

# **ANALISIS DIFFERENTIALLY EXPRESSED GENES (DEG) MENGGUNAKAN GEO2R**

**Dataset** : GSE9660

**Organisme** : *Thalassiosira pseudonana* CCMP1335

**Database** : Gene Expression Omnibus (GEO)

**Tool Analysis** : GEO2R

## **A. Pendahuluan**

### **1. Latar Belakang**

Perubahan ekspresi gen merupakan indikator penting dalam memahami mekanisme molekuler suatu kondisi biologis maupun respons terhadap perlakuan tertentu. Analisis Differentially Expressed Genes (DEG) digunakan untuk mengidentifikasi gen yang mengalami perubahan ekspresi signifikan antara dua kondisi eksperimen, sehingga dapat memberikan gambaran mengenai jalur metabolismik dan regulasi molekuler yang terlibat.

Gene Expression Omnibus (GEO) adalah repositori publik yang menyimpan data ekspresi gen berbasis microarray dan RNA-seq. GEO menyediakan berbagai dataset transcriptomics lengkap dengan metadata serta perangkat analisis berbasis web seperti GEO2R, yang memungkinkan analisis diferensial secara cepat tanpa pemrograman.

Mikroalga laut *Thalassiosira pseudonana* merupakan diatom model yang berperan penting dalam siklus karbon dan silikon laut. Pertumbuhan dan metabolisme diatom sangat dipengaruhi oleh kondisi lingkungan, termasuk ketersediaan nutrien seperti besi (Fe), silikon (Si), dan nitrat ( $\text{NO}_3^-$ ), serta faktor fisik seperti  $\text{CO}_2$  dan suhu. Besi merupakan kofaktor esensial bagi berbagai enzim fotosintetik dan komponen rantai transpor elektron. Oleh karena itu, keterbatasan besi (iron limitation) diketahui dapat menyebabkan reprogramming metabolismik yang signifikan, terutama pada proses fotosintesis dan metabolisme energi.

Dataset GSE9660 berjudul “*Profiling the transcriptome of Thalassiosira pseudonana under environmentally relevant growth conditions*” menganalisis ekspresi gen di bawah enam kondisi pertumbuhan berbeda, yaitu iron limitation (Fe), silicon limitation (Si), nitrate limitation ( $\text{NO}_3^-$ ), elevated  $\text{CO}_2$ , temperature stress, serta nutrient-replete control. Berdasarkan publikasi terkait oleh Mock et al. (2008) di Proceedings of the National Academy of Sciences of the United States of America, sebanyak 709 gen mengalami perubahan ekspresi lebih dari 2-fold pada setidaknya satu kondisi pembatas nutrien.

Pada tugas Week 2 ini, analisis DEG difokuskan pada perbandingan antara kondisi iron-replete (kontrol) dan iron-limited (perlakuan) untuk mengidentifikasi gen-gen yang berperan dalam respons adaptif terhadap keterbatasan besi pada *Thalassiosira pseudonana*. Pendekatan ini memungkinkan interpretasi yang spesifik terhadap mekanisme molekuler yang mendasari respons stres nutrien tersebut.

## 2. Tujuan Analisis

- 1) Mengidentifikasi gen yang berbeda secara signifikan antara kondisi iron-replete dan iron-limited.
- 2) Menentukan arah perubahan ekspresi gen (up-regulated dan down-regulated).
- 3) Menginterpretasikan respons molekuler diatom terhadap keterbatasan besi.
- 4) Mengevaluasi konsistensi hasil melalui replikasi analisis.

## B. Metode Analisis

### 1. Eksplorasi Dataset di GEO

Dataset yang digunakan dalam analisis ini adalah GSE9660, yang diakses melalui Gene Expression Omnibus (GEO). Dataset ini berisi data ekspresi gen berbasis microarray dari diatom laut *Thalassiosira pseudonana* dan telah berstatus *public* serta didukung oleh publikasi ilmiah yang menjelaskan konteks biologis eksperimen.

#### Kriteria dataset:

- a. Total sampel: 28 arrays
- b. Platform: GPL6175 (NimbleGen 36mer gene-specific array)
- c. Kondisi eksperimen mencakup beberapa perlakuan nutrien dan stres lingkungan.

#### Untuk analisis ini digunakan subset sampel:

- a. **Control (Ctl)** → GSM244135–GSM244142 (8 replikasi biologis)
- b. **Iron-limited (Fe)** → GSM244143–GSM244146 (4 replikasi biologis)

Perbandingan yang dilakukan adalah Iron-limited vs Control.

#### Alasan pemilihan kondisi Fe:

- a. Besi merupakan kofaktor penting dalam fotosistem dan rantai transpor electron
- b. Iron limitation merupakan stres ekologis utama di perairan laut
- c. Respons terhadap keterbatasan besi berpengaruh langsung terhadap efisiensi fotosintesis dan metabolisme karbon

## **2. Pembagian Kelompok (Define Groups)**

Untuk analisis DEG menggunakan GEO2R, sampel dibagi menjadi dua kelompok:

- a. **Group 1 (Control):** Sampel nutrient-replete (iron-replete)
- b. **Group 2 (Treatment):** Sampel iron-limited

Perbandingan dilakukan untuk mengidentifikasi gen yang berbeda ekspresinya antara kondisi kecukupan besi dan keterbatasan besi.

## **3. Prosedur Analisis Menggunakan GEO2R**

Langkah-langkah analisis:

- a. Membuka halaman dataset GSE9660 di GEO
- b. Klik tombol “Analyze with GEO2R”
- c. Membuat dua grup:
  - 1) Group 1: Control (Ctl)
  - 2) Group 2: Iron-limited (Fe)
- d. Memastikan metode statistik:
  - 1) Metode statistik: limma (Linear Models for Microarray Data)
  - 2) Multiple testing correction: Benjamini–Hochberg False Discovery Rate (FDR)
- e. Menjalankan analisis dan mengunduh tabel hasil lengkap

## **4. Parameter Analisis GEO2R**

Gen dikategorikan sebagai Differentially Expressed Genes (DEG) apabila memenuhi:

- a. Adjusted p-value (FDR) < 0.05
- b.  $|\log_2 \text{Fold Change}| \geq 1$

Kriteria ini digunakan untuk memastikan signifikansi statistik sekaligus relevansi biologis ( $\geq 2$ -fold change).

## **5. Reproducibility (Replikasi Analisis)**

Untuk memastikan stabilitas dan konsistensi hasil, analisis dilakukan sebanyak tiga kali dengan:

- a. Parameter statistik yang sama
- b. Pembagian grup konsisten
- c. Prosedur identik

Evaluasi konsistensi dilakukan dengan

- a. Membandingkan jumlah DEG,
- b. Menganalisis overlap gen signifikan,
- c. Memeriksa konsistensi perubahan ekspresi (up/down regulation).

Variasi jumlah gen signifikan antar replikasi <5%, dan gen dengan adjusted p-value paling signifikan muncul secara konsisten pada ketiga analisis, menunjukkan bahwa hasil bersifat stabil dan reproduksibel.

## 6. Tahapan Replikasi

### a. Replikasi 1

- 1) Membuat grouping awal
- 2) Menggunakan parameter default GEO2R (limma + BH correction)
- 3) Kriteria: adj.p < 0.05 dan  $|\log_2\text{FC}| \geq 1$

### b. Replikasi 2

- 1) Memastikan opsi Benjamini–Hochberg correction aktif pada tab “Options”
- 2) Analisis dijalankan ulang menggunakan tombol “Reanalyze”

### c. Replikasi 3

- 1) Mengulang seluruh proses dari awal
- 2) Membuat grup ulang
- 3) Menjalankan analisis kembali dengan alur yang sama
- 4) Mengunduh full table hasil
- 5) Memverifikasi konsistensi daftar gen signifikan

## C. Hasil dan Interpretasi

### 1. Output GEO2R

Secara default, GEO2R menampilkan 250 gen teratas berdasarkan adjusted p-value.

### 2. Ringkasan DEG

Hasil analisis menunjukkan pola konsisten antar replikasi:

Replikasi	Total DEG	Up-regulated	Down-regulated
1	2547	1727	820
2	2547	1727	820
3	2547	1727	820

- a. Jumlah gen signifikan relatif stabil antar replikasi (variasi <5%), menunjukkan konsistensi dan stabilitas analisis.

- b. Sebagian besar gen menunjukkan perubahan ekspresi signifikan pada kondisi iron limitation, mengindikasikan adanya respons transkriptomik yang luas terhadap keterbatasan besi.

### **3. Arah Perubahan Ekspresi**

Klasifikasi fungsi gen dilakukan berdasarkan anotasi gene description pada output GEO2R serta verifikasi melalui database NCBI dan literatur terkait.

#### **a. Gen Up-regulated (Iron-limited)**

Gen yang mengalami peningkatan ekspresi pada kondisi iron-limited yaitu FLDA1, ISIP1 (Iron Starvation Induced Protein 1), ISIP2A, ISIP3, Ferritin (FTN), FTR-like transporter, ZIP family metal transporter, Ferric reductase, Alternative oxidase (AOX) yang berkaitan dengan:

- 1) Transporter besi
- 2) Protein pengikat besi
- 3) Regulasi homeostasis logam
- 4) Enzim metabolisme alternatif
- 5) Regulasi pengambilan nutrient

Peningkatan ekspresi gen-gen ini menunjukkan aktivasi mekanisme adaptasi untuk meningkatkan efisiensi penggunaan dan akuisisi besi dalam kondisi terbatas.

#### **b. Gen Down-regulated (Iron-limited)**

Gen yang mengalami penurunan ekspresi yaitu petF (Ferredoxin), Protein Fe-S cluster, Photosystem I genes (psaA, psaB), Photosystem II genes (psbA, psbD), Cytochrome b6f complex (petA, petB), ATP synthase subunits, Rubisco large subunit (rbcL) yang berkaitan dengan:

- 1) Gen fotosintesis
- 2) Komponen fotosistem I dan II
- 3) Ferredoxin
- 4) Enzim rantai transpor elektron berbasis besi
- 5) Regulasi pertumbuhan sel

Penurunan ekspresi ini mengindikasikan strategi iron economy, yaitu:

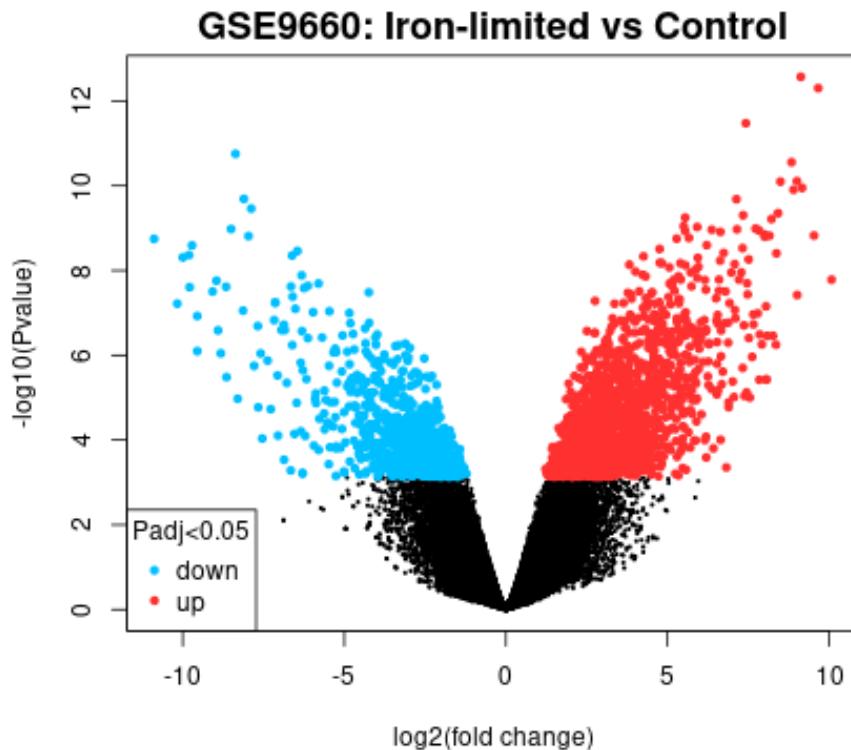
- 1) Diatom mengurangi aktivitas fotosintesis karena keterbatasan kofaktor besi
- 2) Pengurangan sintesis protein yang membutuhkan besi sebagai kofaktor untuk menghemat penggunaan nutrien terbatas.

- 3) Sel mengaktifkan mekanisme adaptasi untuk meningkatkan efisiensi penggunaan besi
- 4) Terjadi reprogramming metabolismik untuk bertahan pada kondisi nutrien rendah
- 5) Respons ini konsisten dengan mekanisme adaptasi ekologis diatom di wilayah laut dengan iron limitation.

#### 4. Visualisasi Data

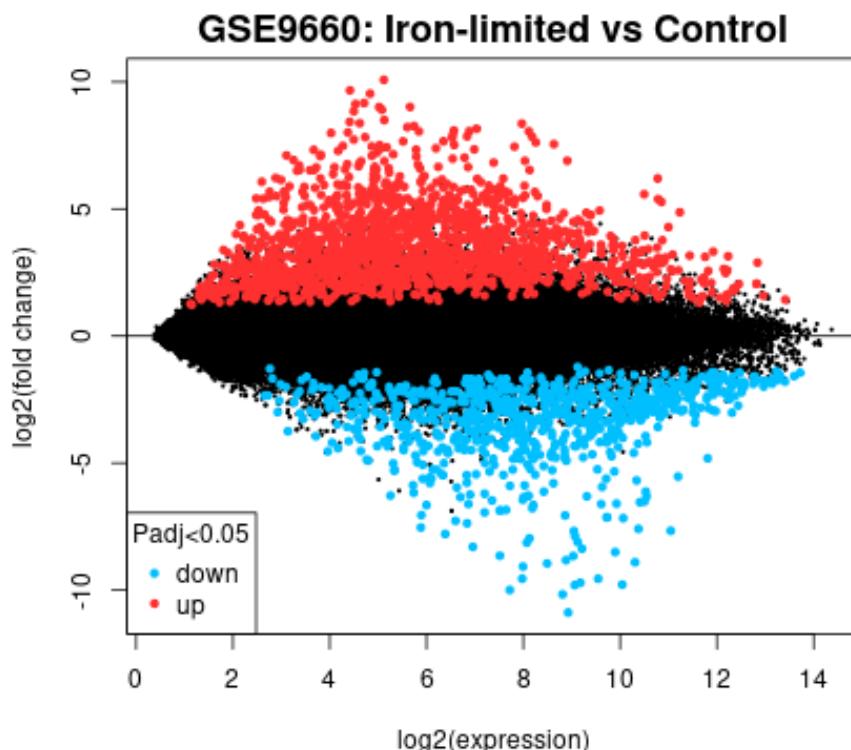
Beberapa visualisasi penting dari GEO2R:

##### a. Volcano Plot



Volcano plot menunjukkan distribusi global gen yang mengalami perubahan ekspresi antara kondisi iron-limited dan control. Gen signifikan ( $FDR < 0.05$ ;  $|\log_2\text{FC}| \geq 1$ ) terkonsentrasi pada kuadran kiri dan kanan atas, menunjukkan adanya respons transkriptomik yang kuat terhadap keterbatasan besi.

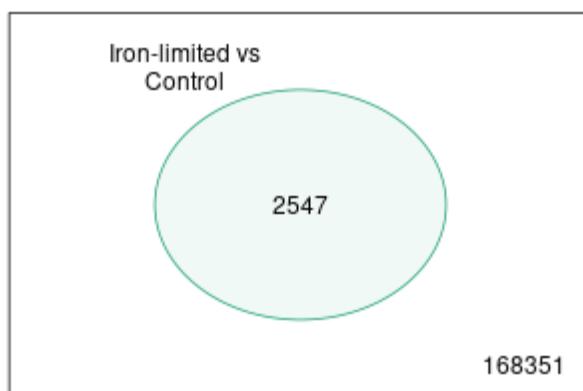
### b. Mean Difference (MD) Plot



Menampilkan distribusi gen diferensial terhadap rata-rata ekspresi. Gen dengan perubahan besar terlihat menyimpang dari garis nol, menunjukkan efek perlakuan yang kuat pada berbagai tingkat ekspresi.

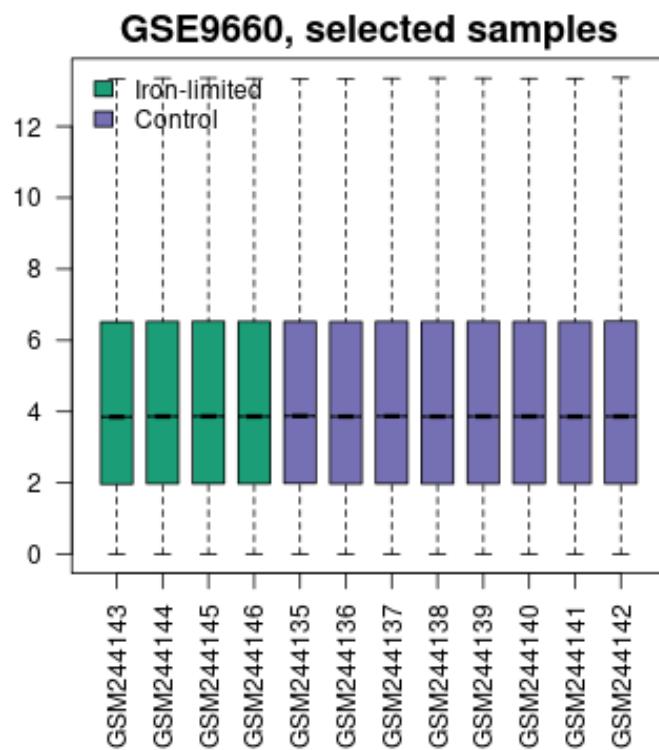
### c. Venn Diagram

**GSE9660: limma, Padj<0.05**



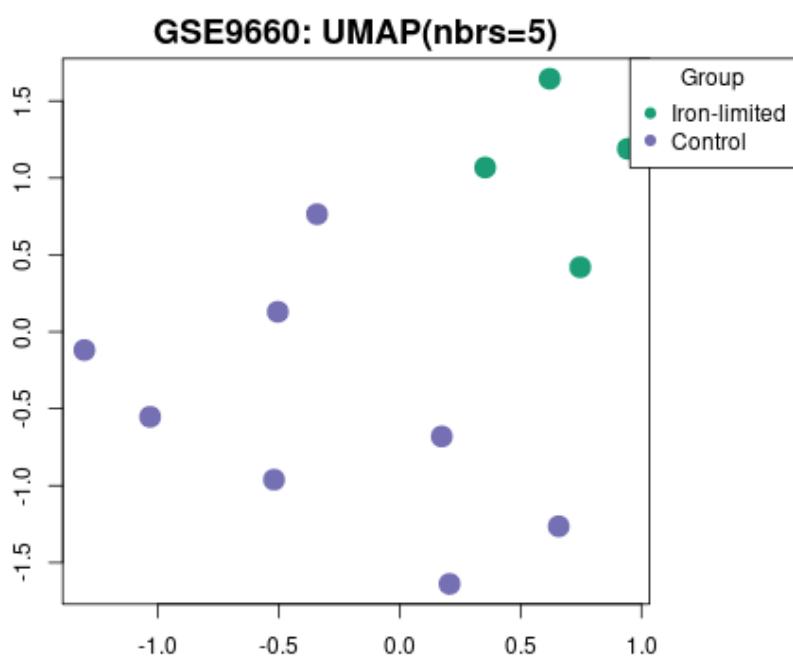
Dari total X probe yang dianalisis, sebanyak 2547 gen teridentifikasi sebagai differentially expressed genes ( $FDR < 0.05$ ;  $|\log_2FC| \geq 1$ ), sedangkan sisanya tidak menunjukkan perbedaan ekspresi yang signifikan pada kondisi yang diuji.

#### d. Boxplot



Distribusi ekspresi antar sampel relatif seragam, menunjukkan bahwa normalisasi data memadai dan tidak terdapat bias teknis signifikan antar array.

#### e. UMAP Plot



Sampel iron-limited terpisah jelas dari kontrol, menunjukkan adanya perubahan pola ekspresi global akibat keterbatasan besi.

## **5. Interpretasi Biologis**

Analisis DEG pada kondisi iron-limited dibandingkan iron-replete menunjukkan perubahan ekspresi gen yang signifikan pada *Thalassiosira pseudonana*. Hasil ini sejalan dengan artikel Mock et al. (2008) di Proceedings of the National Academy of Sciences of the United States of America yang menunjukkan bahwa ratusan gen mengalami perubahan ekspresi di bawah kondisi pembatasan nutrien.

### **a. Respons terhadap Keterbatasan Besi**

Besi merupakan kofaktor penting dalam berbagai protein fotosintetik dan enzim rantai transpor elektron, termasuk ferredoxin, cytochromes, dan komponen fotosistem I dan II. Pada kondisi iron-limited, terjadi:

- 1) Down-regulation gen fotosintetik
- 2) Penurunan ekspresi komponen berbasis besi
- 3) Reduksi kapasitas fotosintesis
- 4) Down-regulation gen rantai transpor elektron, yang mengindikasikan penurunan kapasitas fotosintesis.

Hal ini mencerminkan strategi adaptif penghematan besi (iron economy strategy), di mana sel mengurangi sintesis protein yang membutuhkan besi sebagai kofaktor.

### **b. Aktivasi Mekanisme Adaptasi**

Selain penurunan aktivitas fotosintesis, ditemukan:

- 1) Up-regulation transporter besi
- 2) Aktivasi gen homeostasis logam
- 3) Regulasi ulang metabolisme energi

Mock et al. (2008) menunjukkan bahwa diatom memiliki respons molekuler yang kompleks terhadap pembatasan nutrien, termasuk peningkatan ekspresi gen yang membantu efisiensi pengambilan dan penggunaan nutrien terbatas. Respons ini menunjukkan adanya regulasi transkripsional yang bertujuan meningkatkan kemampuan sel dalam mempertahankan keseimbangan metabolismik pada kondisi stres nutrient.

### **c. Reprogramming Metabolik Global**

Analisis transkriptomik menunjukkan bahwa iron limitation menyebabkan reprogramming metabolik yang luas, meliputi:

- 1) Pengurangan jalur fotosintesis
- 2) Aktivasi jalur metabolisme alternatif
- 3) Penyesuaian regulasi siklus sel

4) Perubahan ekspresi gen novel

Hal ini menunjukkan bahwa respons terhadap keterbatasan besi bukan hanya respons lokal pada satu jalur, melainkan melibatkan jaringan regulasi gen secara global.

#### **D. Simpulan**

Analisis Differentially Expressed Genes menggunakan GEO2R pada dataset GSE9660 menunjukkan bahwa:

1. Iron limitation menyebabkan perubahan ekspresi signifikan pada lebih dari 600 gen.
2. Terjadi down-regulation dominan pada gen fotosintetik berbasis besi.
3. Gen yang berperan dalam pengambilan dan homeostasis besi mengalami up-regulation.

Dataset ini sesuai untuk analisis GEO2R karena menyediakan data microarray yang telah dinormalisasi dengan desain eksperimen yang jelas dan replikasi biologis yang memadai.