

TP3 Analyse différentielle

Justine Guégan - j.guegan-ihu@icm-institute.org
Guillaume Meurice - guillaume.meurice@gustaveroussy.fr

08 décembre 2016

Les données nécessaires à cette séance se trouvent sur le site web :

<https://pf-bb.github.io/CentraleSupelec-R-genomics/TP4.html>.

Objectif

Dans les TP2 et 3, nous avons appris à manipuler des données d'expressions, à évaluer la variabilité de profils transcriptionnels, et à rechercher des gènes différentiellement exprimés entre deux groupes d'échantillons. Aujourd'hui, nous nous pencherons sur l'analyse fonctionnelle de ces gènes : quels sont les mécanismes biologiques sous-jacents, si tant est que nous puissions les identifier ?

Répondez aux questions dans un document Rmarkdown produisant un fichier **PDF** ou **HTML**.

Données

Les ARNm de 17 échantillons de tumeurs du sein de 3 types, HER2 positif (HER2), triple négatif (TNBC), non triple négatif (Non-TNBC), et 3 échantillons de sein normal (NBS) ont été séquencés par Illumina HiSeq2000.

Analyse des données

Lors du TP précédent, vous aviez généré les MA-plot, des volcanoplots et des heatmap permettant de représenter les gènes différentiellement exprimés entre deux groupes d'échantillons. Afin de repartir sur la même base, nous vous donnons aujourd'hui les fichiers suivants :

- Fichiers MA-plot : ces fichiers contiennent en ligne les gènes, et en colonnes, les valeurs **M**, **A**, ainsi que les p.value et p.value ajustée.
 - dataMAplot_HER2_vs_NBS.txt
 - dataMAplot_TNBC_vs_NBS.txt
 - dataMAplot_NTNBC_vs_NBS.txt
- Fichiers heatmap : ces fichiers contiennent un sous-ensemble de la matrice de comptage, réduit aux seuls gènes différentiellement exprimés pour les 3 contrastes (seuils utilisés : $|\log FC| > 1$ et p.value ajustée ≤ 0.05)
 - count_HER2_vs_NBS
 - count_TNBC_vs_NBS
 - count_NTNBC_vs_NBS
- Fichiers de liste de gènes : ces fichiers contiennent simplement la liste des gènes sur- et sous-exprimés dans les 3 contrastes :

- gene_list_HER2_vs_NBS_up.txt
- gene_list_HER2_vs_NBS_dn.txt
- gene_list_TNBC_vs_NBS_up.txt
- gene_list_TNBC_vs_NBS_dn.txt
- gene_list_NTNBC_vs_NBS_up.txt
- gene_list_NTNBC_vs_NBS_dn.txt

1. Lecture des fichiers de données