

Profil DEG GSE10072: Studi transkriptomik *lung adenocarcinoma*

Pendahuluan

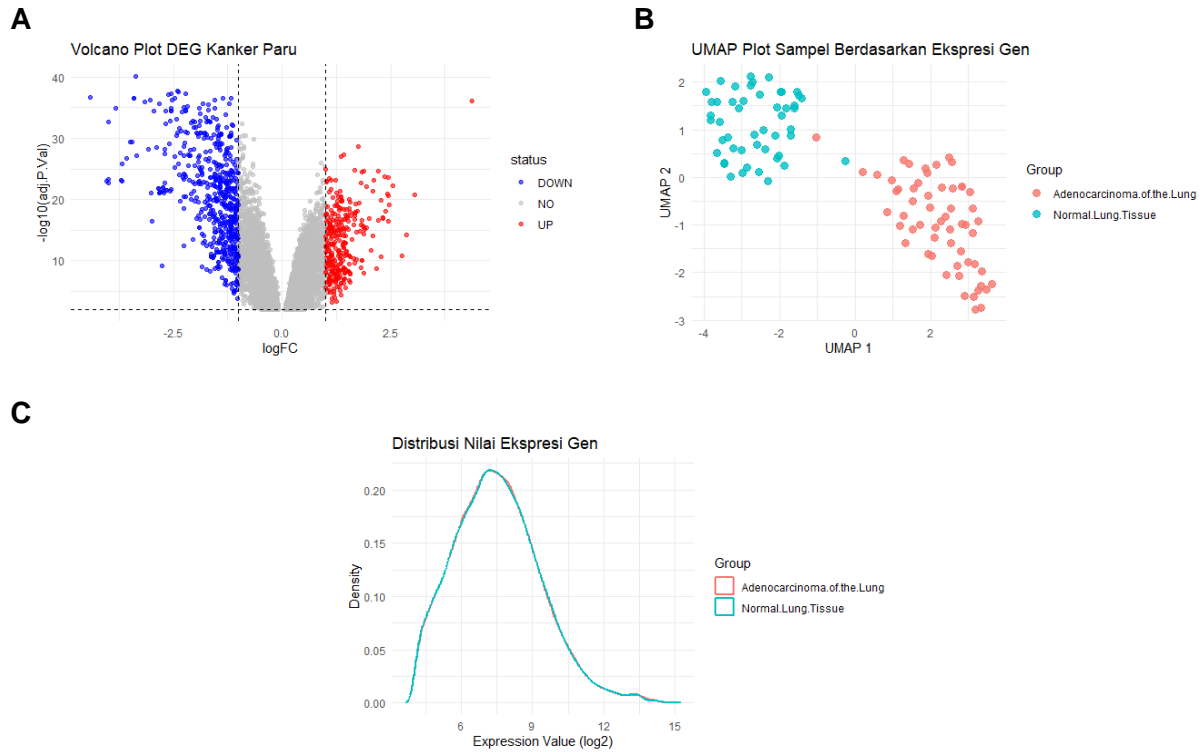
Merokok merupakan faktor risiko utama *lung adenocarcinoma* dan diketahui menyebabkan perubahan molekuler yang memengaruhi regulasi ekspresi gen. Namun, mekanisme transkriptomik yang membedakan jaringan tumor pada perokok dan non-perokok masih belum sepenuhnya dipahami. Pemanfaatan data microarray publik dari *Gene Expression Omnibus* (GEO) memungkinkan eksplorasi perubahan ekspresi gen secara komprehensif dan berbasis data nyata. Analisis ini bertujuan untuk mengidentifikasi *Differentially Expressed Genes* (DEG) yang terkait dengan status merokok pada *lung adenocarcinoma* menggunakan R, serta mengidentifikasi gen sesuai dengan fungsinya menggunakan analisis *gen ontology* (GO) serta yang mengaitkannya dengan jalur metabolisme biokimia menggunakan KEGG pathway. Dengan pendekatan statistik berbasis limma, analisis ini diharapkan dapat mengungkap gen dan proses biologis kunci yang berkaitan dengan siklus sel dan proliferasi, serta menunjukkan penerapan workflow analisis DEG yang reproducible menggunakan data GEO, serta pengelompokan fungsi dan kaitannya dengan jalur metabolisme biokimia.

Metode

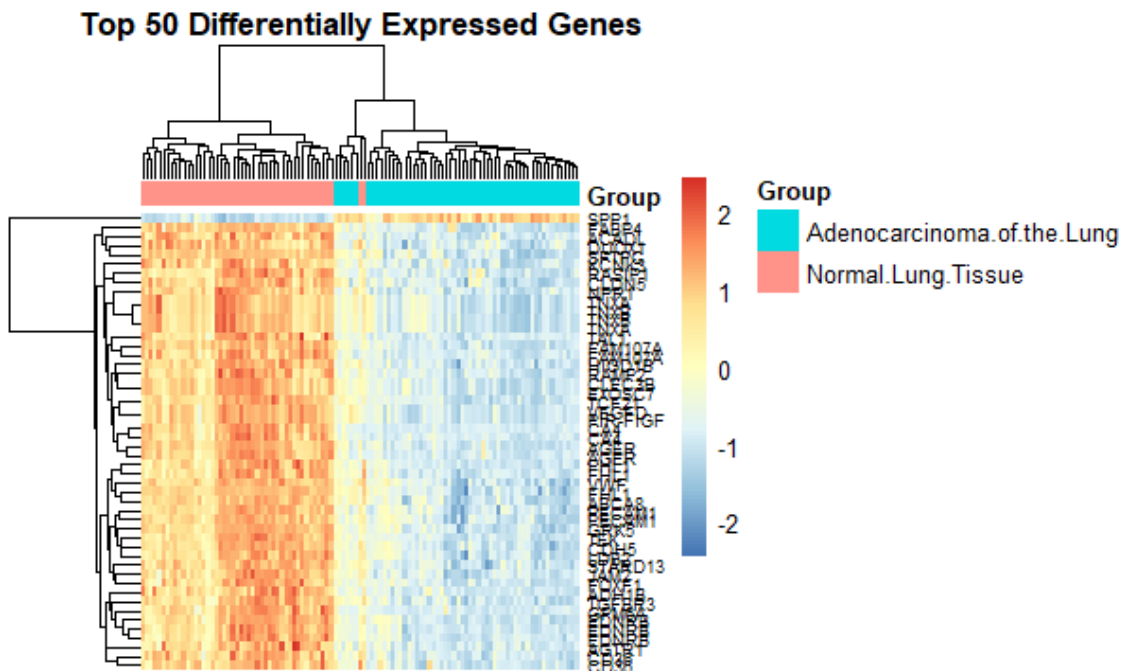
Analisis ekspresi gen ini menggunakan kode akses GEO GSE10072 dengan analisis DEG menggunakan Rstudio yang diukur dengan metode microarray (HG-U113A) (Landi et al., 2008). Terdapat 2 kelompok, yakni kelompok 1 (kontrol/normal) dan kelompok 2 (pengidap *lung adenocarcinoma*). Ekspresi gen dikatakan berbeda secara signifikan apabila memenuhi kedua syarat berikut, yakni $\text{Adj. p-value} < 0.05$ dan $\log_2\text{FC} \geq 1$ untuk gen yang terekspresi secara *up regulated* dan $\log_2\text{FC} \leq -1$ untuk *down regulated*. Setelah analisis DEG, dilanjutkan dengan analisis *gen ontology* (GO) untuk mengetahui pengelompokan gen berdasarkan fungsinya yang ekspresinya berbeda signifikan menggunakan g:Profiler dengan 3 parameter pengelompokan utama (*Molecular function* (MF), *biological process* (BP) dan *cellular component* (CC)) dan dilanjutkan dengan analisis KEGG untuk mengetahui jalur biokimia yang memetakan gen terkait dengan menggunakan tools KEGG color Mapper.

Hasil dan Interpretasi

Berdasarkan hasil dari ekspresi gen menggunakan R dengan kriteria default ($\text{adj. p-value} < 0.05$ dan $\log_2\text{FC}$) didapati terdapat beberapa gen signifikan dengan pola *upregulated* dan *downregulated*. Ekspresi perubahan gen membentuk pola *up* dan *down regulated* yang ditandai dengan dots berwarna biru (down) dan merah (up) pada volcano plot (**Gambar 1. A**). UMAP Plot menunjukkan bahwa kelompok kontrol dan kelompok perlakuan defisiensi nitrogen membentuk 2 cluster berbeda yang tidak bercampur (overlap) yang mengindikasikan bahwa ekspresi gen secara keseluruhan berbeda jelas antara kedua kelompok tersebut (**Gambar 1.B**). Hasil ini memperkuat validitas hasil pada volcano plot yang menunjukkan banyaknya gen yang terekspresi berbeda. Data yang dihasilkan juga menunjukkan adanya distribusi ekspresi yang konsisten dan data sudah ter-normalisasi dengan baik (**Gambar 1.C**). Heatmap dari 50 gen dengan ekspresi diferensial tertinggi menunjukkan pemisahan yang jelas antara perlakuan normal dan pengidap *lung adenocarcinoma* (**Gambar 2**). Pola heatmap yang dihasilkan sesuai dengan UMAP plot, di mana sampel membentuk dua cluster yang terpisah tanpa overlap signifikan. Hal ini menunjukkan bahwa perbedaan ekspresi gen pada kedua kelompok cukup kuat untuk membedakan status biologis sampel dan mencerminkan perubahan molekuler yang khas pada penderita *lung adenocarcinoma*.



Gambar 1. Hasil Analisis DEG GSE10072 menggunakan R dengan kriteria signifikan. Keterangan: A (Volcano plot), B (UMAP Plot), C (distribusi nilai ekspresi gen).

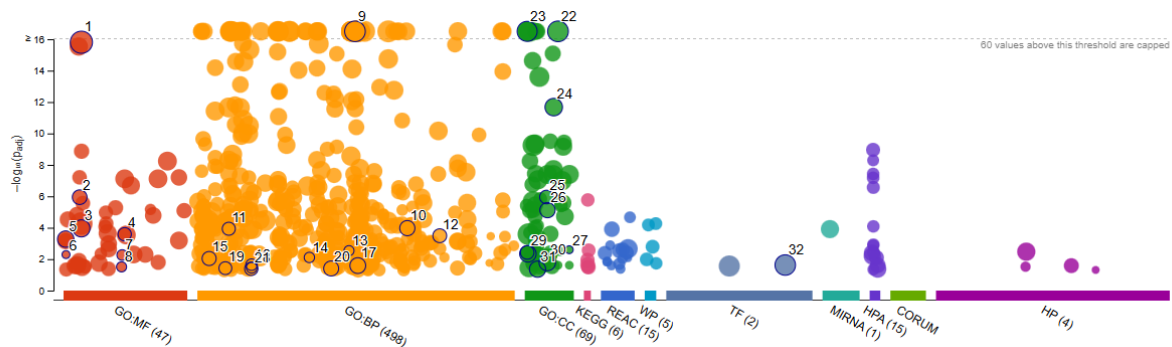


Gambar 2. Hasil Analisis heatmap terhadap 50 gen yang terekspresi berbeda.

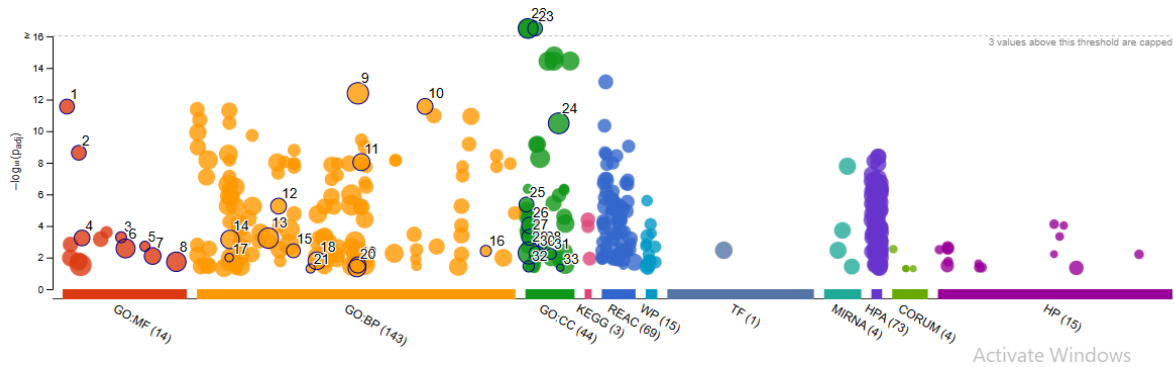
Analisis dengan gen ontology (GO) menunjukkan sebaran fungsi gen berbeda sesuai dengan pola regulasi gen (**Gambar 3.A**). Gen dengan pola regulasi *downregulated* menunjukkan fungsi yang terkait dengan molekuler (MF), yakni yang terkait dengan *protein binding* (signifikansi tinggi), *extracellular matrix structural constituent*, *calcium ion binding*, *cargo receptor activity*, *DNA-binding transcription activator activity*, *adrenomedullin receptor activity*, *nuclear glucocorticoid receptor binding*, dan *type I transforming growth factor beta receptor binding* yang menunjukkan penurunan interaksi protein, sinyal pertumbuhan, dan kontrol transkripsi. Sebanyak 12 fungsi yang terkait dengan biological process (BP) dengan signifikansi tertinggi pada *anatomical structure development*. Fungsi lainnya seperti *ROS metabolic process*, *complement activation*, *cell-substrate junction organization* menunjukkan signifikansi yang relatif tinggi, hasil ini mengindikasikan gangguan pada pemeliharaan struktur jaringan, pertahanan terhadap stres oksidatif, serta adhesi sel dan matriks. Sementara untuk fungsi CC, didapati sebanyak 10 fungsi yang terkait dengan *downregulated*, dengan urutan sangat signifikan ke signifikansi tinggi berturut-turut terkait dengan *cell periphery*, *extracellular region*, *anchoring junction*, *actomyosin* dan *contractile muscle fiber*, menunjukkan melemahnya arsitektur epitel dan stabilitas sitoskeleton. Secara keseluruhan, pola ini mencerminkan penurunan fungsi protektif dan struktural jaringan paru yang dapat menciptakan lingkungan mikro yang lebih rentan terhadap transformasi ganas.

Fungsi berbeda ditunjukkan oleh gen dengan regulasi ekspresi *upregulated* (**Gambar 3.B**), dimana terdapat 8 fungsi yang terkait dengan MF, dengan signifikansi yang tinggi pada fungsi *antigen binding* dan *extracellular matrix structural constituent* yang menunjukkan aktivasi komponen imun humoral dan remodeling matriks. Sebanyak 13 fungsi terkait dengan BP dan 12 fungsi terkait dengan CC. *Response to stimulus*, *nuclear chromosome segregation*, *cell division* dan *extracellular matrix organization* menunjukkan fungsi dengan signifikansi tinggi di kelompok BP, menegaskan peningkatan aktivitas pembelahan sel dan reorganisasi lingkungan tumor. Sedangkan pada fungsi CC dengan signifikansi tinggi, didapati gen yang terkait dengan fungsi *extracellular space*, *immunoglobulin complex* dan *cell periphery*, menunjukkan peningkatan aktivitas sekretori dan interaksi sel dengan mikroenvironment. Secara umum, pola ini menggambarkan aktivasi proliferasi sel dan perubahan lingkungan ekstraseluler yang mendukung perkembangan dan progresi adenokarsinoma paru.

A



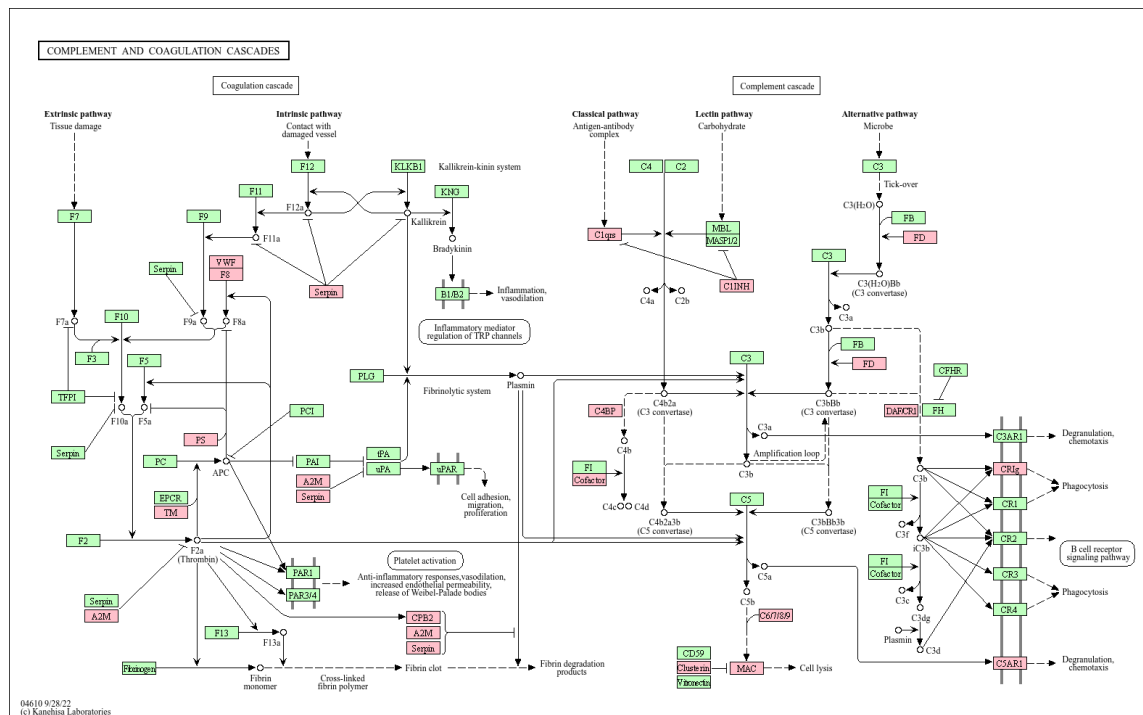
B



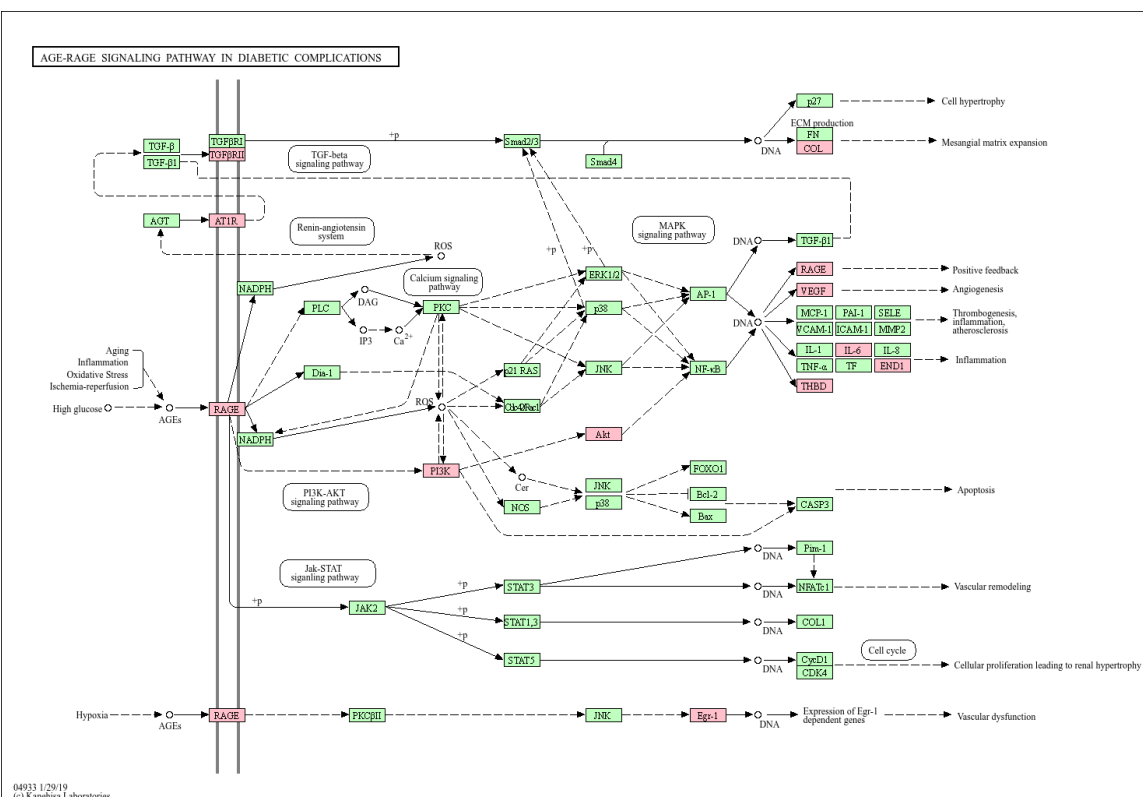
Gambar 3. Analisis Gen Ontology (GO) pada DEG dengan regulasi *downregulated* (A) dan *upregulated* (B)

DEG dengan regulasi *downregulated* menunjukkan keterkaitan dengan signifikansi tinggi terhadap 6 jalur metabolisme, yakni pada jalur metabolisme *complement and coagulation cascades* – 86 genes hit (KEGG:04610), *AGE-RAGE signalling in diabetic complications* – 100 genes hit (KEGG:04933), *vascular smooth muscle contraction* - 133 genes hit (KEGG:04270), *Focal adhesion* - 202 genes hit (KEGG:04510), *relaxin signaling pathway* – 129 genes hit (KEGG:04926), serta *adherens junction* - 93 genes hit (KEGG:04520) (**Gambar 4**). Secara umum, jalur-jalur ini berperan dalam regulasi respon imun, regulasi inflamasi, integritas struktur epitel, interaksi sel dan matriks ekstraseluler, serta homeostasis vaskular. Penurunan ekspresi pada jalur tersebut mengindikasikan terjadinya disfungsi mekanisme protektif jaringan paru, termasuk gangguan adhesi sel dan stabilitas arsitektur epitel, yang berpotensi menciptakan lingkungan mikro yang lebih permisif terhadap transformasi ganas. Hasil analisis ini sejalan dengan pembahasan oleh Landi et al. (2008), yang menunjukkan bahwa paparan rokok memicu perubahan ekspresi gen luas pada jaringan paru, termasuk gangguan pada respon imun dan struktur jaringan bahkan sebelum perkembangan tumor yang nyata, mendukung konsep *field cancerization* pada karsinogenesis paru.

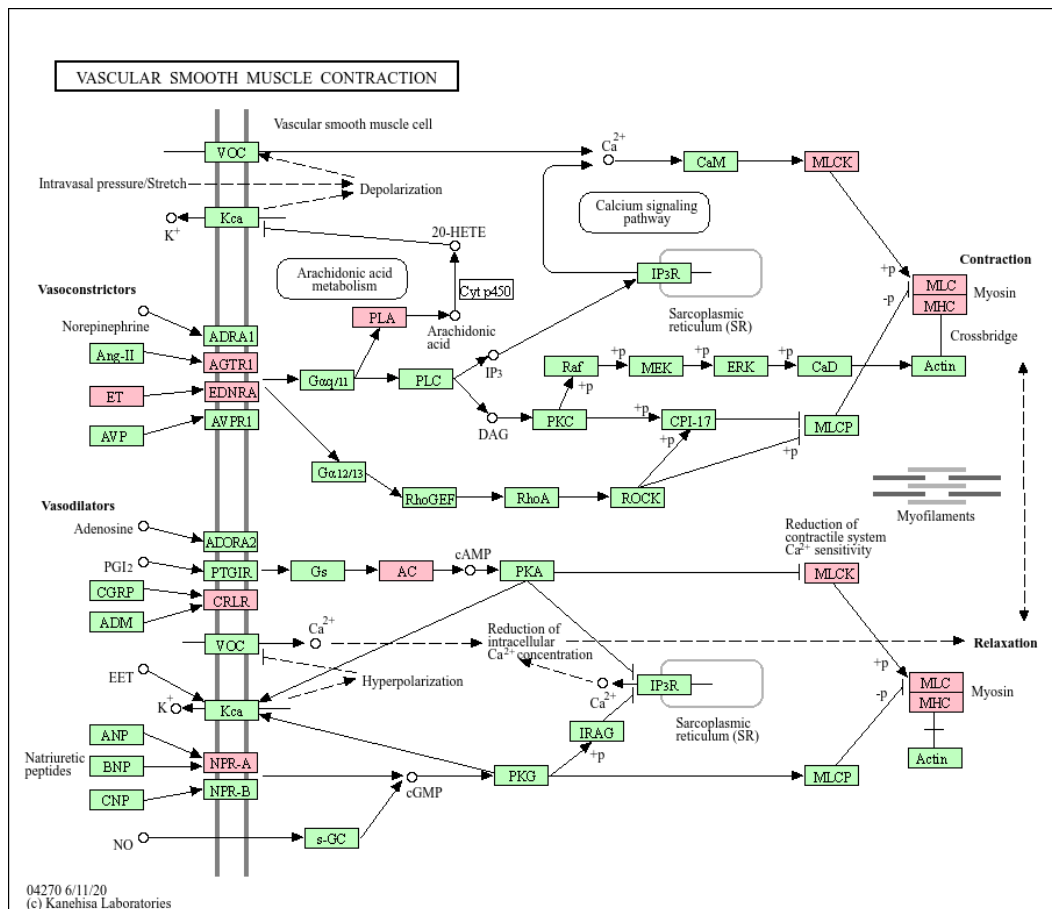
A.



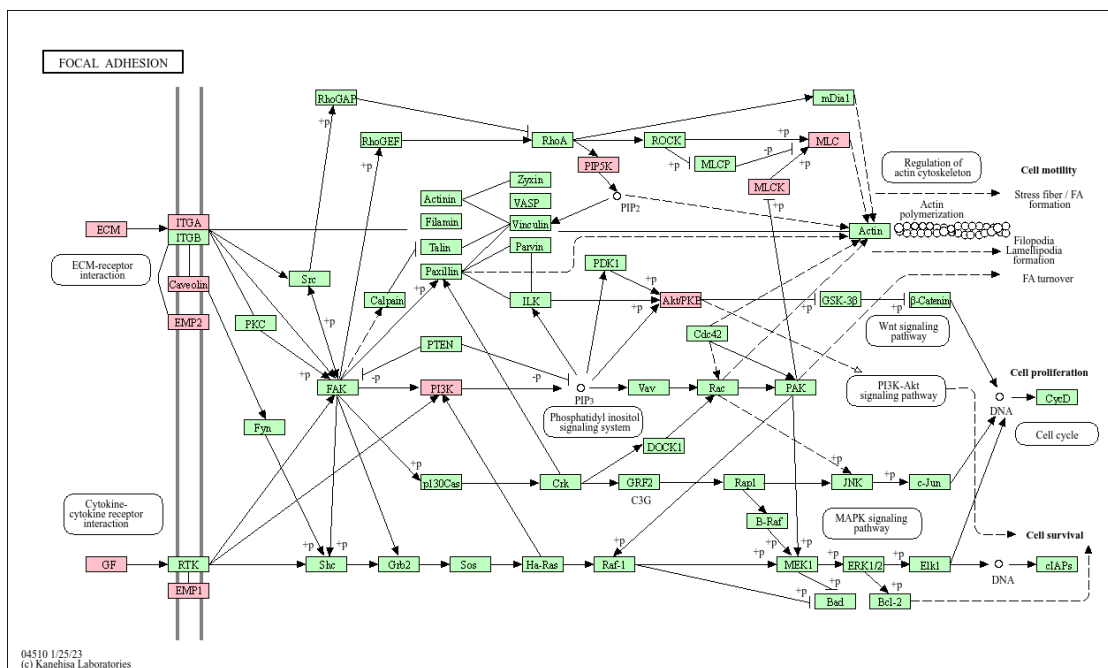
B.



C.



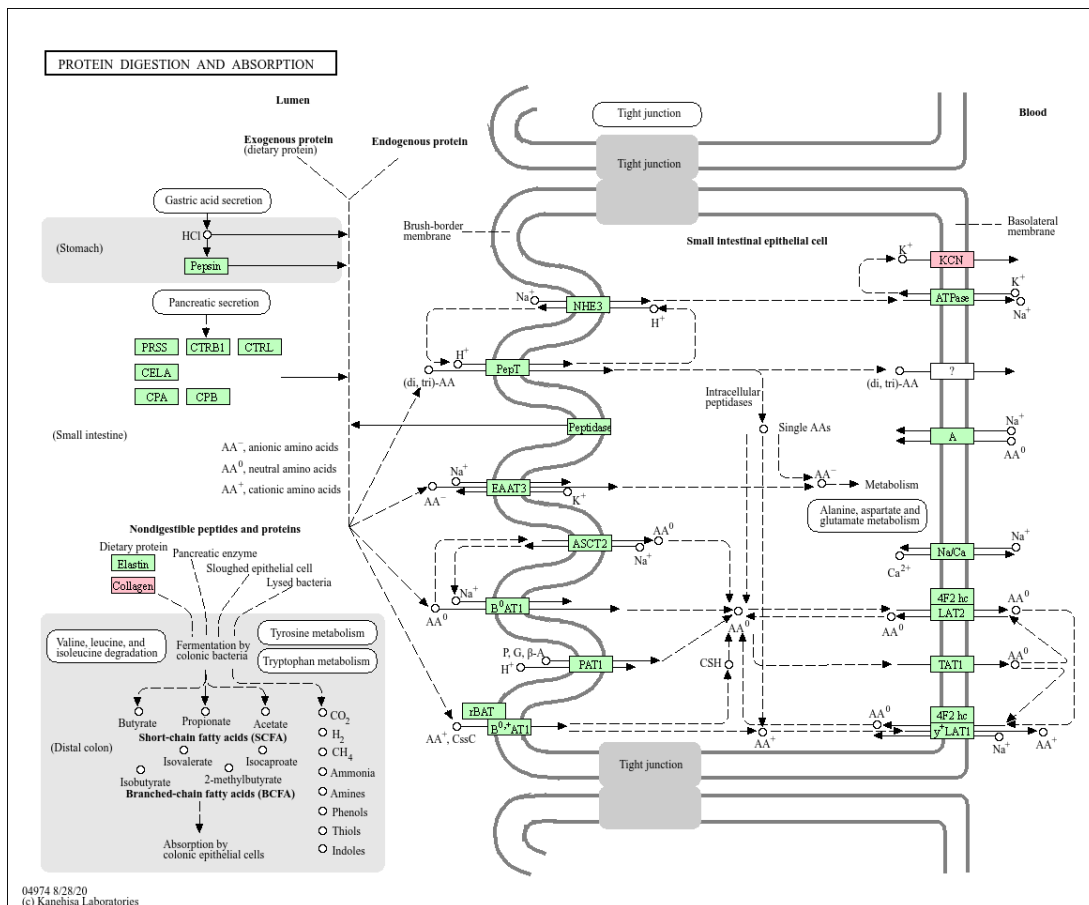
D.



Gambar 4. KEGG pathway dengan keterkaitan tinggi terhadap DEG downregulated. A (*complement and coagulation cascades*) B. (*AGE-RAGE signalling in diabetic complications*), C (*vascular smooth muscle contraction*), D (*Focal adhesion*), E (*relaxin signaling pathway*), F (*adherenes junction*).

Keterkaitan tinggi dengan jalur metabolisme juga didapat pada regulasi *upregulated*, dimana terkait dengan 3 jalur metabolisme, yakni pada *protein digestion and absorption* – 103 genes hit (KEGG:04974), *p53 signalling pathway* – 74 genes hit (KEGG:04115) dan *cell cycle* – 157 genes hit (KEGG:04110) (**Gambar 5.**). Ketiga jalur ini secara biologis berkaitan dengan peningkatan proliferasi sel, respons terhadap stres genotoksik, serta perubahan komposisi matriks ekstraseluler. Aktivasi jalur siklus sel mencerminkan karakteristik utama sel kanker berupa proliferasi tidak terkendali, sementara keterlibatan jalur p53 menunjukkan adanya respons terhadap kerusakan DNA akibat paparan karsinogen rokok, meskipun dalam konteks kanker mekanisme kontrol ini sering kali tidak berjalan efektif. Peningkatan ekspresi gen kolagen dan komponen matriks pada jalur protein digestion and absorption juga mengindikasikan perubahan lingkungan mikro tumor yang mendukung invasi dan progresi. Secara keseluruhan, pola upregulation ini memperkuat temuan Landi et al. (2008) bahwa signature molekuler akibat rokok mencakup aktivasi proliferasi dan perubahan regulasi pertumbuhan sel, yang bersama dengan disfungsi mekanisme protektif jaringan, berkontribusi terhadap perkembangan dan progresi *lung adenocarcinoma*.

A.



Daftar Pustaka

Landi MT, Dracheva T, Rotunno M, Figueroa JD, Liu H, et al. (2008) Gene Expression Signature of Cigarette Smoking and Its Role in Lung Adenocarcinoma Development and Survival. PLOS ONE 3(2): e1651. <https://doi.org/10.1371/journal.pone.0001651>