**Eindopdracht bRNA – Transcriptoom Analyse**

Voor de eindopdracht wordt middels transcriptoom analyse het effect van een behandeling op chronische inflammatie in kaart gebracht.

Van vier 'celtypes' is zowel een onbehandelde ('control') als een behandelde sample ('treated') aanwezig. Informatie hierover is te vinden in het bestand *chronische\_inflammatie\_metadata.csv*. Een RNA-seq experiment is uitgevoerd waarna de kwaliteit van de ruwe reads is bepaald. De reads zijn vervolgens ge-aligned tegen een humaan referentie genoom wat resulteerde in de count tabel die te vinden is in het bestand *chronische\_inflammatie\_rawcounts.csv*. De counts zijn bepaald op gen-niveau. Voor annotatie kan het bestand *annotables\_grch37.csv* gebruikt worden.

***Opdracht omschrijving***

* Deze opdracht is een groepsopdracht en mag in tweetallen uitgevoerd en ingeleverd worden.
* **Schrijf een script** waarmee de genexpressie verschillen tussen behandelde en onbehandelde cellen wordt bepaald. Hierbij moet rekening gehouden worden met celtype-specifieke effecten (gepaarde analyse).
* **Schrijf een rapport** waarbij de belangrijkste uitkomsten uit je genexpressie analyse workflow geïnterpreteerd worden. Gebruik hierbij waar nodig minimaal? 1 a 2 relevante figuren per onderdeel.
* **De onderwerpen/stappen die minimaal aan de orde moeten komen in zowel script als rapport zijn:**
  + Het verwijderen van genen die laag tot expressie komen
  + QC op de count data
  + Hiërarchische clustering
  + Normalisatie
  + Testen van differentiële genexpressie
  + Gen set enrichment of pathway level analysis
* **In het rapport moet, naast het bovenstaande, ook aandacht worden besteed aan de biologische interpretatie van de resultaten:**
  + Korte terugkoppeling gene set enrichment/pathway analyse resultaten naar de behandeling van chronische inflammatie (maximale lengte 1 pagina).
    - **Alternatief**: vergelijk de lijst met significante genen uit het artikel (zie ELO) met de door jou gegenereerde lijst. Leg uit hoe en waarom deze van elkaar verschillen (maximale lengte 1 pagina).

***Beoordeling***

* **Om deelpunten te krijgen moet je script voor dat stukje van de analyse foutloos runnen, dit is een voorwaardelijke eis.**
* Voor juistheid van de interpretatie in het rapport van de hierboven genoemde onderwerpen/stappen kunnen per stap punten gehaald worden. In totaal zijn er 12 punten te behalen. Voor een voldoende beoordeling (V) moeten minimaal 9 punten behaald worden.
* Toelichting cesuur: voor de uitvoering van deze opdracht mag gebruik gemaakt worden van de vele aanwezige voorbeeld scripts en uitgewerkte analyses op internet. De cesuur is hier echter wel op aangepast en ligt derhalve wat hoger dan normaal.

**Inhoudelijke eisen:**

Filteren van lage expressie

a) script runt foutloos *voorwaardelijk*

b) interpretatie in rapport voldoende *1 punt*

QC op count data

a) script runt foutloos *voorwaardelijk*

b) interpretatie in rapport voldoende *1 punt*

Hierarchische clustering

a) script runt foutloos *voorwaardelijk*

b) interpretatie in rapport voldoende *2 punten*

Normalisatie

a) script runt foutloos *voorwaardelijk*

b) interpretatie in rapport voldoende *2 punten*

DEG testing

a) script runt foutloos *voorwaardelijk*

b) interpretatie in rapport voldoende *3 punten*

GSEA/Pathwaylevel analyse

a) script runt foutloos *voorwaardelijk*

b) interpretatie in rapport voldoende *3 punten*

**TOTAAL: 12 punten**