GO:0005176	DNA binding transcription activator activity	3.36e-159
5	GO:MT	GO:0005212	structural constituent of eye lens	1.48e-158
6	GO:MI	GO:0008827	cargo receptor activity	1.68e-153
7	GO:MF	GO:0008827	G protein coupled amine receptor activity	3.65e-153
8	GO:MF	GO:0008823	steroid hydroxylase activity	3.65e-153
9	GO:MF	GO:0008823	transcription coactivator binding	3.877e-153
10	GO:BP	GO:0002501	multicellular organismal process	3.228e-150
11	GO:BP	GO:0016061	viral process	3.117e-149
12	GO:BP	GO:0031926	sensory perception of pain	3.034e-148
13	GO:BP	GO:0031926	pre-mRNA processing involved in metabolic process	3.395e-147
14	GO:BP	GO:0009755	one-carbon compound transport	3.469e-147
15	GO:BP	GO:0007213	methionine biosynthesis metabolic process	3.479e-147
16	GO:BP	GO:0031982	cytoskeletal structure organization	3.521e-147
17	GO:BP	GO:0046058	cAMP metabolic process	3.517e-147
18	GO:BP	GO:0001165	cellular response to UV-B	3.025e-146
19	GO:BP	GO:0035473	pigmentation	2.010e-145
20	GO:BP	GO:0035334	cholesterol efflux	2.173e-145
21	GO:BP	GO:0009852	T cell receptor signaling pathway	2.240e-145
22	GO:BP	GO:0004816	regulation of long-term neuronal synaptic plasticity	2.911e-145
23	GO:BP	GO:0004693	visual perception	3.530e-145
24	GO:BP	GO:0008811	rhythmic process	3.556e-145
25	GO:BP	GO:0009852	protein input into nucleus	3.699e-145
26	GO:BP	GO:0003170	regulation of CD4-positive, alpha-beta T cell differentiation	4.441e-145
27	GO:CC	GO:0009117	cysteine	4.441e-145
28	GO:CC	GO:0004239	receptor complex	4.441e-145
29	GO:CC	GO:0005667	transcription regulator complex	4.550e-145
30	GO:CC	GO:0001838	haptoglobin-hemoglobin complex	4.835e-145
31	GO:CC	GO:0039262	chylomicron	4.852e-145
32	GO:CC	GO:0005387	collagen type IV trimer	4.900e-145
33	GO:CC	GO:0009296	protein lipid complex	5.000e-145
34	GO:CC	GO:0009065	glycoprotein complex	5.066e-145
35	GO:CC	GO:0003481	very low density lipoprotein particle	5.099e-145

- **Lampiran 9:** Visualisasi pathway KEGG Mapper pada Gen Down-Regulated (Tight Junction)

