

# BRIE2\_tutorial

Anne Brussaard

2025-05-21

## 0.1 Deelvraag: Kan ik door het volgen van een tutorial met BRIE2 data analyse uitvoeren?

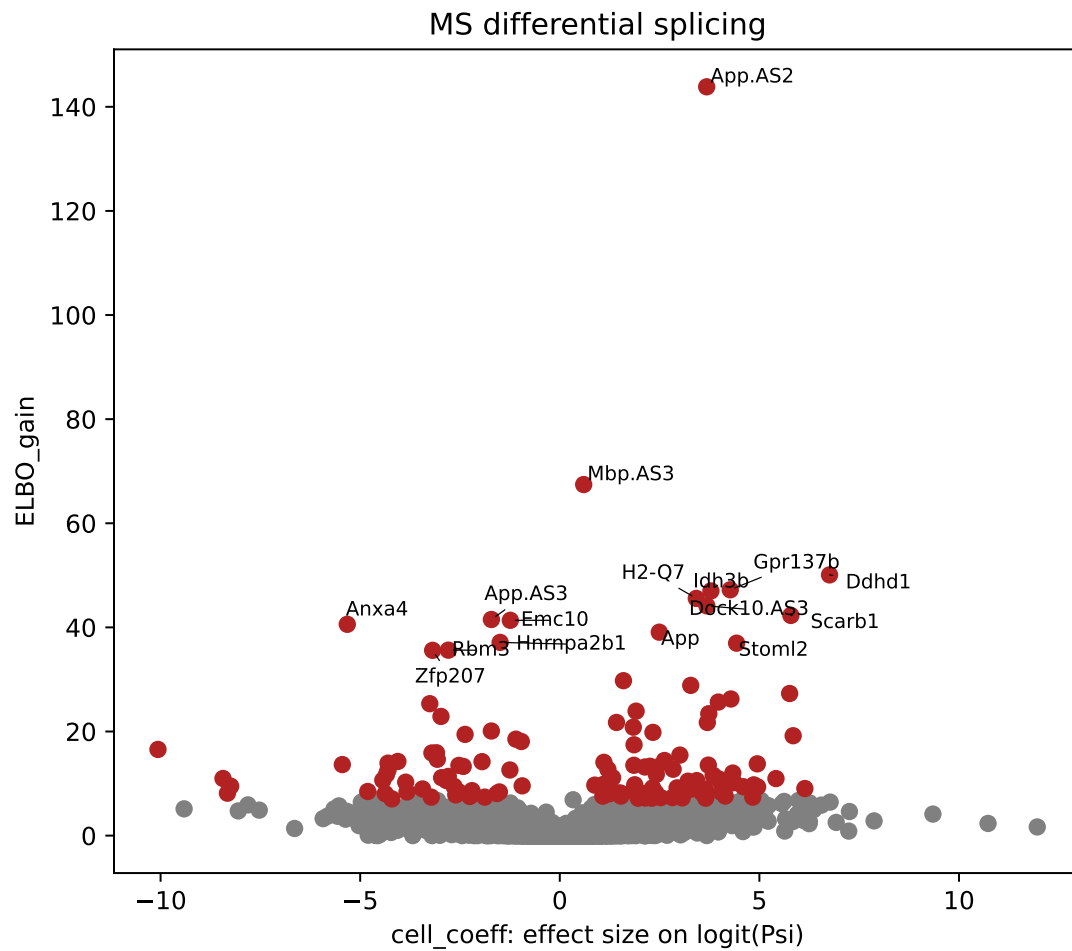
De data die is gebruikt voor het uitvoeren van deze tutorial is van Falcao et al, 2018. Het bevat 2208 muizencellen die onderzocht zijn met SMART-seq2 waarbij de ene helft experimentele auto0immuun encefalomyelitis (EAE) cellen zijn. Deze cellen bootsen multiple sclerose na, de andere helft zijn controle cellen. Het wordt gebruikt om na te bootsen hoe BRIE2 gebruikt kan worden om differentiele splicing events tussen twee groepen cellen te detecteren.

Eerst worden de packages geladen die nodig zijn voor het uitvoeren van de analyse. dit gaat o.a. om umap, os, brie, numpy, pandas, scanpy en matplotlib.pyplot.

## 0.2 BRIE2 optie 1

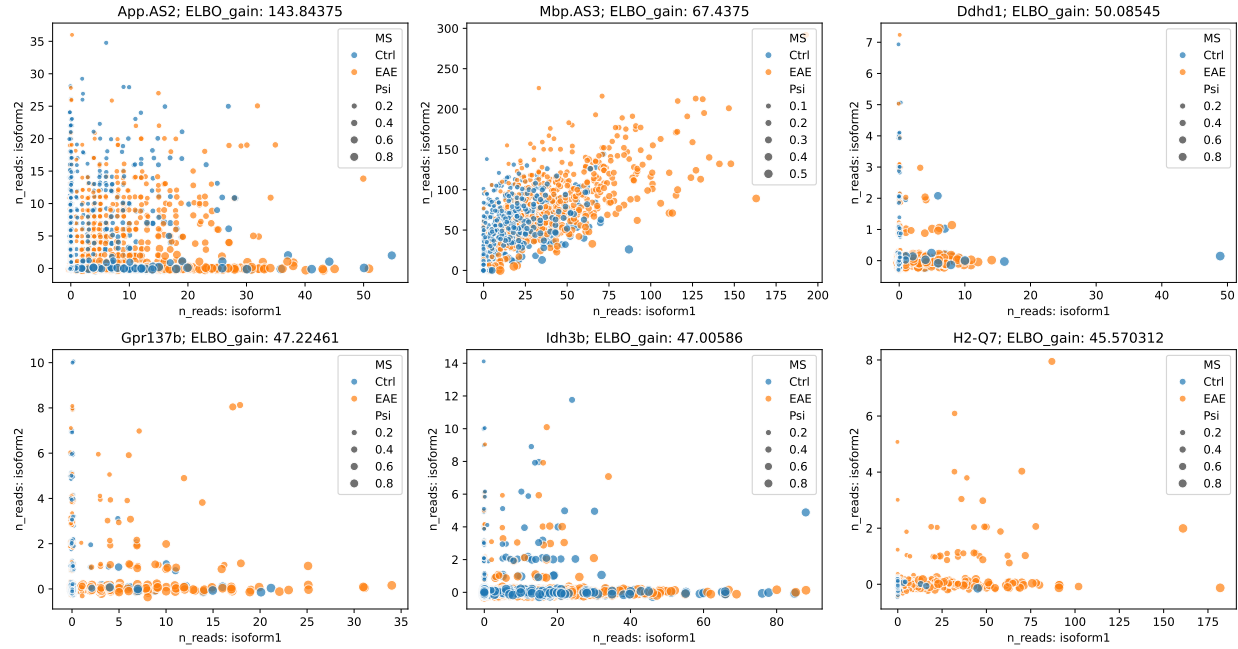
Hierna volgt de eerste mogelijkheid van analyseren met BRIE2. Optie 1= differential splicing events. Deze modus gebruikt statistische analyse (regressie) om de splicing in de twee groepen te vergelijken. In deze modus wordt ook rekening gehouden met celtype zodat het effect apart kan worden bekeken per celtype.

De resultaten van optie 1 worden eerst weergegeven in een volcano plot. In deze plot is Cell\_coeff een statistische maat die aangeeft hoe sterk de splicing (PSI) verandert tussen twee groepen. Een positieve Cell\_coeff geeft aan dat de PSI waarden in de EAE-groep hoger is dan in de controle groep. PSI geeft hierin aan wat de kans is dat een bepaald splicing event voorkomt.



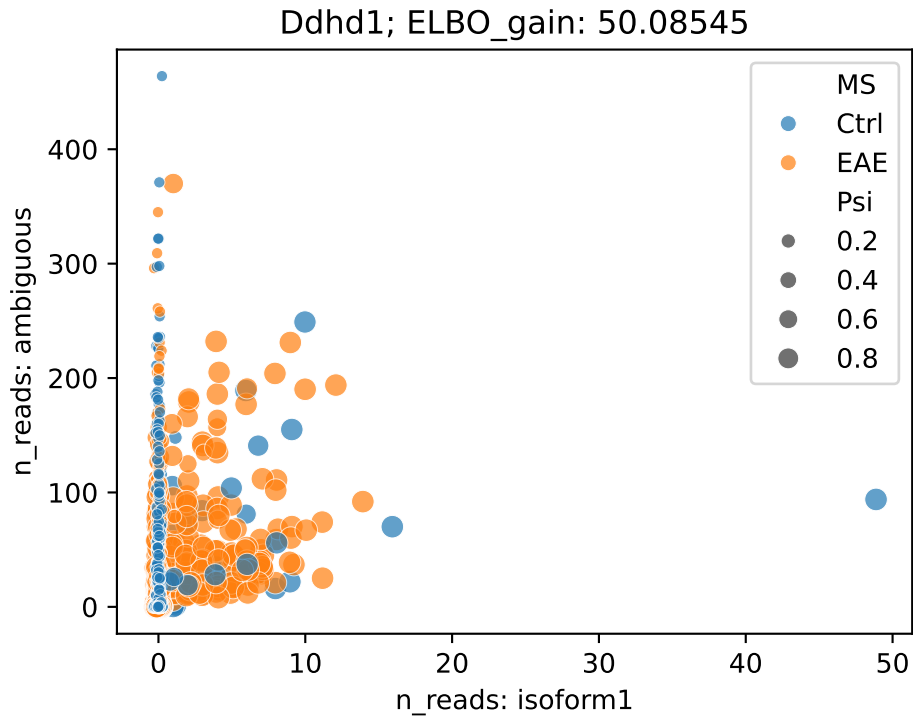
Vervolgens worden de splicing events met een ELBO\_gain van  $> 7$  geselecteerd dat zijn de differentiale splicing events. Er wordt gekeken naar hoeveel events dat zijn en hoeveel genen ze representeren zodat deze gebruikt kunnen worden in verdere analyse.

Daarna volgt de visualisatie van de ruwe counts van de differentiale splicing events (DSE). Dit wordt gedaan om te controleren of wat BRIE2 berekend ook zichtbaar is in de data



Deze plots laten de condities zien met de controle groep blauw en de EAE groep oranje.

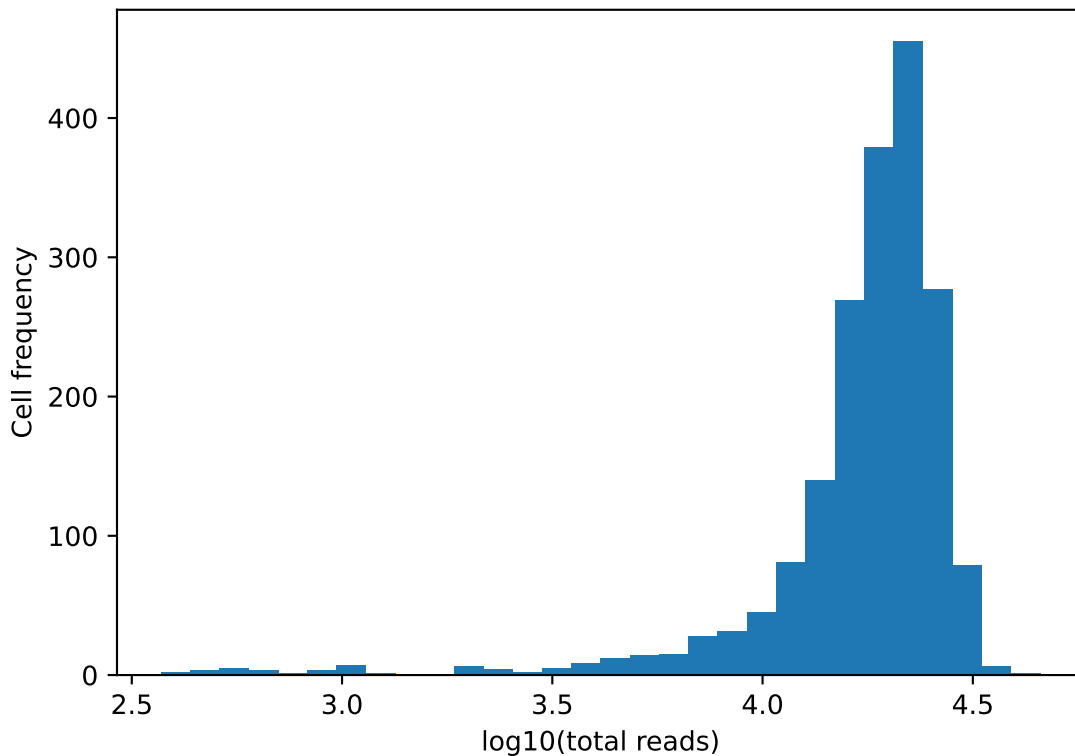
Sommige splicing events kunnen niet direct aan een isovorm worden toegewezen en hebben ambiguos (onduidelijke) reads. De oorzaken hiervan kunnen verschillen van lage kwaliteit, hoge sequentie gelijkenis, geen goed referentiegenoom, of overlappende genen.



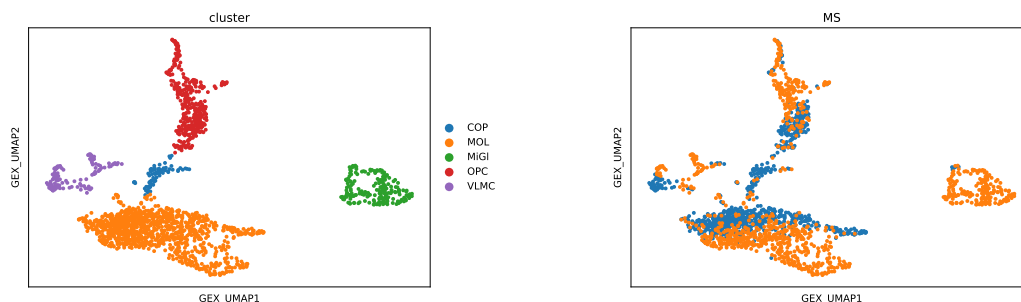
### 0.3 BRIE2 optie 2

Met de tweede optie van BRIE2 kan de kwantificatie van de splicing events uitgevoerd worden. In dit geval wordt dit gedaan met de aggregatie van de cellen om te voorkomen dat dit invloed heeft op je biologische hypothese. Dit houdt in dat de cellen als groepen worden geanalyseerd en niet als individuele cellen zodat gekeken kan worden naar verschillende celtypes.

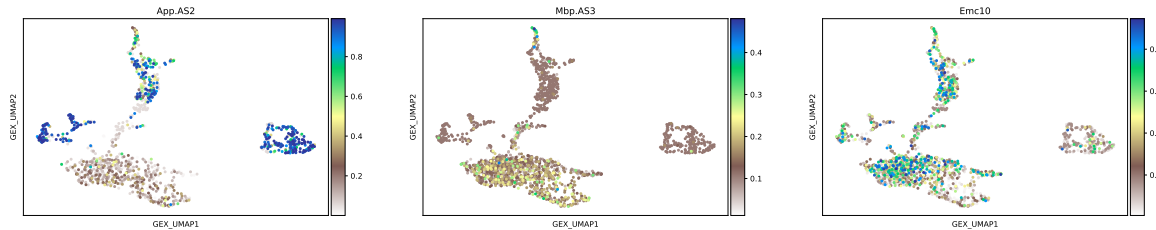
Ook worden de cellen gefilterd op basis van leesdiepte. In de onderstaande histogram zijn het aantal reads per cel te zien. Hierna worden alleen de cellen met voldoende reads ( $>3000$  reads) meegenomen om te zorgen voor hogere betrouwbaarheid.



Daarna worden de splicing fenotypes gevisualiseerd op basis van gen expressie in een Umap. Eerst met de verdeling van clusters en visualisatie van de EAE groep en controle groep. Iedere punt in de UMAP is een cel en zo wordt de verdeling van de cellen over de condities bekeken. Te zien is dat er 3 PC's zijn met een grote ratio, deze zullen gebruik worden voor verdere analyse.

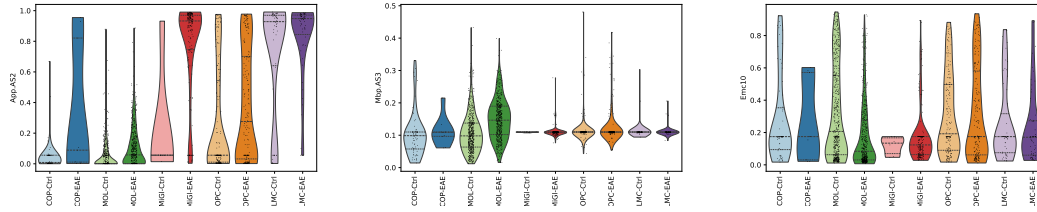


Vervolgens worden de genen App.As2, Mbp.AS3 en Emc10 weergegeven in een scatterplot en violin plot. Voor de genen App.AS2, Mbp.AS3 en Emc10 wordt de geschatte splicing weergegeven. Hierin laat de kleur zien wat de geschatte PSI waarde is en kan gekeken worden waar bepaalde splicing events meer of minder voorkomen.



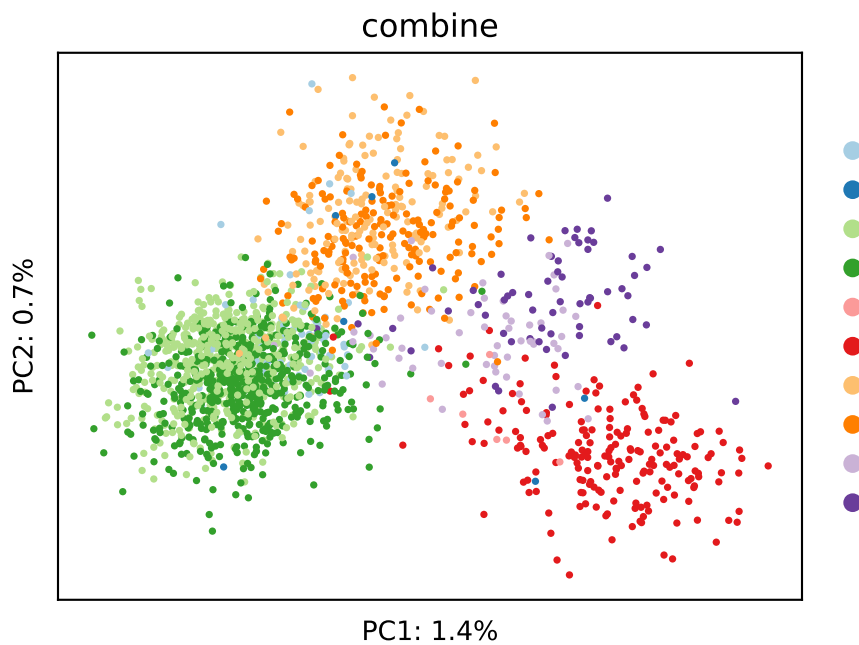
In de violinplots wordt per groep en conditie de verdeling in PSI waarde weergegeven.

```
## /home/anne.brussaard/miniconda3/envs/TFProb/lib/python3.11/site-packages/scanpy/plotting/_anndata.py
##
## Passing 'palette' without assigning 'hue' is deprecated and will be removed in v0.14.0. Assign the '
##
## ax = sns.violinplot(
## /home/anne.brussaard/miniconda3/envs/TFProb/lib/python3.11/site-packages/scanpy/plotting/_anndata.py
## ax = sns.violinplot(
## /home/anne.brussaard/miniconda3/envs/TFProb/lib/python3.11/site-packages/scanpy/plotting/_anndata.py
##
## The 'scale' parameter has been renamed and will be removed in v0.15.0. Pass 'density_norm='width' f
## ax = sns.violinplot(
## /home/anne.brussaard/miniconda3/envs/TFProb/lib/python3.11/site-packages/scanpy/plotting/_anndata.py
##
## Passing 'palette' without assigning 'hue' is deprecated and will be removed in v0.14.0. Assign the '
##
## ax = sns.violinplot(
## /home/anne.brussaard/miniconda3/envs/TFProb/lib/python3.11/site-packages/scanpy/plotting/_anndata.py
## ax = sns.violinplot(
## /home/anne.brussaard/miniconda3/envs/TFProb/lib/python3.11/site-packages/scanpy/plotting/_anndata.py
##
## The 'scale' parameter has been renamed and will be removed in v0.15.0. Pass 'density_norm='width' f
## ax = sns.violinplot(
## /home/anne.brussaard/miniconda3/envs/TFProb/lib/python3.11/site-packages/scanpy/plotting/_anndata.py
##
## Passing 'palette' without assigning 'hue' is deprecated and will be removed in v0.14.0. Assign the '
##
## ax = sns.violinplot(
## /home/anne.brussaard/miniconda3/envs/TFProb/lib/python3.11/site-packages/scanpy/plotting/_anndata.py
## ax = sns.violinplot(
## /home/anne.brussaard/miniconda3/envs/TFProb/lib/python3.11/site-packages/scanpy/plotting/_anndata.py
##
## The 'scale' parameter has been renamed and will be removed in v0.15.0. Pass 'density_norm='width' f
## ax = sns.violinplot(
```

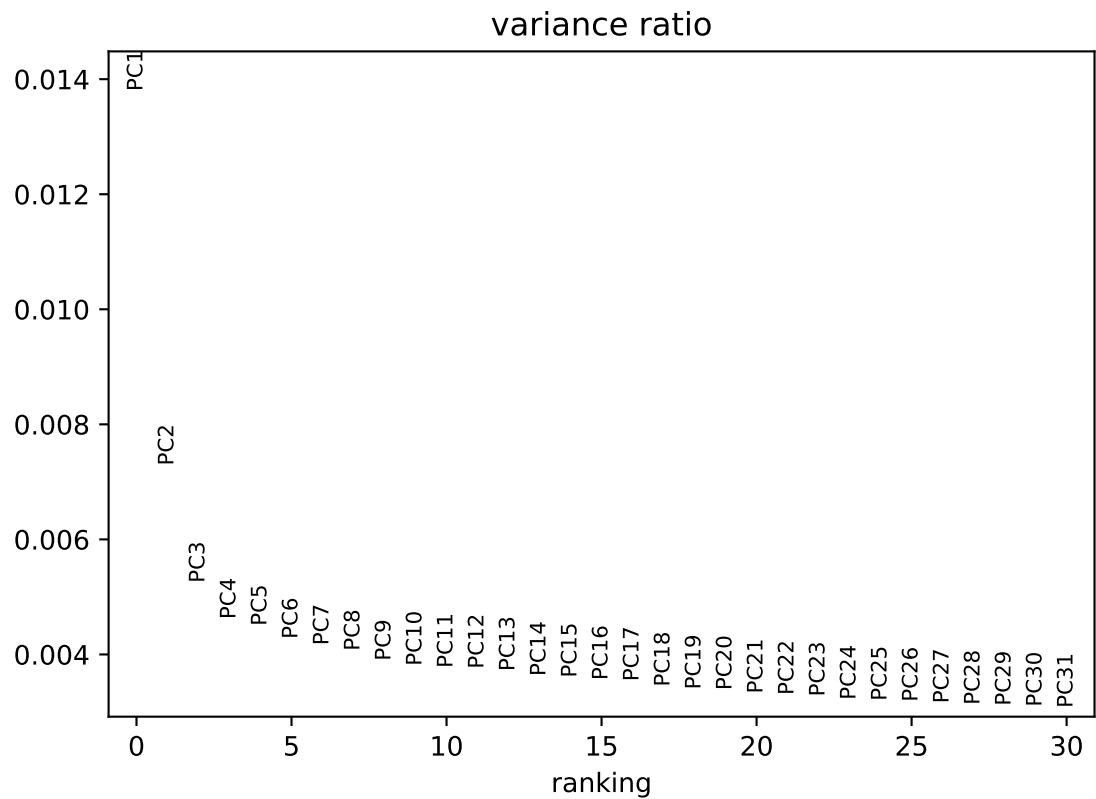


Voor verdere downstream analyse worden alleen de splicing events meegenomen die gedetecteerd zijn als differentiaile splicing events.

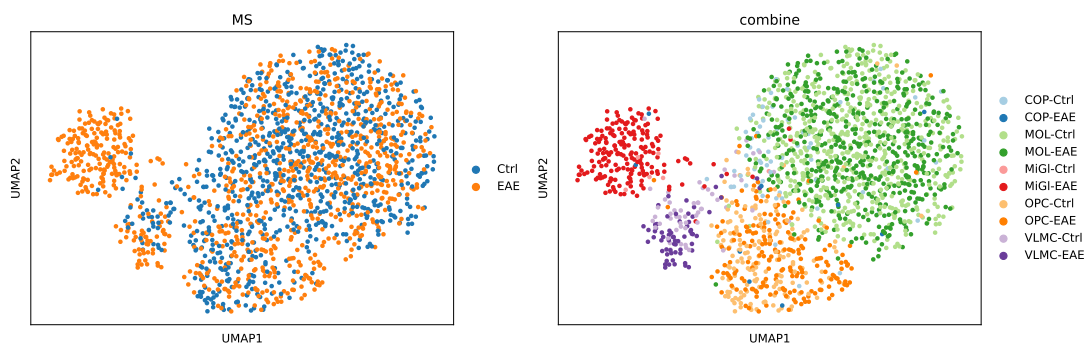
## (1845, 3780)



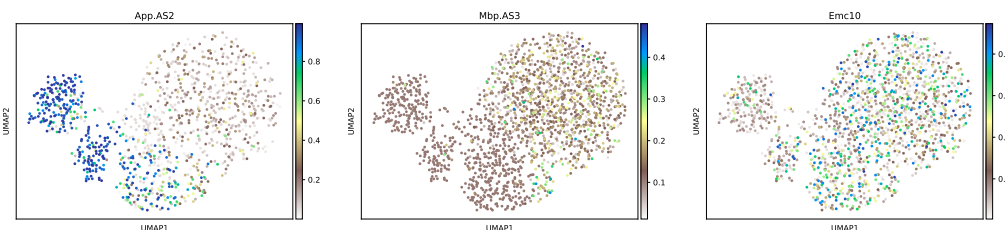
Ook wordt de variance ratio tussen de PSI waarden berekend. Om dit te doen is eerst de splicing data omgezet naar componenten (PC). Dit houdt in dat er gekeken wordt welke PSI waarden verschillen tussen de twee groepen en statistisch relevant zijn.



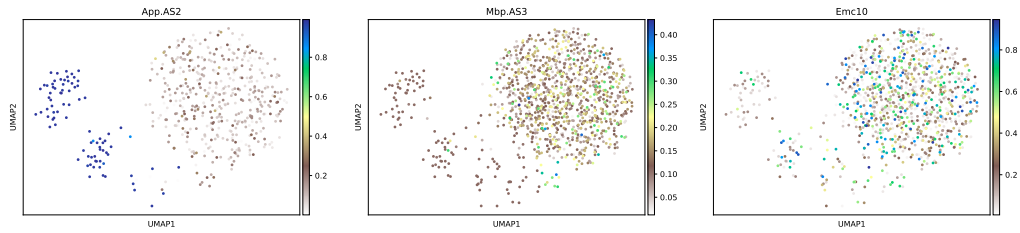
Dit kan vervolgens gebruikt worden om clusters te maken op basis van splicing in plaats van expressie. In de onderstaande UMAP is te zien wat de clusters zijn op basis van PSI waarden. Hieruit kan ook gezien worden of de splicing overeenkomt met de MS/controlle.



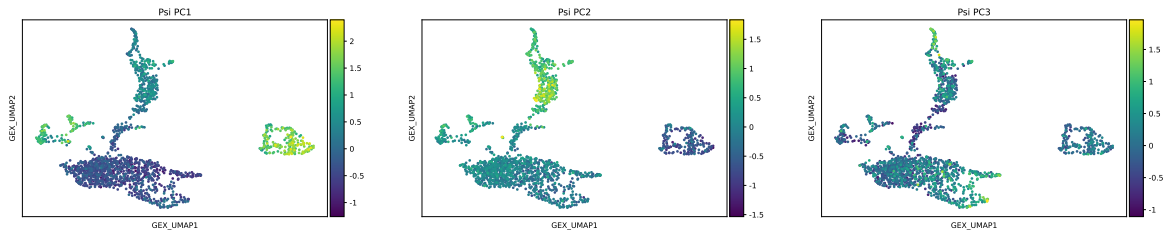
Hier wordt dit uitgezet voor de genen App.AS2, Mbp.AS3 en Emc10



Daarna wordt de UMAP nogmaals gemaakt, dit keer bevat de UMAP alleen cellen met een betrouwbare PSI, voor een nauwkeurige visualisatie van de splicing.



Vervolgens zijn de eerste 3 PC geplot tegen de oorspronkelijke genexpressie UMAP. Dit kan laten zien of de splicing overeenkomt met het expressie cluster.



## 0.4 Conclusie

In deze tutorial is BRIE2 toegepast op een dataset van muizencellen om differentiële splicing te analyseren tussen controle- en EAE-cellen. Zowel op individueel celniveau als op geaggregeerd niveau kon BRIE2 onderscheid maken tussen splicing events.