

Transcriptomic Signature of Early Antiviral Response Following Influenza Vaccination: A Bioinformatics Approach

A. Pendahuluan

Gen merupakan “instruksi” dalam tubuh. Namun, tidak semua instruksi dijalankan setiap saat. Differentially Expressed Genes (DEG) membantu peneliti mengetahui Gen yang mengalami peningkatan atau penurunan ekspresi ketika seseorang berada dalam kondisi tertentu, misalnya saat sakit atau setelah menerima vaksin.

Analisis DEG memerlukan perangkat lunak pendukung seperti limma atau DESeq2 untuk menangani data berskala besar yang tidak memungkinkan dihitung secara manual atau menggunakan Excel, seperti data ekspresi gen yang umumnya terdiri atas puluhan ribu baris dan banyak sampel. Selain itu, penggunaan tools bioinformatika memungkinkan hasil penelitian direplikasi oleh peneliti lain sehingga mendukung prinsip reproducibility.

Dataset yang digunakan adalah GSE48018, yang merupakan bagian dari studi respons sistem kekebalan tubuh terhadap vaksinasi. Dataset ini berfokus pada perubahan ekspresi gen dalam sel darah setelah seseorang menerima vaksin influenza. Analisis ini bertujuan mengidentifikasi Differentially Expressed Genes (DEGs), yaitu gen-gen yang menunjukkan perubahan ekspresi secara signifikan akibat pemberian vaksin. Melalui analisis ini, mekanisme molekuler respons imun tubuh serta potensi penanda genetik (biomarker) yang berkaitan dengan efektivitas vaksinasi dapat dipahami.

B. Metode

Data yang dianalisis difokuskan pada gen yang mengalami peningkatan ekspresi (up-regulated) dan penurunan ekspresi (down-regulated) antara Day 0 (kondisi sebelum vaksinasi) dan Day 1 (24 jam setelah vaksinasi). Perbandingan ini dilakukan untuk mengidentifikasi perbedaan ekspresi gen (DEG) dan memahami respons awal tubuh terhadap vaksin.

Langkah-langkah analisis dalam R Studio dilakukan sebagai berikut: Persiapan library → Pengambilan dataset (GEO) → Penentuan grup (Design & Contrast) → Analisis statistik (limma) → Identifikasi gen signifikan (TopTable & pelabelan) → Analisis fungsi biologis (GO & KEGG) → Visualisasi pengelompokan sampel (UMAP).

1. Persiapan Alat (Instalasi dan Load Library)

Pada tahap awal, dilakukan instalasi BiocManager untuk mengakses paket bioinformatika seperti limma, GEOquery, dan illuminaHumanv4.db, serta penggunaan install.packages untuk paket umum seperti ggplot2 dan dplyr.

2. Pengambilan dan Pemeriksaan Data

Dataset diambil dari database NCBI (GEO) dan dimasukkan ke dalam R untuk dilakukan pemeriksaan awal. Fungsi exprs dan pData digunakan untuk memisahkan data menjadi dua bagian, yaitu matriks ekspresi gen dan metadata sampel (misalnya informasi hari pengambilan sampel). Fungsi quantile digunakan untuk memeriksa rentang nilai ekspresi serta distribusi data, yang umumnya berada dalam skala log2.

3. Desain Eksperimen (Grouping)

Untuk membedakan sampel Day 0 dan Day 1, fungsi sub digunakan untuk mengekstraksi informasi hari dari judul sampel.

Selanjutnya, fungsi model.matrix digunakan untuk membuat tabel design yang memetakan setiap sampel ke dalam grup masing-masing.

Fungsi makeContrasts kemudian digunakan untuk menentukan perbandingan spesifik yang akan diuji, yaitu Day 1 dikurangi Day 0.

4. Analisis Statistik (DEG dengan limma)

Tahap ini merupakan inti analisis. Fungsi lmFit digunakan untuk mencocokkan data ekspresi gen dengan desain eksperimen. Fungsi eBayes diterapkan untuk menstabilkan estimasi statistik sehingga nilai p-value menjadi lebih akurat melalui pendekatan empirical Bayes. Selanjutnya, fungsi topTable digunakan untuk menghasilkan tabel gen yang menunjukkan perbedaan ekspresi paling signifikan.

5. Visualisasi Awal (Volcano Plot dan Heatmap)

Sebelum membuat volcano plot, gen diklasifikasikan sebagai meningkat (UP) jika memiliki logFC positif dan signifikan, serta menurun (DOWN) jika logFC negatif dan signifikan. Volcano plot dibuat menggunakan fungsi ggplot untuk menampilkan sebaran tingkat signifikansi dan besar perubahan ekspresi gen. Heatmap dibuat menggunakan fungsi pheatmap untuk memvisualisasikan pola ekspresi gen-gen teratas sehingga perbedaan pola antara Day 0 dan Day 1 dapat terlihat secara lebih jelas.

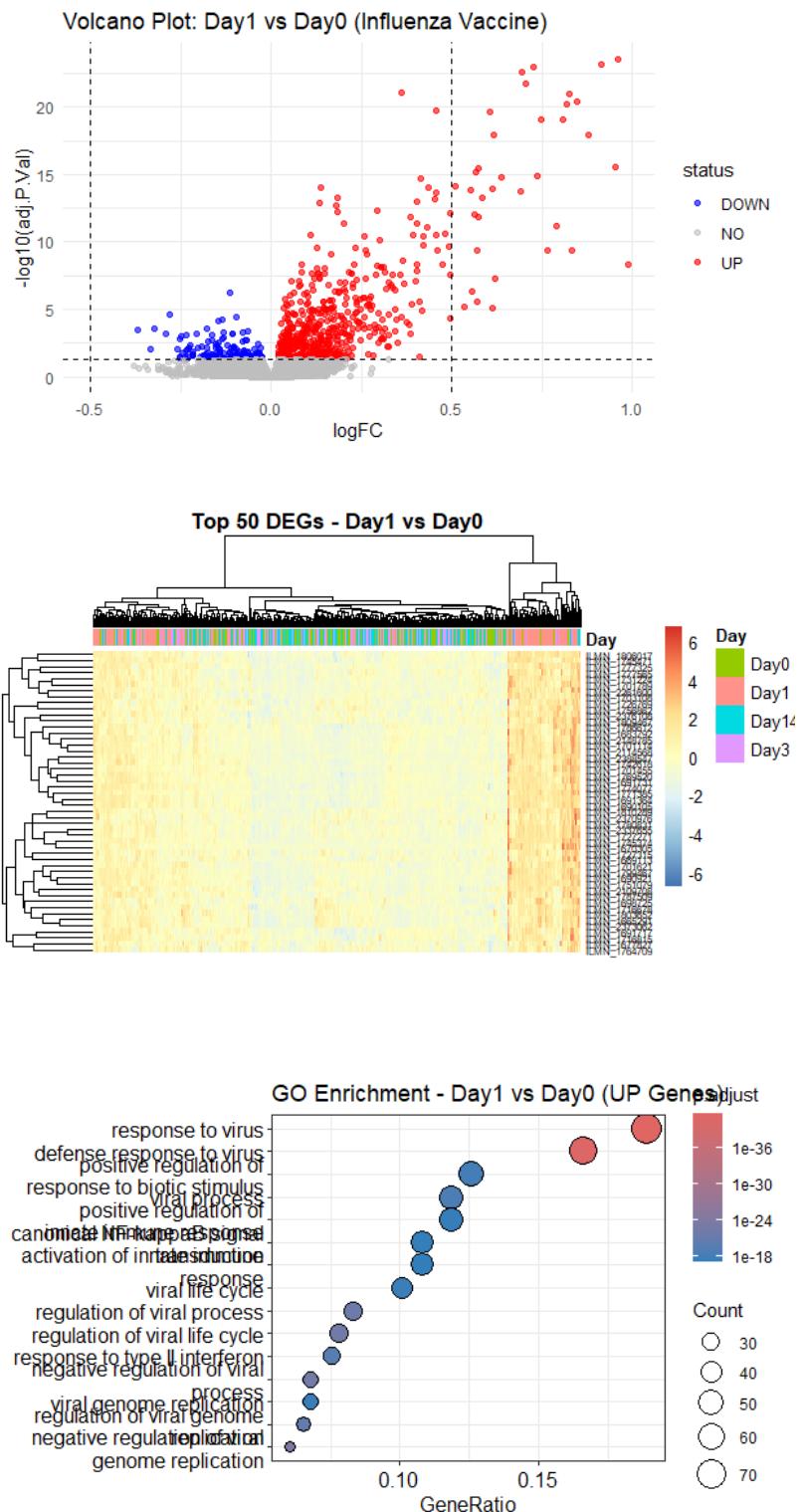
6. Analisis Pengayaan (Enrichment Analysis)

Gen-gen yang signifikan dianalisis lebih lanjut untuk mengetahui fungsi biologisnya. Fungsi bitr digunakan untuk mengonversi ID probe menjadi ID Entrez agar kompatibel dengan basis data biologis. Fungsi enrichGO digunakan untuk mengidentifikasi proses biologis yang dominan, sedangkan enrichKEGG digunakan untuk mengidentifikasi jalur molekuler yang terlibat.

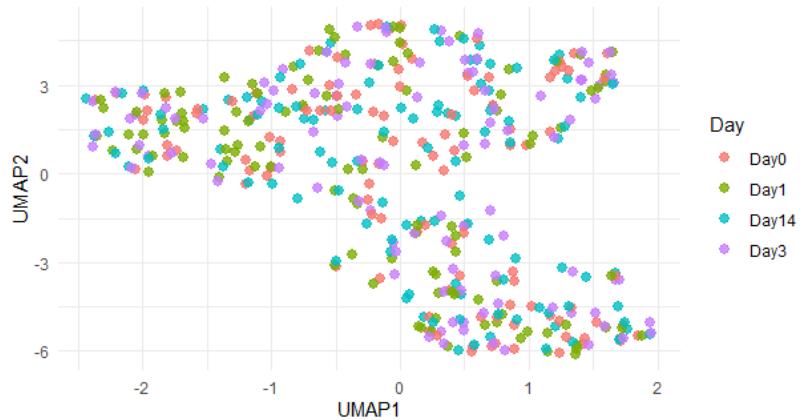
7. Reduksi Dimensi (UMAP)

UMAP digunakan untuk memvisualisasikan kemiripan global antar sampel dengan mereduksi puluhan ribu dimensi gen menjadi dua dimensi (koordinat X dan Y). Tiga versi UMAP dibuat, yaitu menggunakan seluruh gen, 1000 gen paling variabel, dan hanya gen signifikan (DEG), untuk memperjelas struktur data.

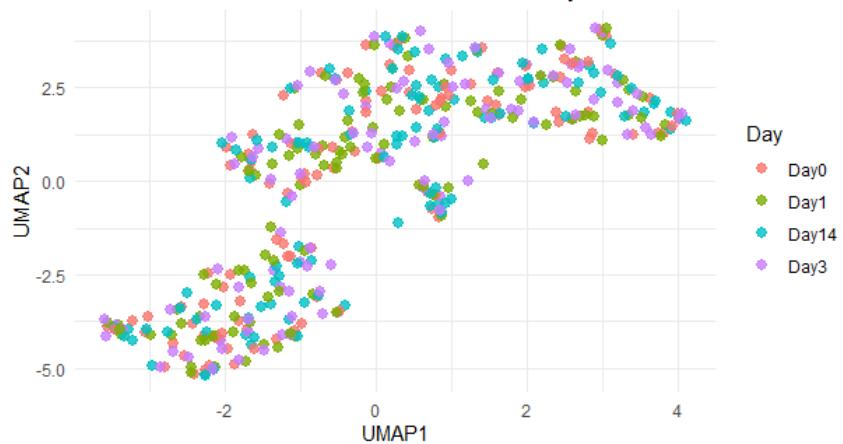
C. Hasil dan Interpretasi



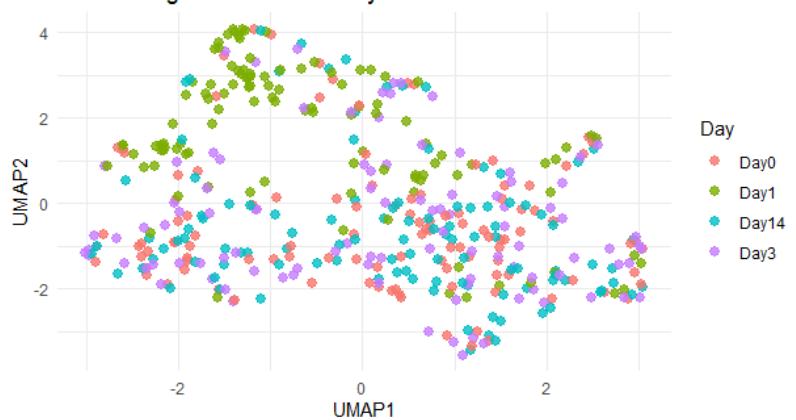
UMAP - Top 1000 Most Variable Genes



UMAP Plot - Influenza Vaccine Response



UMAP - Significant DEGs Only



Volcano plot menunjukkan peningkatan signifikan pada Day 1 dibandingkan Day 0, dengan 525 gen meningkat dan 141 gen menurun. Hal ini menunjukkan adanya respons imun yang kuat setelah vaksinasi. Gen-gen dengan logFC tinggi dan FDR rendah didominasi oleh interferon-stimulated genes, yang merupakan ciri khas respons antivirus pada fase awal imun bawaan.

Heatmap 50 gen paling signifikan menunjukkan pola peningkatan ekspresi yang konsisten pada sampel Day 1, terutama pada gen-gen yang berperan dalam respons interferon dan pertahanan terhadap virus. Pola pengelompokan menunjukkan adanya aktivasi imun yang terkoordinasi meskipun terdapat variasi antar individu.

Analisis GO Biological Process menunjukkan pengayaan signifikan pada kategori “respon terhadap virus” dan “respon terhadap interferon tipe I”, yang mengindikasikan bahwa vaksin influenza memicu aktivasi jalur pertahanan antivirus dalam 24 jam pertama. Nilai p.adjust yang sangat rendah menunjukkan bahwa pengayaan ini bukan kebetulan statistik, melainkan sinyal biologis yang kuat dan konsisten.

KEGG pathway menunjukkan pengayaan pada jalur Influenza A serta beberapa jalur infeksi virus lainnya. Hal ini mencerminkan aktivasi modul pertahanan antivirus yang bersifat umum, terutama interferon signaling dan antigen presentation, yang digunakan tubuh untuk melawan berbagai infeksi virus.

Analisis UMAP menunjukkan bahwa pemisahan antar kondisi tidak terlalu jelas ketika seluruh gen digunakan, yang mencerminkan dominasi variasi antar individu. Namun, ketika analisis difokuskan pada gen paling variabel atau gen signifikan, struktur data menjadi lebih informatif, menunjukkan bahwa perubahan ekspresi gen interferon merupakan faktor utama yang membedakan kondisi pasca vaksinasi.

Secara keseluruhan, analisis menunjukkan bahwa 24 jam setelah vaksinasi influenza terjadi aktivasi kuat respons imun innate yang ditandai dengan peningkatan ekspresi gen-gen interferon-stimulated. Analisis functional enrichment mengonfirmasi pengayaan jalur pertahanan antivirus dan interferon signaling. Temuan ini konsisten dengan mekanisme biologis di mana vaksin memicu keadaan antivirus sementara sebelum berkembangnya respons imun adaptif.

D. Kesimpulan

Berdasarkan analisis differential gene expression terhadap dataset GSE48018, ditemukan bahwa 24 jam setelah vaksinasi influenza terjadi aktivasi kuat dari respons imun bawaan, ditandai dengan peningkatan signifikan gen-gen interferon-stimulated. Analisis pengayaan fungsi biologis dengan metode GO dan KEGG mendukung penemuan ini dengan menunjukkan dominasi jalur respons terhadap virus dan interferon tipe I, yang merupakan mekanisme utama pertahanan antivirus pada tahap awal setelah stimulasi imun.

Hasil ini menunjukkan bahwa vaksin influenza memicu keadaan “antiviral state” secara sistemik dalam waktu singkat, sebelum berkembangnya respons imun adaptif. Pola ekspresi gen yang terkoordinasi pada Day 1 mencerminkan aktivasi jalur molekuler yang spesifik dan konsisten secara biologis.

Peran bioinformatika dalam penelitian ini sangat krusial, karena memungkinkan analisis ribuan gen secara simultan menggunakan pendekatan statistik yang terstandarisasi dan reproduksibel. Melalui tools seperti limma, enrichment analysis, serta visualisasi multidimensi (UMAP), bioinformatika tidak hanya membantu mengidentifikasi gen yang berubah secara signifikan, tetapi juga mengungkap konteks biologis dan jalur molekuler yang mendasarinya. Oleh karena itu, bioinformatika berfungsi sebagai penghubung antara data genomik berskala besar dan pemahaman mekanisme imun dengan cara yang sistematis dan berbasis bukti.