

# M2 Bio-Informatic Internship

03/10/2025 Presentation

Jordan Dutel

Université Claude Bernard Lyon 1

Centre de Recherche en Cancérologie de Lyon (CRCL)

Team : Dr Pierre Saintigny

Tutor : Dr Pierre Martinez

March 6, 2025



## 1 Introduction

## 2 Workflow

## 3 Results

## 4 Outlook

## 1 Introduction

## ② Workflow

## 3 Results

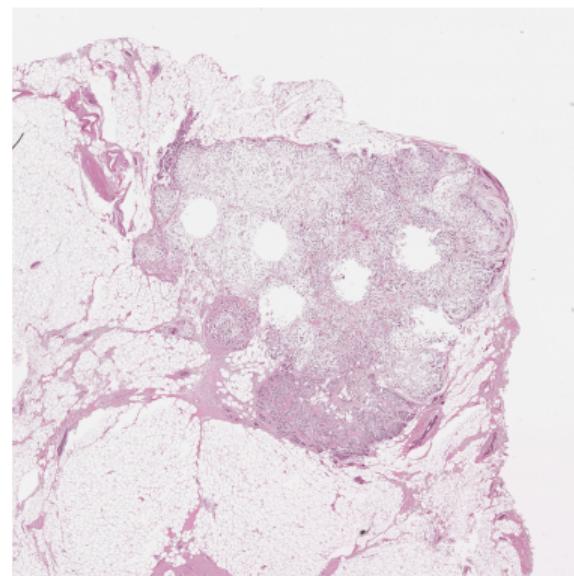
4 Outlook

Data

- 15 patients
  - Mixte MpBC (Metaplastic Breast Cancer) samples
  - Different tumor phenotypes
    - Squamous
    - Epithelial
    - Mesenchymal
    - Spindle-like (fusiform)
    - Chondroid
  - Spatial transcriptomic counts (*Visium*, 10X Genomics)

# Images

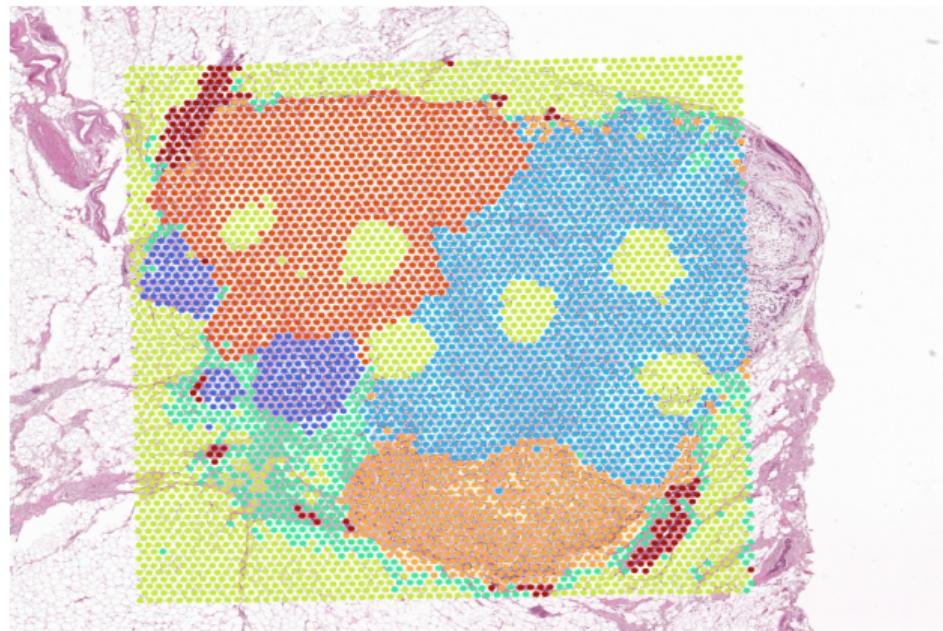
## MpBC9 patient



**Figure 1:** Raw MpBC9 tissue

## Images

MpBC9 patient



- Artifacts
  - Classical chondrosarcoma cells
  - Epithelial tumor cells
  - Intermediate tumour cells
  - Mixoid chondrosarcoma cells
  - Normal fibrous tissue
  - Normal epithelium

**Figure 2:** Annotated MpBC9 tissue

## Research questions

- ① Determine specific markers of different tumor cell phenotypes (for diagnosis)
  - ② Define the genes and pathways involved in trans-differentiation
  - ③ Analyze the genomic and microenvironmental differences between compartments

## 1 Introduction

## ② Workflow

## 3 Results

4 Outlook

## Pipeline

## Steps

- ① Load data
  - ② Merging objects
  - ③ Normalisation
  - ④ PCA
  - ⑤ Harmony (Batch effect correction)
  - ⑥ UMAP
  - ⑦ Find cell-clusters
  - ⑧ Find cluster-specific markers

## Tools

<b>Step</b>	<b>R Packages/Functions</b>
Read Data	Seurat
Normalization	NormalizeData() + ScaleData() (No SCTransform !!!)
Batch effect correction	RunHarmony()

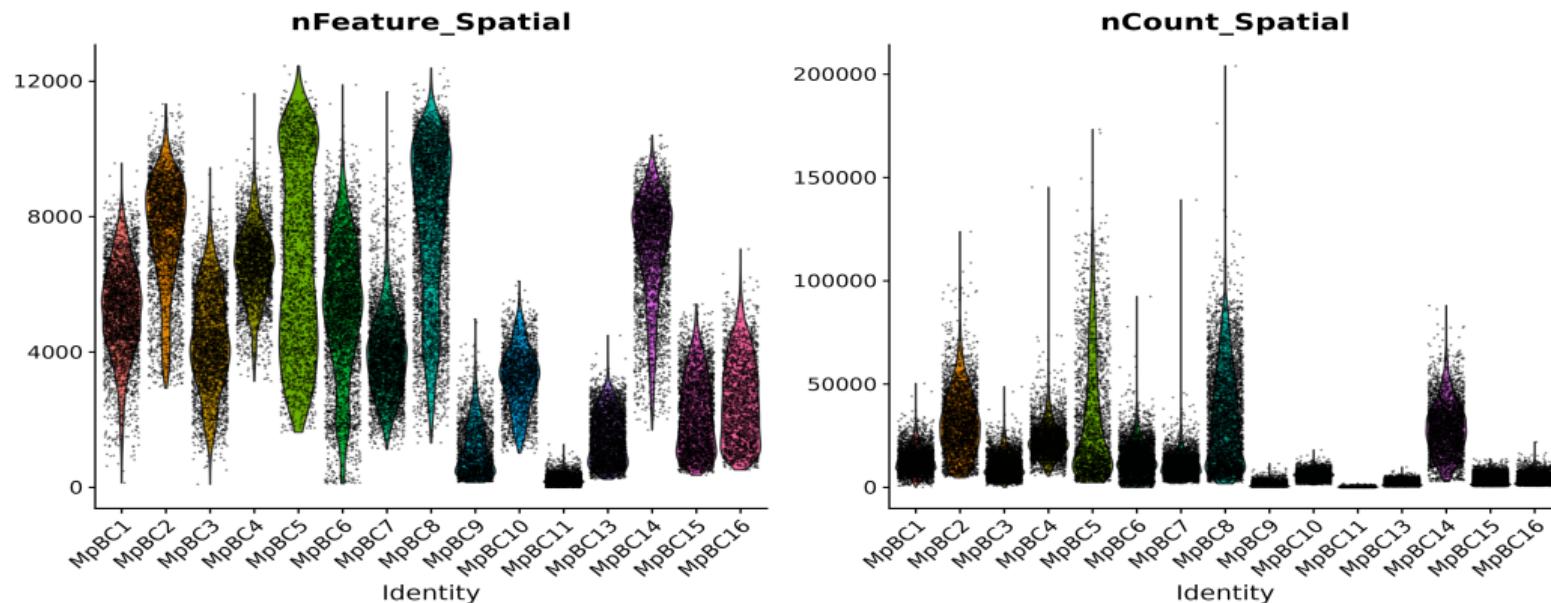
## 1 Introduction

## 2 Workflow

## 3 Results

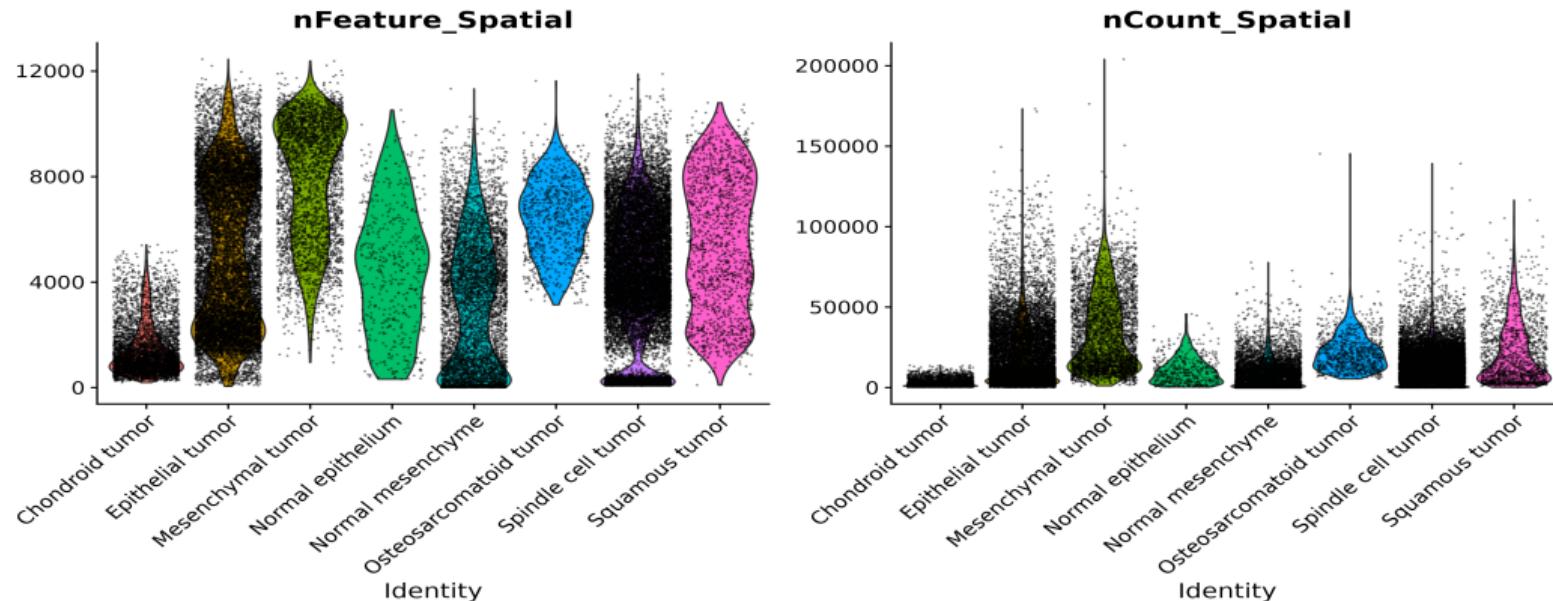
## 4 Outlook

## Quality Control



**Figure 3:** Genes and UMI counts by patients

## Quality Control



**Figure 4:** Genes and UMI counts by tumor cell types

## Normalization

## Points to discuss

- Using Seurat::NormalizeData() + Seurat::ScaleData() (rather than Seurat::SCTransform())
  - Log Normalization
  - Several parameters (scale.factor, number of variable features...) empirically determined -> 750 genes
  - Attempt to normalize individually each object (rather than normalize the whole object) -> No significant differences

## Dimension reduction : PCA

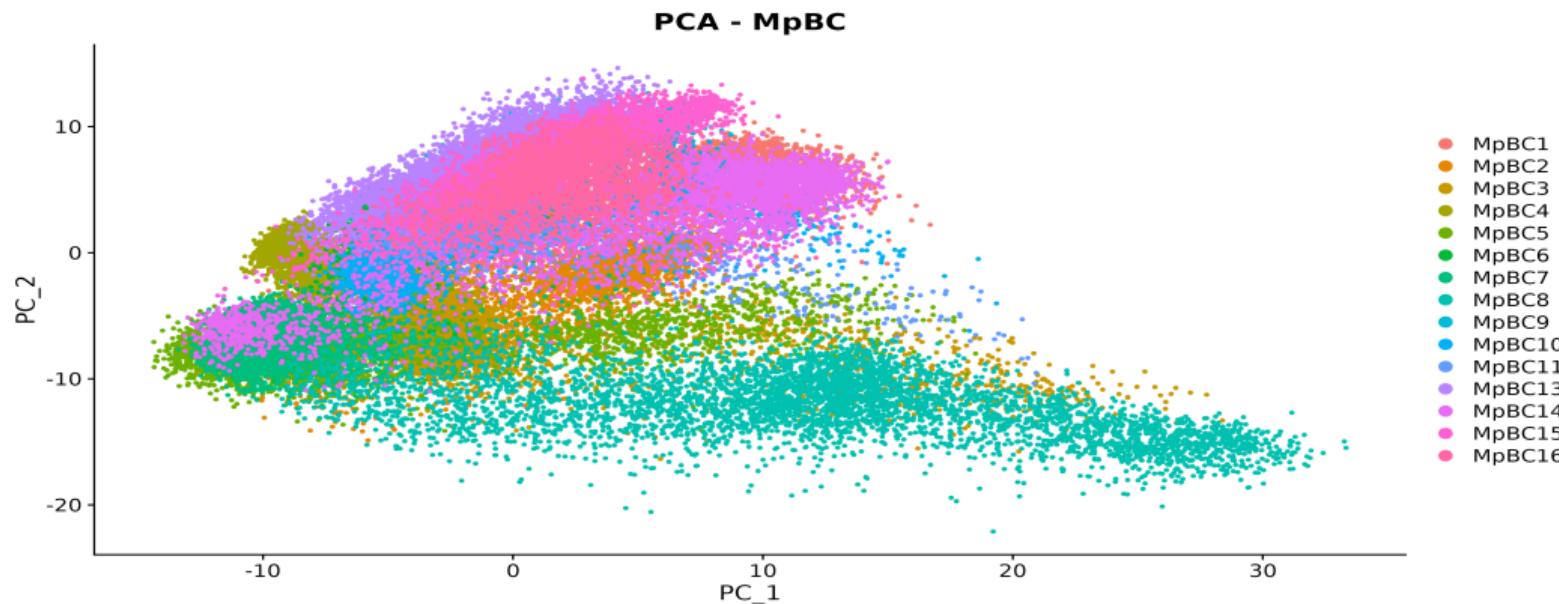
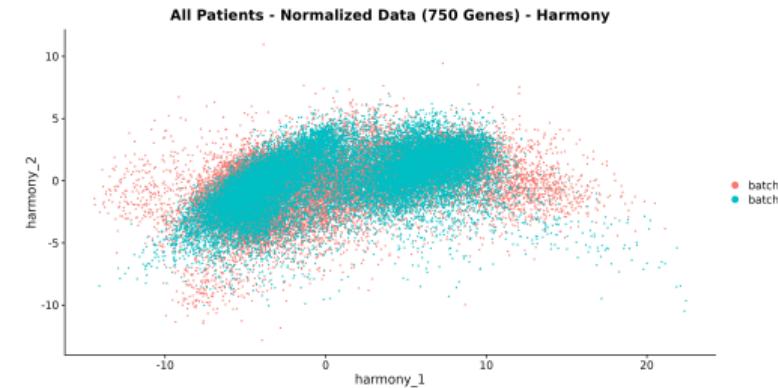
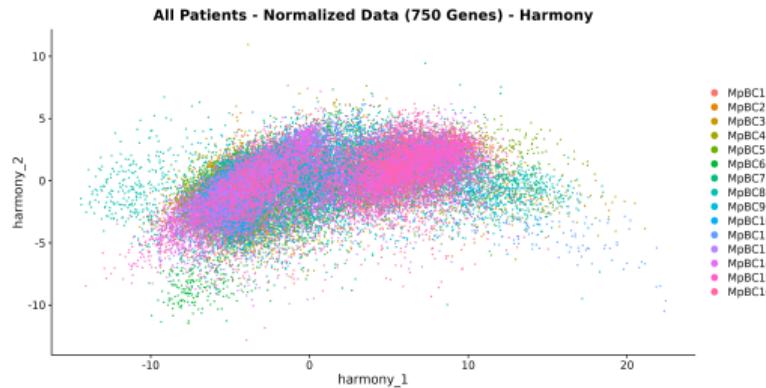
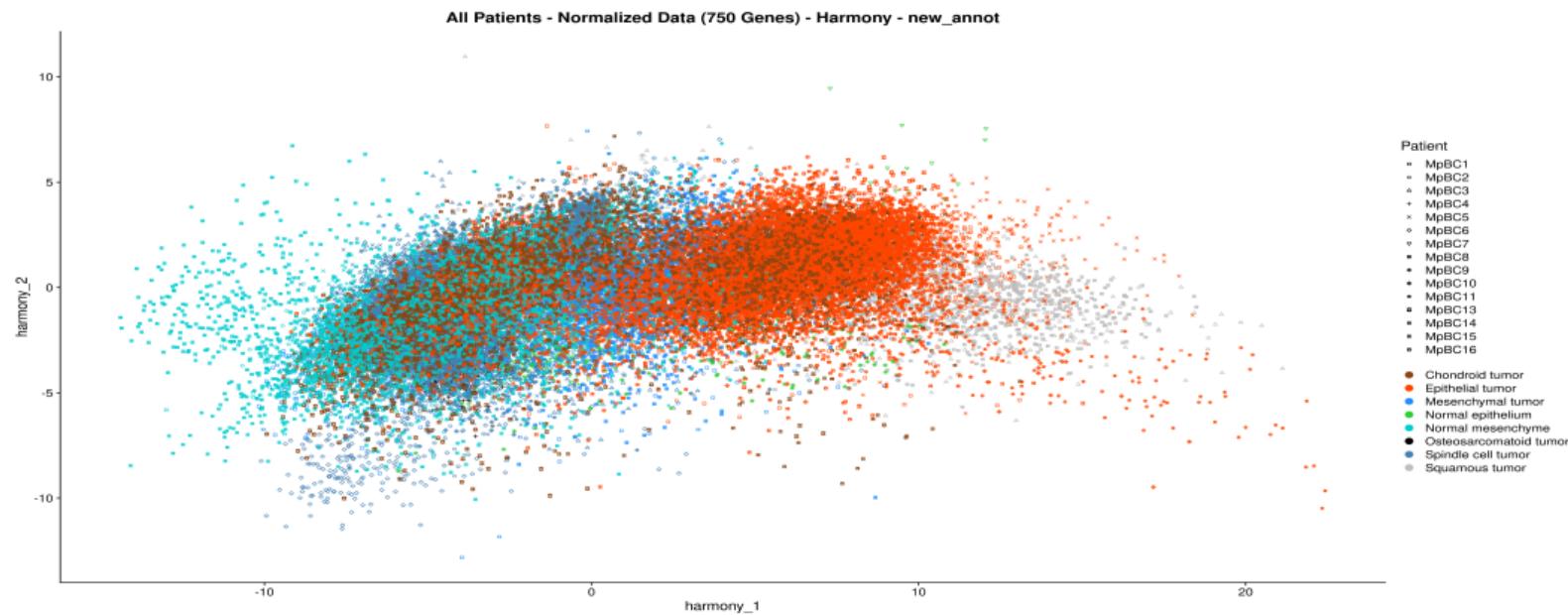


Figure 5: PCA reduction

## Batch effect correction : Harmony



## Batch effect correction : Harmony



**Figure 6:** Harmony correction by tumor cell types

# UMAP Projection

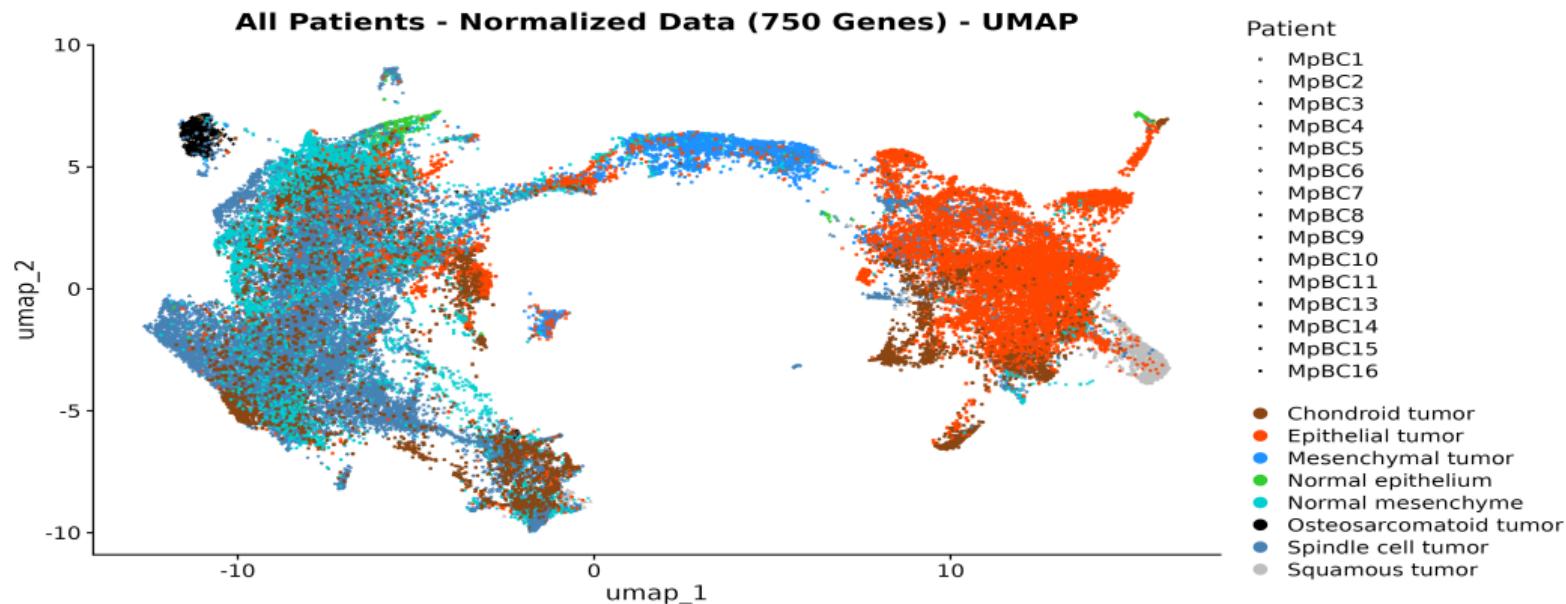


Figure 7: UMAP projection

## UMAP Projection

## Points to discuss

- Some points/cells are still patient-specific (Chondroid and Epithelial cells).
    - Enhance Harmony correction (but jeopardize biological signal) ?
  - Chondroid cells difficult to regroup and Spindle cells really heterogenous...

## Solutions

- Single nuclei and spot deconvolution to improve cell groups according to cell types/annotations

# UMAP Projection

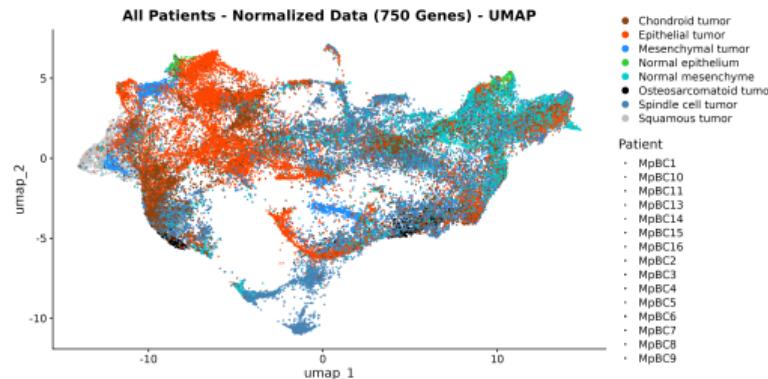
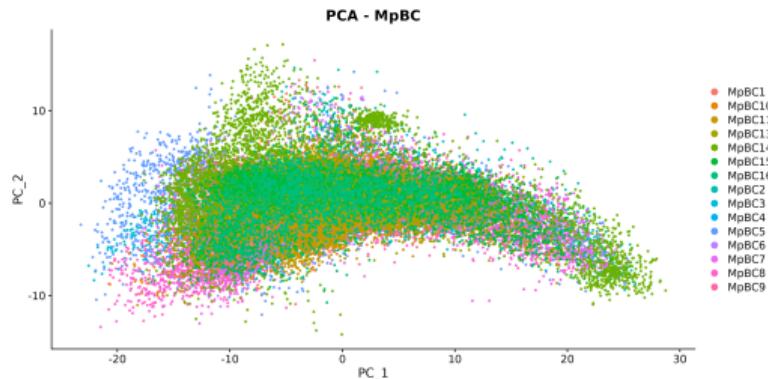
## Question

- As samples have different tumor composition, how do we manage the risk of overcorrection with Harmony and loss of biological signal ?

## Alternatives

- Seurat Integration (Seurat::FindAnchors())
    - Not conclusive (for our dataset)
    - Samples with different tumor cell types -> Too few common anchors (features) shared between all samples

## UMAP Projection



Best option with Harmony correction

# Clustering

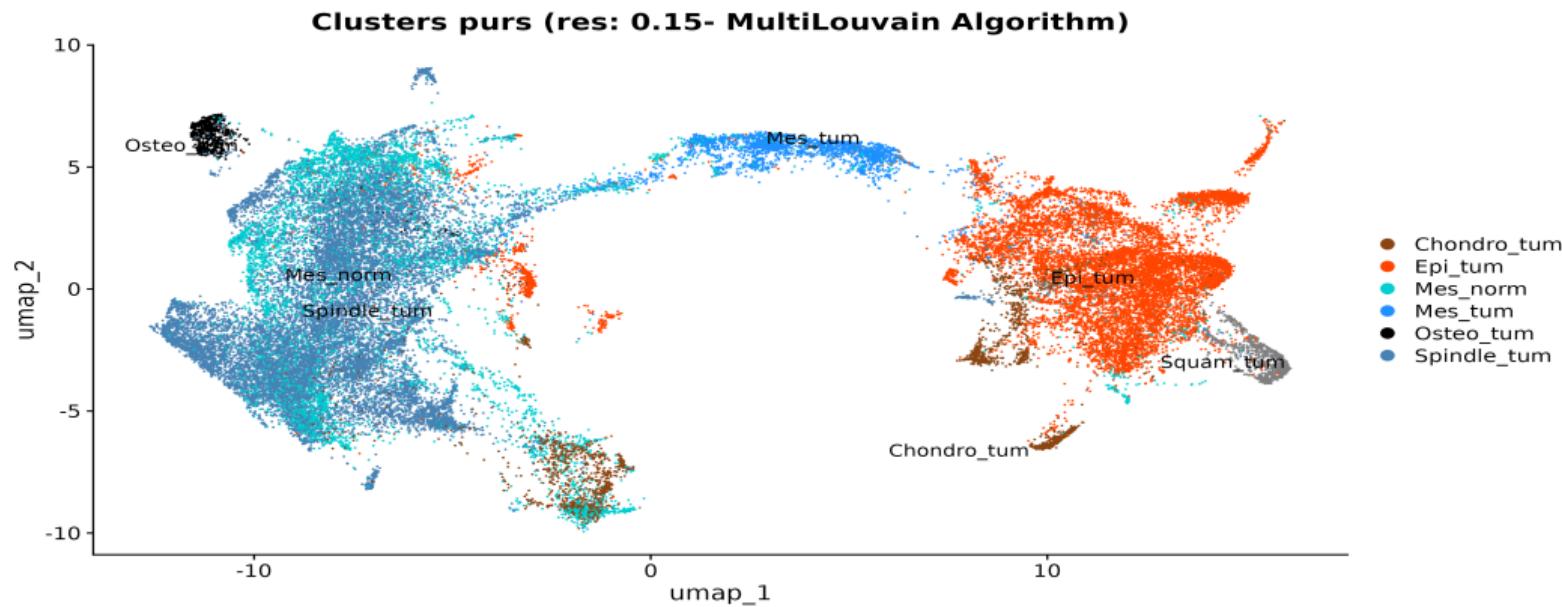


Figure 8: Multi-Louvain clustering after annotations intersection

# Clustering

## Points to discuss

- Mesenchymal tumor cluster close to Squamous tumor...
  - Strange, because phenotypically very different
  - Come from the same patient (MpBC8)
  - Harmony correction not enough ?
- Some clusters are still patient-specific more than cell-types-specific (e.g. Epi-tum)

# Feature Plots

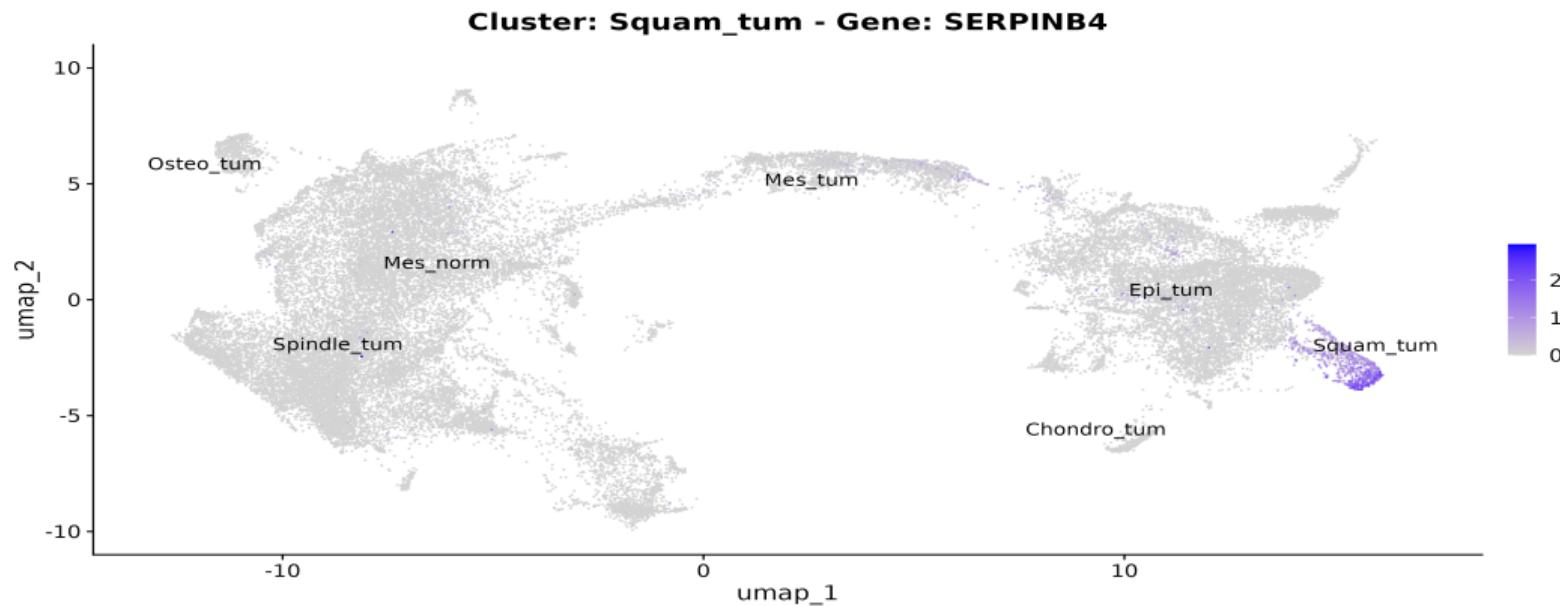


Figure 9: Feature Plot of SERPINB4 gene (Squamous cluster)

# Feature Plots

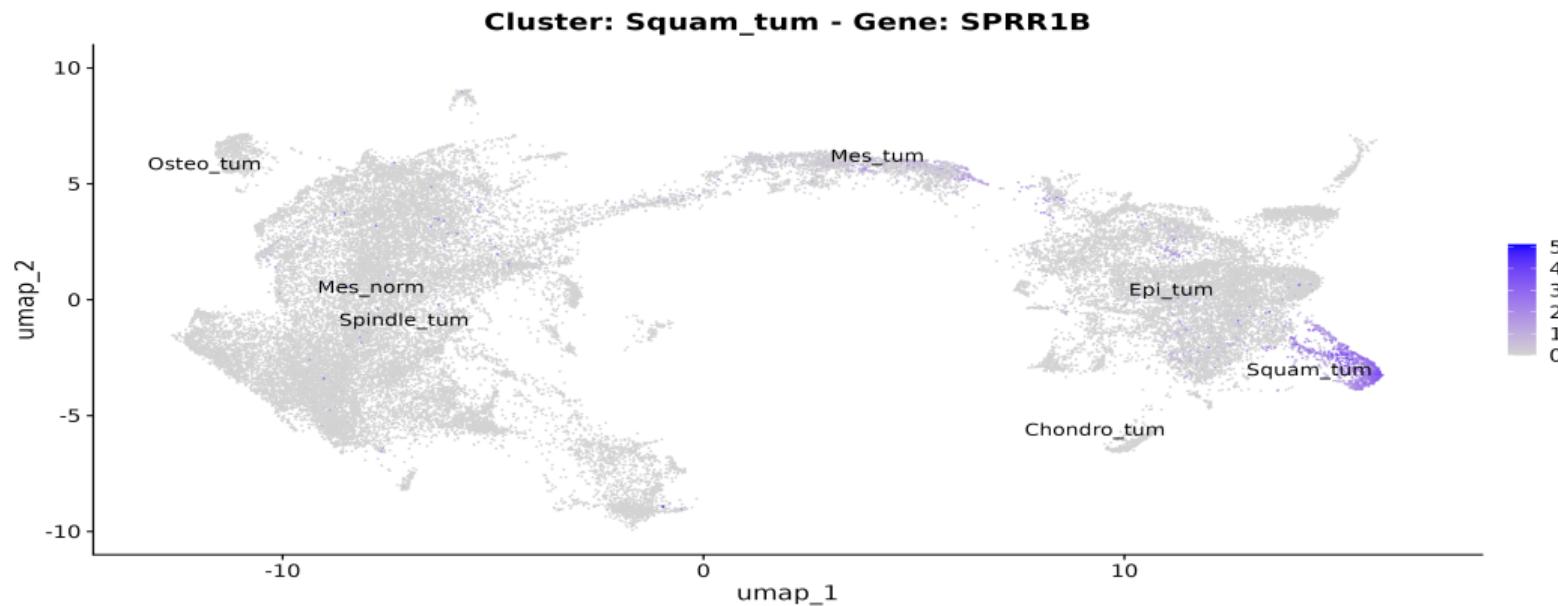


Figure 10: Feature Plot of SPRR1B gene (Squamous cluster)

# Feature Plots

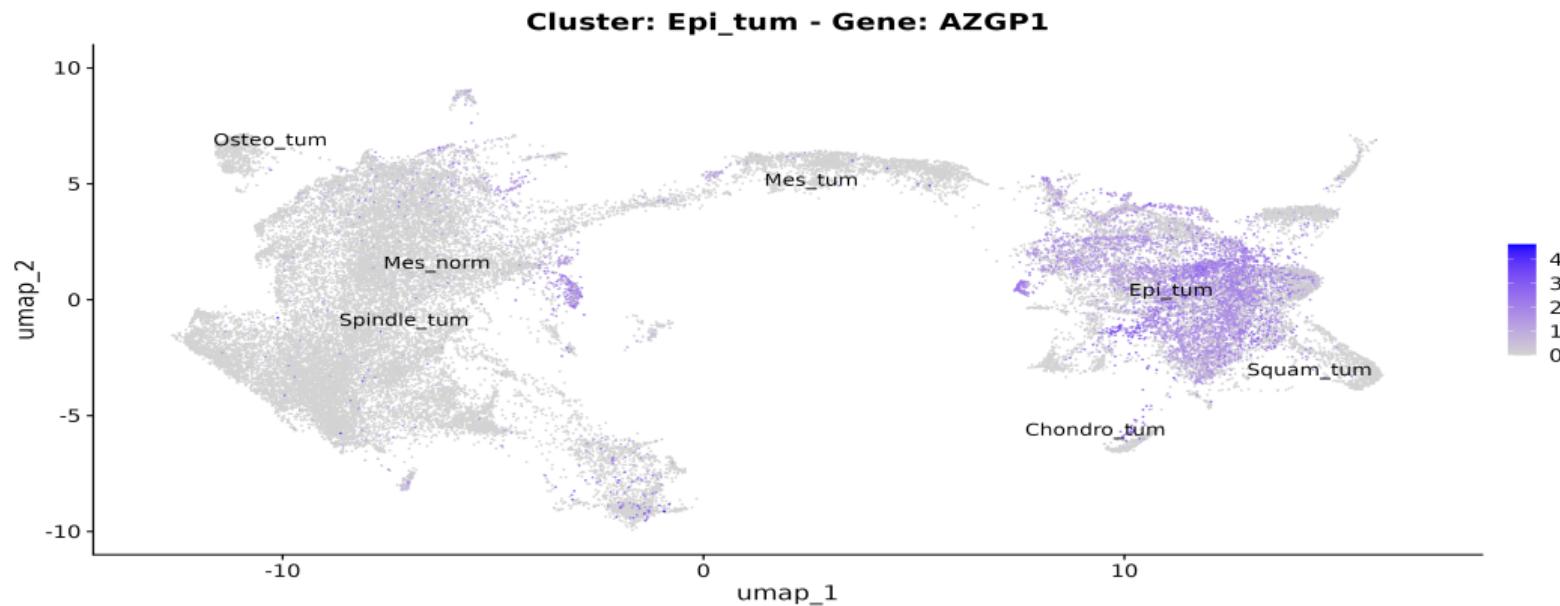


Figure 11: Feature Plot of AZGP1 gene (Epithelial cluster)

# Feature Plots

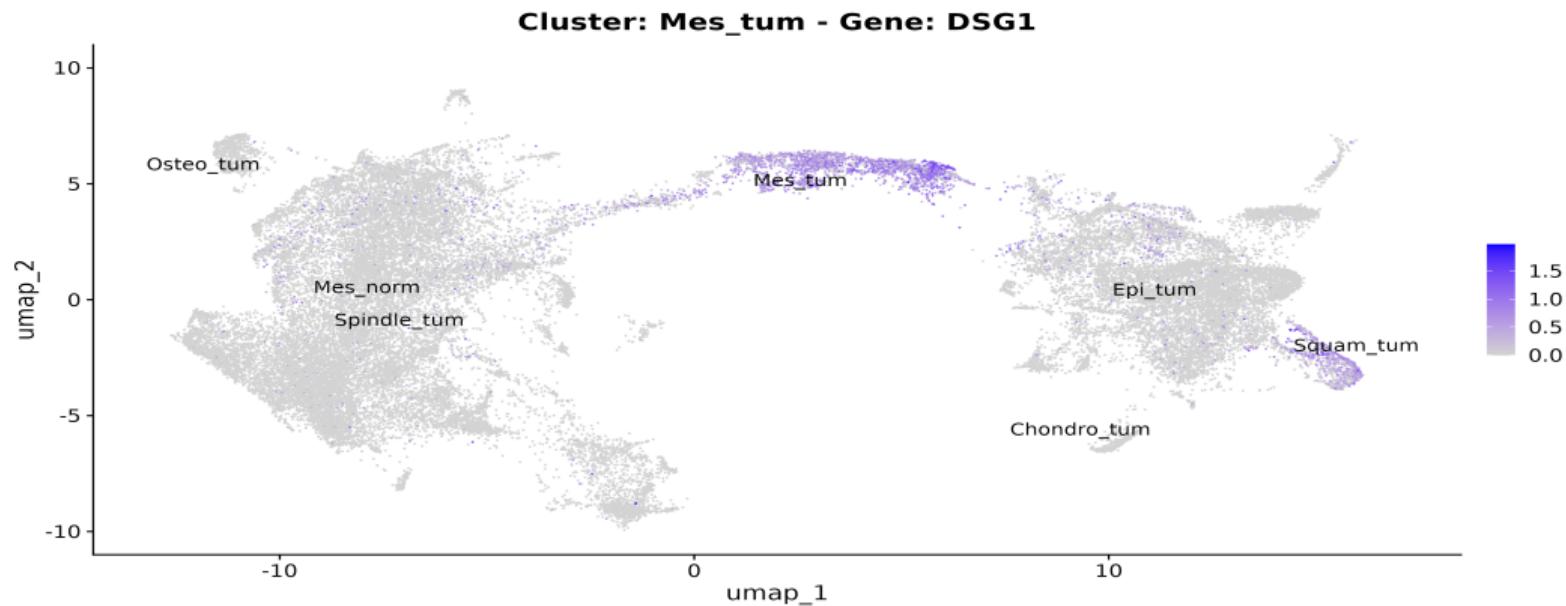


Figure 12: Feature Plot of DSG1 gene (Mesenchymal tumor cluster)

# Feature Plots

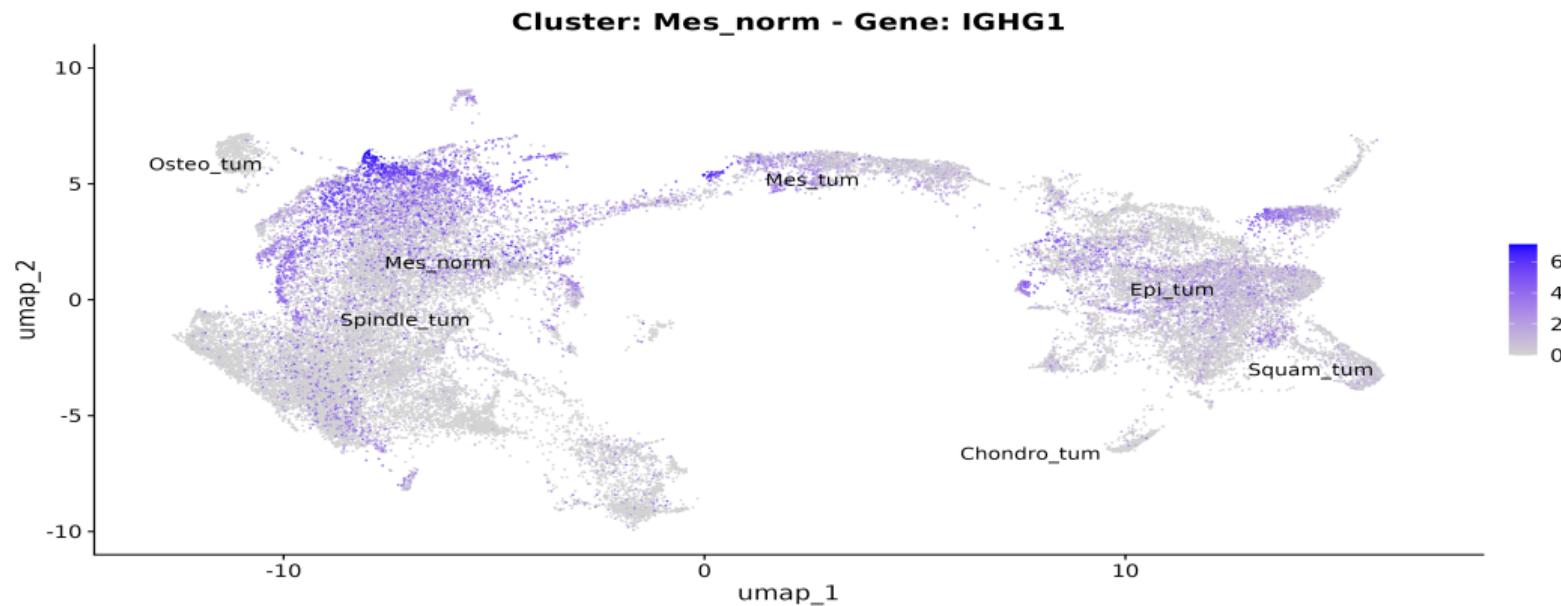


Figure 13: Feature Plot of IGHG1 gene (Mesenchymal normal cluster)

# Feature Plots

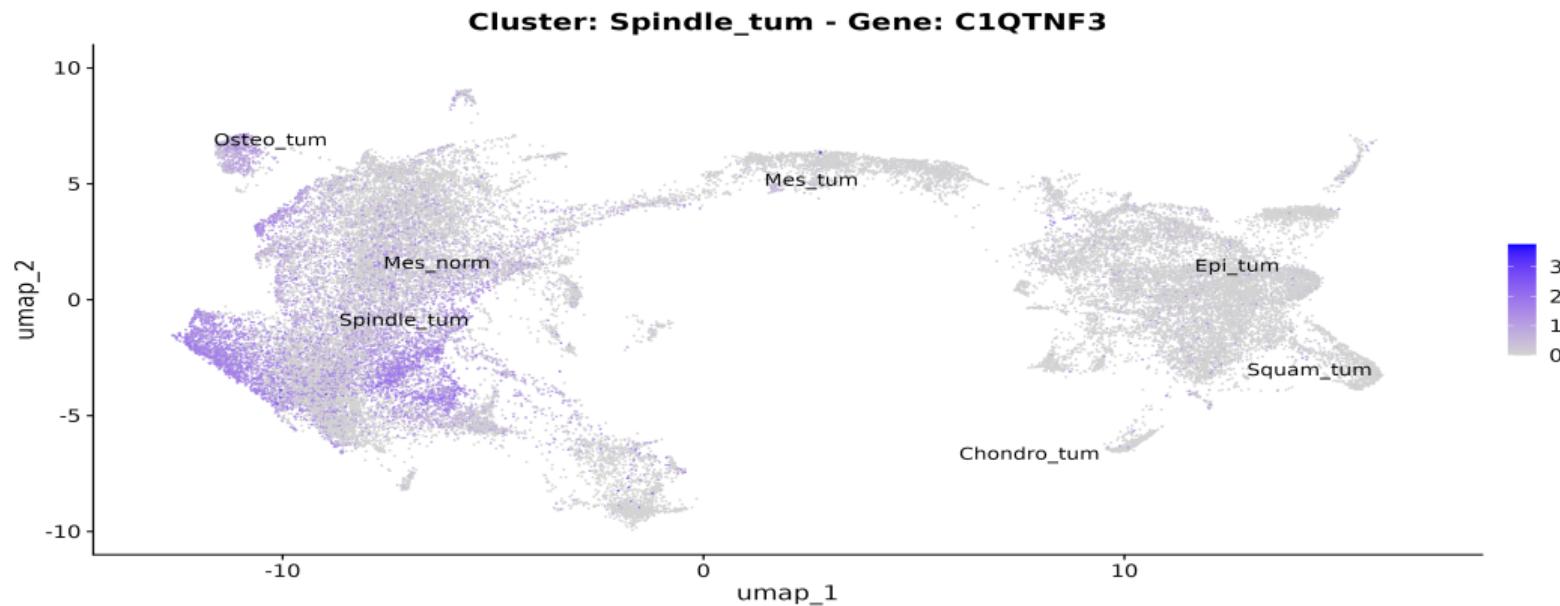


Figure 14: Feature Plot of C1QTNF3 gene (Spindle cluster)

# Feature Plots

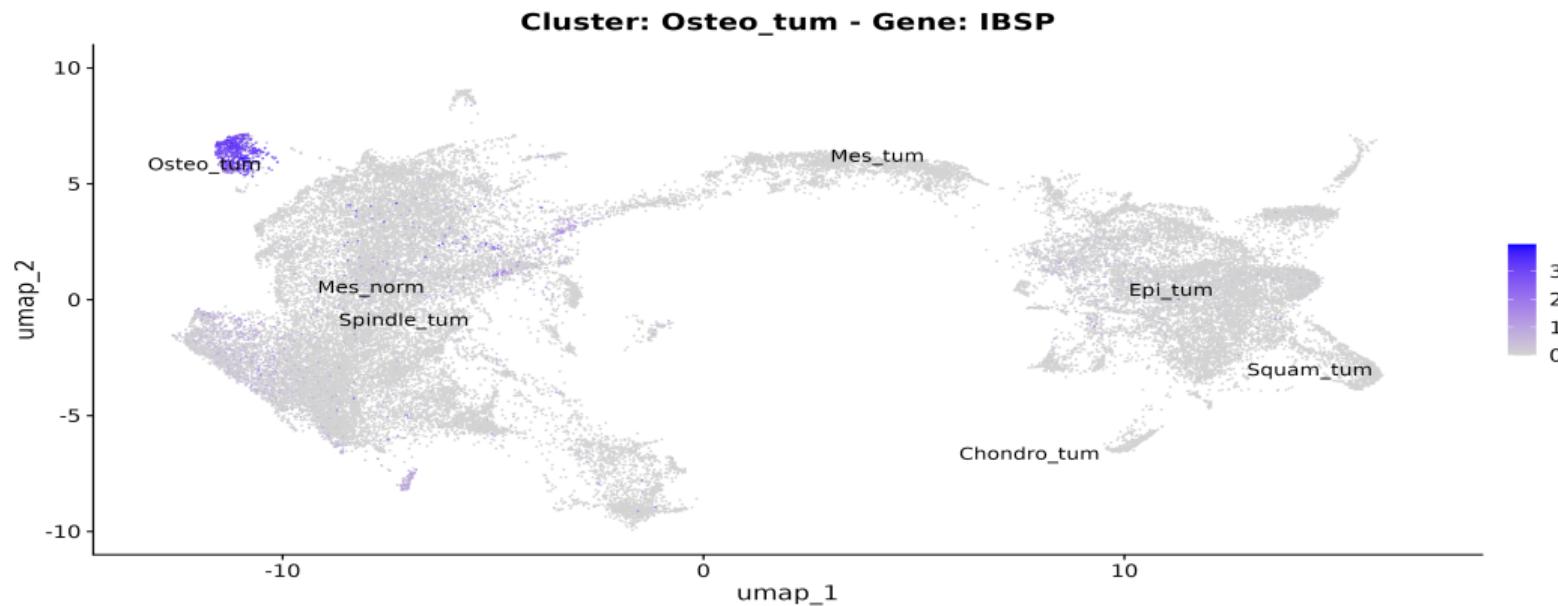


Figure 15: Feature Plot of IBSP gene (Osteosarcomatoid cluster)

# Feature Plots

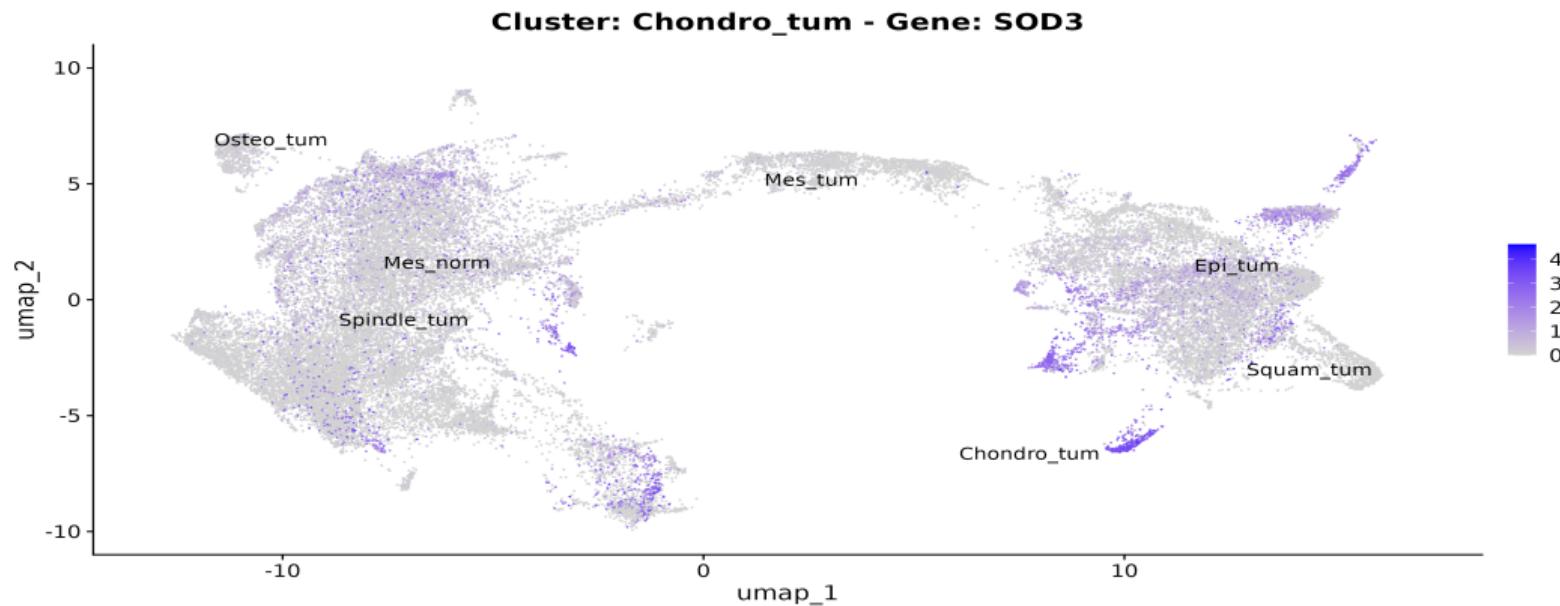


Figure 16: Feature Plot of SOD3 gene (Chondroid cluster)

## Dot Plot

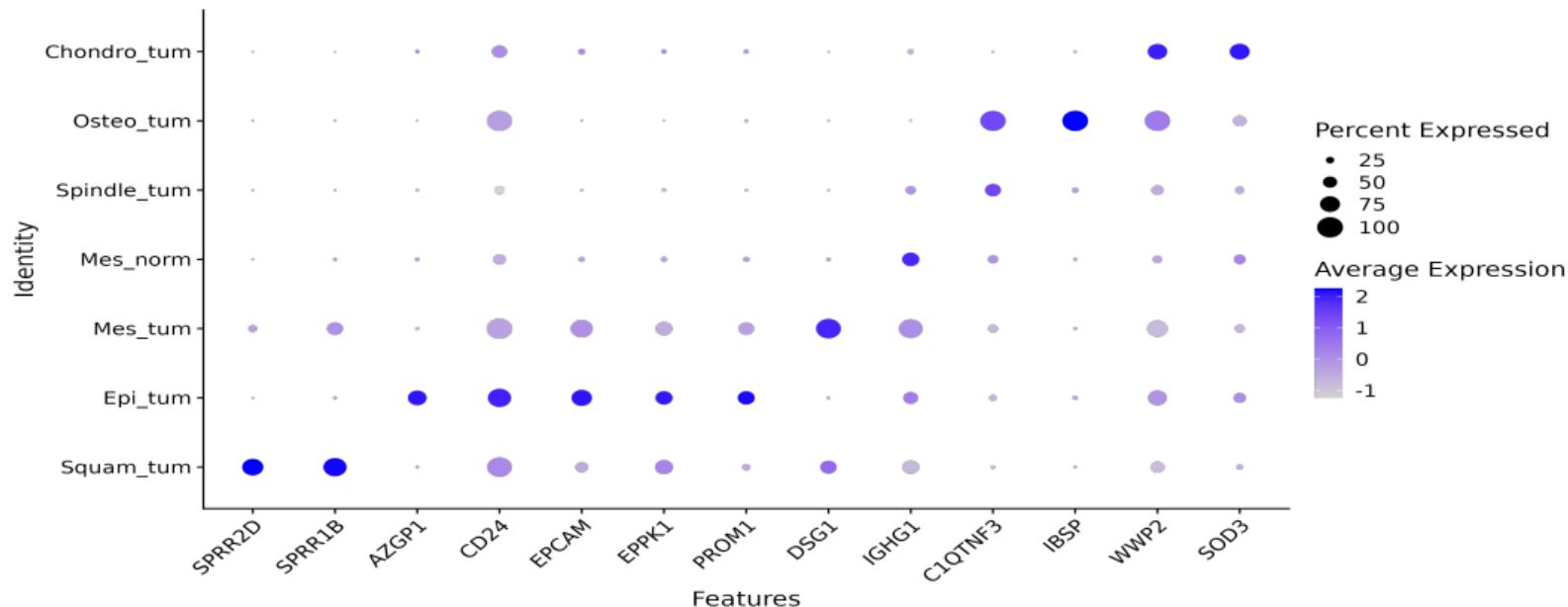


Figure 17: DotPlot of few cluster-specific markers

## 1 Introduction

## 2 Workflow

## 3 Results

## 4 Outlook

In the next months,

- Perform cell-deconvolution with sNuc-seq (single nuclei)-> Improve cluster resolutions and markers
- Realize cell-trajectory analysis -> Slingshot
- Search for Copy Number Alteration (CNA) in specific tumor cell types -> InferCNVplus
- Epigenomic and tumoral microenvironment analysis

**Thank you for listening !**

Tutor : Dr Pierre Martinez

jordan.dutel@lyon.unicancer.fr