

Differentially Expressed Genes Analysis Using the Web-Based GEO2R Tool

1.) Pendahuluan (latar belakang singkat dan tujuan analisis)

Tuberculosis (TB) adalah penyakit infeksi menular yang disebabkan oleh bakteri *Mycobacterium tuberculosis* yang terutama menyerang paru-paru, tetapi juga dapat memengaruhi organ lain seperti kelenjar getah bening, tulang, dan sistem saraf. TB masih menjadi salah satu penyebab utama kematian akibat penyakit infeksi di dunia, terutama di negara berkembang. Infeksi TB memicu berbagai respon imun pada tubuh manusia, termasuk perubahan ekspresi gen pada sel imun dan jaringan yang terinfeksi. Perubahan ini dapat dipelajari menggunakan pendekatan transcriptomics untuk mengidentifikasi gen-gen yang mengalami perubahan ekspresi signifikan antara individu sehat dan pasien TB.

Analisis Differentially Expressed Genes (DEG) merupakan metode penting dalam bioinformatika untuk mengidentifikasi gen yang mengalami perubahan ekspresi secara signifikan antara dua kondisi biologis yang berbeda. Identifikasi DEG dapat membantu memahami mekanisme molekuler penyakit, menemukan biomarker potensial, serta mendukung pengembangan strategi diagnostik dan terapeutik.

Tujuan dari analisis ini adalah untuk mengidentifikasi gen-gen yang mengalami perubahan ekspresi secara signifikan antara individu sehat dan pasien tuberculosis menggunakan dataset transcriptomics publik GSE313408 dan menganalisisnya menggunakan tool berbasis web GEO2R.

2.) Metode (dataset yang digunakan, pembagian grup, parameter utama GEO2R, dan skema replikasi)

1) Dataset

Dataset yang digunakan adalah GSE313408 yang diperoleh dari Gene Expression Omnibus (GEO), yang berisi data ekspresi gen dari sampel individu sehat dan pasien tuberculosis. Dataset ini memungkinkan analisis perbandingan ekspresi gen antara kondisi healthy (kontrol) dan disease (tuberculosis).

2) Metode Analisis

Analisis DEG dilakukan menggunakan GEO2R yang menggunakan paket statistik limma. Sampel dibagi menjadi dua kelompok, yaitu healthy dan tuberculosis. Analisis menggunakan

koreksi multiple testing Benjamini–Hochberg False Discovery Rate (FDR). Kriteria gen signifikan ditetapkan sebagai adjusted p-value < 0.05 dan $|\log_2 \text{fold change}| \geq 1$.

3) Reproducibility (Replikasi Analisis)

Analisis dilakukan sebanyak tiga kali menggunakan parameter dan pembagian kelompok yang sama. Hasil menunjukkan pola yang konsisten dalam distribusi gen signifikan, volcano plot, dan pemisahan klaster UMAP, yang menunjukkan bahwa analisis memiliki tingkat reproduktibilitas yang baik.

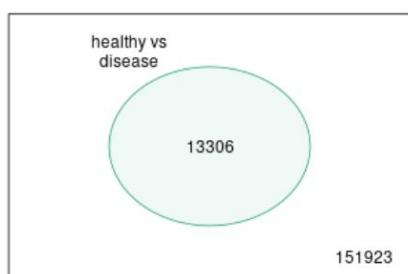
3.) Hasil dan Interpretasi (ringkasan output DEG + interpretasi biologis)

Berdasarkan analisis Differentially Expressed Genes (DEG) menggunakan GEO2R pada dataset GSE313408 yang membandingkan kelompok healthy dan tuberculosis patients,

Venn diagram

Explore and download

GSE313408: limma, Padj<0.05



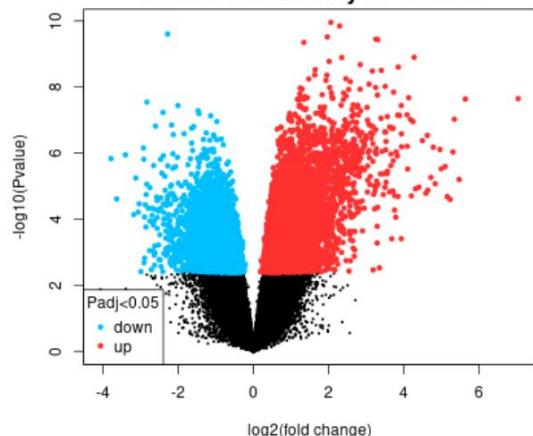
diperoleh sejumlah gen yang menunjukkan perbedaan ekspresi signifikan. Berdasarkan Venn diagram dengan kriteria adjusted p-value < 0.05 , teridentifikasi sebanyak 13.306 gen yang mengalami perubahan ekspresi signifikan antara kelompok healthy dan disease, dari total sekitar 151.923 gen yang dianalisis. Hal ini menunjukkan bahwa infeksi tuberculosis menyebabkan perubahan ekspresi gen yang luas pada tingkat transcriptome.

Visualisasi menggunakan volcano plot menunjukkan distribusi gen berdasarkan nilai log₂ fold change dan signifikansi statistik. Titik berwarna merah menunjukkan gen yang mengalami up-regulation, sedangkan titik berwarna biru menunjukkan gen yang mengalami down-regulation pada pasien tuberculosis dibandingkan individu sehat. Jumlah gen yang mengalami up-regulation terlihat lebih dominan dibandingkan gen yang mengalami down-regulation, yang menunjukkan adanya aktivasi ekspresi gen tertentu sebagai bagian dari respon biologis terhadap infeksi tuberculosis. Selain itu, beberapa gen menunjukkan nilai log₂ fold

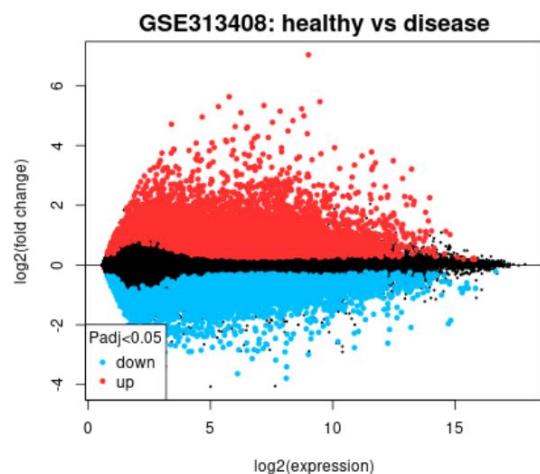
Volcano plot

Explore and download

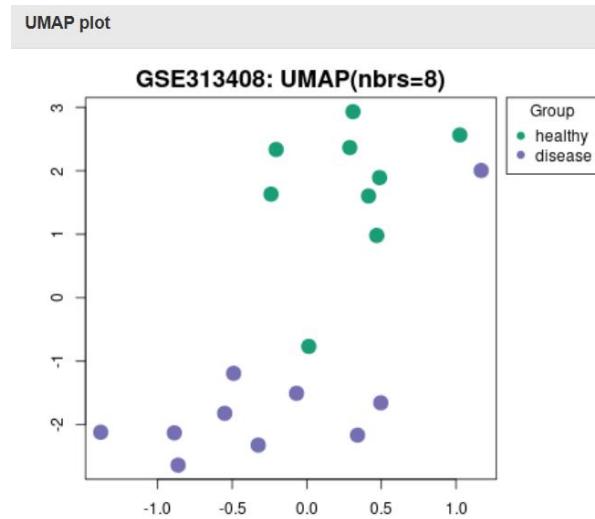
GSE313408: healthy vs disease



change yang tinggi, yang mengindikasikan perubahan ekspresi yang kuat secara biologis.



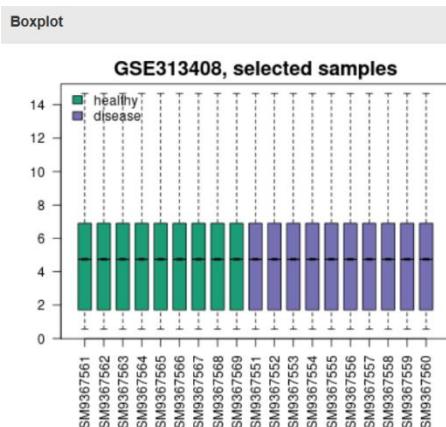
nol, yang menunjukkan adanya perubahan ekspresi gen yang konsisten antara kedua kelompok. Hal ini menunjukkan bahwa infeksi tuberculosis memengaruhi berbagai kelompok gen tanpa bergantung pada tingkat ekspresi awal gen tersebut.



MA plot (Mean Average plot) menunjukkan hubungan antara tingkat ekspresi gen rata-rata dan perubahan ekspresi (log fold change). Pada plot ini terlihat bahwa gen-gen signifikan tersebar pada berbagai tingkat ekspresi, baik pada gen dengan ekspresi rendah maupun tinggi. Gen up-regulated (ditandai warna merah) dan down-regulated (warna biru) tersebar secara jelas dari garis

sistematis pada ekspresi gen, sehingga sampel dapat dibedakan secara jelas berdasarkan status penyakit.

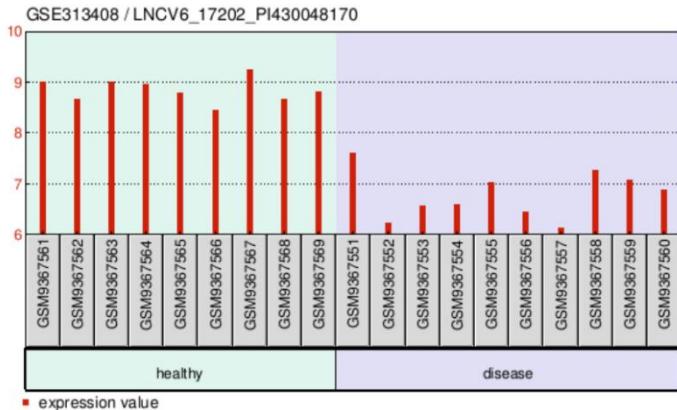
UMAP plot menunjukkan pemisahan yang jelas antara kelompok healthy dan disease berdasarkan profil ekspresi gen. Sampel healthy (warna hijau) membentuk klaster yang berbeda dari sampel tuberculosis (warna ungu), yang menunjukkan bahwa profil transcriptome kedua kelompok memiliki karakteristik yang berbeda secara signifikan. Pemisahan klaster ini menunjukkan bahwa kondisi tuberculosis menyebabkan perubahan



Boxplot distribusi ekspresi gen menunjukkan bahwa distribusi nilai ekspresi antar sampel relatif seragam dan telah terstandardisasi dengan baik. Median dan distribusi ekspresi gen pada setiap sampel terlihat konsisten, yang menunjukkan bahwa data telah mengalami normalisasi dengan baik dan tidak terdapat bias teknis yang signifikan antar sampel. Hal ini penting untuk memastikan bahwa perbedaan

ekspresi gen yang terdeteksi benar-benar disebabkan oleh perbedaan kondisi biologis, bukan oleh variasi teknis.

Expression plot pada salah satu gen representatif juga menunjukkan bahwa tingkat ekspresi gen pada kelompok healthy secara umum lebih tinggi dibandingkan kelompok disease, yang



mengindikasikan adanya gen yang mengalami down-regulation pada pasien tuberculosis. Pola ini konsisten dengan hasil analisis DEG secara keseluruhan yang menunjukkan adanya perubahan ekspresi gen signifikan antara kedua kelompok.

Secara biologis, perubahan ekspresi gen yang signifikan ini mencerminkan respon imun tubuh terhadap infeksi *Mycobacterium tuberculosis*. Infeksi tuberculosis diketahui mengaktifkan berbagai jalur imun, termasuk respon inflamasi, aktivasi makrofag, dan produksi sitokin. Gen-gen yang mengalami up-regulation kemungkinan berperan dalam mekanisme pertahanan tubuh terhadap infeksi, sedangkan gen yang mengalami down-regulation mungkin terkait dengan gangguan fungsi normal sel atau strategi bakteri untuk menghindari respon imun host. Secara keseluruhan, hasil ini menunjukkan bahwa tuberculosis menyebabkan perubahan signifikan pada profil ekspresi gen, yang mencerminkan interaksi kompleks antara patogen dan sistem imun host.

4.) Kesimpulan

Analisis Differentially Expressed Genes menggunakan dataset GSE313408 dan tool GEO2R berhasil mengidentifikasi sejumlah gen yang mengalami perubahan ekspresi signifikan antara individu sehat dan pasien tuberculosis. Gen-gen yang mengalami up-regulation menunjukkan aktivasi respon imun terhadap infeksi *Mycobacterium tuberculosis*, sedangkan gen yang mengalami down-regulation menunjukkan kemungkinan gangguan fungsi biologis normal akibat infeksi.