

Exploración de datos de expresión.

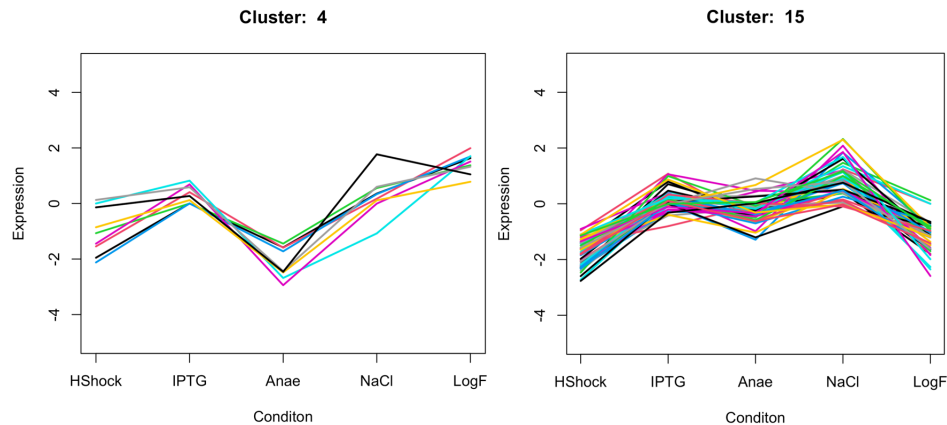
Objetivo:

Determinar si existe una estructura útil en los datos proporcionados en el archivo “datos_lcg.txt”.

Estrategia:

Utilizar el programa de agrupamiento como “herramienta de exploración”. Se corre el *clustering* y se determina si existen grupos limpios donde todos los genes muestran un perfil de expresión similar.

Correr K-means para un número relativamente grande de *clusters* ($k \geq 20$), seleccionar *clusters* pequeños (4-25 genes) y graficar los valores de expresión de cada gene a través de las 5 condiciones de expresión. Por ejemplo:



¿Cuál es el *cluster* con perfil de expresión más consistente que pueden encontrar?

Si no logran encontrar *clusters* suficientemente pequeños para este análisis, incrementen el número de *clusters* que le piden al método de agrupamiento.