Page de garde

Résumé

Table des matières

Table des figures

Liste des abréviations utilisées

TNBC = Triple Negative Breast Cancr

MpBC = Metaplasic Breast Carcinoma

CNA = Copy Number Alteration

CNV = Copy Number Variant

CLB = Centre Léon Bérard

H&E = Hematoxiline & Eosine

FFPE = Formalin-Fixed, Paraffin-Embedded

UMI = Unique Molecular Identifier

UMAP = Uniform Manifold Approximation and Projection

PCA = Principal Component Analysis

snRNA-seq = single nuclei RNA-sequencing

RCTD = Robust Cell Type Decomposition

Liste des logiciels et bases de données utilisés (avec numéros de versions)

Section Principale

# Introduction

## Contexte et état de l’art

Selon l’OMS (Organisation Mondiale de la Santé) en 2022, le cancer du sein était la première cause de cancer chez les femmes dans 157 pays sur 185 [https://www.who.int/fr/news-room/fact-sheets/detail/breast-cancer], et on considère qu’environ 1 femme sur 12 en sera diagnostiqué au cours de sa vie [https://www.who.int/fr/news-room/fact-sheets/detail/breast-cancer]. Aucun facteur de risque spécifique autre que le sexe et l’âge n’est connus [https://www.who.int/fr/news-room/fact-sheets/detail/breast-cancer], faisant donc de cette pathologie un enjeu majeure de santé publique dans les années à venir. Parmi les cancers du sein, on distingue les types TNBC (Triple Negative Breast Cancer), caractérisés par l’absence de certains types de récepteurs aux œstrogènes et progestérone. Les plus connus de ces récepteurs sont la protéine HER2 (Human Epidermal growth factor Receptor 2) [https://cancer.ca/fr/cancer-information/cancer-types/breast/what-is-breast-cancer/cancerous-tumours/triple-negative-breast-cancer#:~:text=Le%20cancer%20du%20sein%20triple,ses%20propres%20options%20de%20traitement.] et BRCA1/2 (Breast Cancer 1/2) [https://www.brcafrance.fr/cancer-hereditaire-sein-et-ovaire/]. De plus, de par l’absence de ces récepteurs pour les patientes atteintes de TNBC, cela rend le traitement par des thérapies ciblées extrêmement compliquées. Ainsi, les TNBC sont la plupart du temps des tumeurs agressives et de haut grade, c’est-à-dire, qu’elles présentent d’importante modifications cellulaires dysplasiques (dont le phénotype est très différent de celle des cellules normales) et disposant d’une capacité à se développer et à se propager très rapidement [https://cancer.ca/fr/cancer-information/what-is-cancer/stage-and-grade/grading]. Enfin, ces TNBC sont caractérisés par une variété impressionnante de sous-type, avec des phénotypes très hétérogènes [ref]. Parmi eux, nous retrouvons des sous-types rares, tel que les Basal-Like (Basal-Like Breast Cancer, BLBC) et les carcinomes du sein métaplasiques (Metaplastic Breast Cancer, MpBC) [ref]. Ces derniers, sont des cas rares et complexe de TNBC, encore aujourd’hui très mal compris et avec des marqueurs pour le diagnotique non caractérisés. Les patientes sont donc confrontées aujourd’hui à un manque important d’option thérapeutique, ce qui en fait une forme de cancer très agressive et avec une forte mortalité.

Peut-être expliquer le processus de cancérisation dans le cancer du sein (cancer épiT)

## Présentation des MpBC

Dans ce projet de recherche nous nous intéressons plus spécifiquement aux MpBC car les cellules composant ce sous-type disposent d’une capacité remarquable à se trans-différencier. Selon le CNRS, la transdifférenciation est « la conversion d'un type cellulaire entièrement différencié en un autre type » [https://www.insb.cnrs.fr/fr/cnrsinfo/est-ce-quune-division-change-les-etapes-de-reprogrammation-pour-une-cellule]. En effet, les MpBC sont caractérisés par une importante plasticité cellulaire aboutissant à des compartiments tissulaires phénotypiquement très distinct, au sein même de la tumeur. Les échantillons MpBC présentant au moins 2 compartiments cellulaires au sein de la tumeur sont dits « mixtes », en opposition aux MpBC « pures », ne présentant pas encore de cellules cancéreuses trans-différenciées.

Actuellement, on dénombre 5 types de trans-différenciation des MpBC à partir des cellules tumorales épithéliales : Malpighienne (ou Squammeuse), Mésenchymateuse, Spindle-like (ou Fusiforme), Chondroïde et Osteosarcomatoïde. Selon le type de la trans-différenciation, les cellules tumorales peuvent présenter un phénotype, par exemple, plus épithélial que les cellules cancéreuses initiales (trans-différenciation malpighienne), ou au contraire présenter un phénotype plus stromal (trans-différenciation mésenchymateuse). Certaines cellules, lors de ce processus peuvent également prendre une forme allongée et fusiforme, en forme d’épingle (trans-différenciation Spindle-like), ou bien créer abondamment de la matrice extracellulaire autour d’elles et s’enfermer dans une logette cartilagineuse (trans-différenciation chondroïde), voire même une matrice ostéoïde ressemblant à de l’os immature (trans-différenciation ostéosarcomatoïde). Cette large diversité de différenciation des MpBC n’est, encore aujourd’hui, pas comprise et les mécanismes restent inexpliqués. C’est à ce jour, une lacune importante dans notre compréhension de la plasticité des cancers. Or, en appréhendant l’origine de ces mécanismes, nous serions plus à même de développer des options de diagnostic moléculaire précis et d’élargir notre panel de cibles thérapeutiques pour ces cancers agressifs, qui pour l’instant n’ont aucuns traitements disponibles.

\*Peut être rajouté une photo/figure qui montre à quoi ressemble un MpBC

## Problématique(s)

Dans ce projet de recherche, afin de comprendre l’origine des mécanismes sous-tendant l’apparition de ces différents compartiments cellulaires au sein des MpBC, mon travail s’est articulé autour de 2 principales objectifs. Dans un premier temps, je me suis intéressé, grâce à des données de transcriptomique spatiale, au profil transcriptionnel de chaque sous-type cellulaire au sein des MpBC, afin de trouver des marqueurs phénotypiques spécifiques. Dans un second temps, je me suis focalisé sur les possibles causes génétiques en effectuant une analyse des altérations du nombre de copie (CNA) dans le génome de chaque sous-type cellulaire.

# Matériels et méthodes

## Echantillons MpBC

### Prélèvement/Collecte et fixation

#### Provenance des échantillons

Les échantillons MpBC proviennent de 16 patientes diagnostiquées pour un cancer du sein et ayant un suivi régulier au Centre Léon Bérard (CLB). Chacune des patientes présentent des MpBC mixtes (à verif si pas que certaines pures) avec des trans-différenciations de différents types. La collecte des biopsies s’est faite…… et la participation au projet de recherche s’est faite avec la non opposition des patientes et dans le cadre éthique du recueil et l’analyses de matériels biologiques imposé par l’hôpital CLB.

\*Plus d’infos, voire avec Pierre

#### FFPE fixation et coloration

Une fois les prélèvements cliniques réalisé pour le diagnostic, les biopsies ont suivies un protocole FFPE (Formalin-Fixed, Paraffin-Embedded) de fixation et d’inclusion en paraffine afin de conserver durablement l’échantillon et de faciliter sa réutilisation à des fins de recherche. Pour chaque patientes, une partie du tissu a été découpé au microtome puis placé sur une lame histologique afin de réaliser une coloration H&E (Hématoxiline et Eosine). Le reste de la tumeur a été conservé dans des congélateurs ?

#### 

### Séquençage

#### Référence de la techno, profondeur ?, kit use

Les tissus prélevés ont été ensuite placé sur lame de séquençage de la société privé 10X Genomics. Pour notre projet nous avons utilisé la technologie Visium, permettant d’apporter une information spatiale via des spots de cellules (1 spot représentant environ 10 cellules), en plus de nos données transcriptomiques, pour chaque échantillon. Le séquençage s’est fait sur un séquenceur de type ???, avec le kit de séquençage ??? et à une profondeur de ???. Parmi les 16 échantillons, 1 seule présente des statistiques de contrôle qualité non optimales (MpBC13). Nous l’avons donc exclu des analyses.

### Annotation phénotypique des spots

#### Clustering k-means (Loupe Browser) & expert anatomopathologiste

Afin de disposer de l’information spatiale, nous avons annoté chacun des spots pour chaque échantillon avec les sous-types cellulaires observés sur la lame avec la coloration H&E. Pour cela nous avons combiné les informations du sous-type cellulaire fournies par 2 méthodes différentes. Une première annotation des spots a été réalisé par le logiciel « Loupe Browser » de la société 10X Genomic, permettant de clustériser les spots présentant des profils transcriptomiques similaires, via un clustering k-means. Puis dans un deuxième temps, à l’aide de ces clusters trouvés, les spots ont été individuellement vérifiés et manuellement annoté par un expert anatomopathologiste de la plateforme histologique du CLB. Cela nous garantit donc une bonne confiance quant à l’identité des cellules séquencés sur la lame, pour chaque patientes.

## Software et packages

### RStudio et language R

#### Version du logiciel

### Analyse des marqueurs phénotypiques

#### Loupe Browser

##### Version du logiciel

#### Seurat

##### Version du package

##### Contrôle qualité et filtrage des spots

###### UMI & Features count,

###### % gènes mitochondriaux

###### …

##### Normalisation et Scaling

###### Nombre de gène sélectionné

##### Déterminations des paramètres UMAP de façon empirique

##### Clustering

###### Type de l’algorithme use (=Louvain)

#### Harmony

##### Version du package

##### Paramètres de la correction batch-effect

### Analyse des marqueurs génotypiques

#### InferCNVPlus

##### Constitution du groupe de cellule de références

##### Normalisation

##### Tests statistiques

# Résultats

## Analyse des marqueurs phénotypiques

### Réduction PCA harmony

#### Figure 1 : Réduction PCA après correction Harmony

### Projection UMAP (après Clustering)

#### Figure 2 : Projection UMAP (après épuration des spots contradictoires entre annotations expert et clustering seurat)

### Feature Plot

#### Figure 3 : Illumination de la UMAP selon les gènes marqueurs

##### Rajouter sur les figures, des \* avec les termes GO (fonctions) pour les gènes représentés

### Dot Plot

#### Figure 4 : Features selon types cellulaires

#### Figure 5 : Features selon les patients

## Analyse des marqueurs génotypiques

### Visualisation des différences de score CNA entre compartiment

#### Figure 6 : Shift Plot

#### Figure 7 : Effect\_size en fonction p\_val

##### Modifier la figure pour la mettre sous forme de VolcanoPlot (abscisse = Size Effect, ordo = Log10(Pval\_adj), Différence sans val absolue)

#### Figure 8 : P\_val barplot par bras

#### Figure 9 : Heatmap Plot

#### Figure 10 : Correlation Plot

# Discussions

## Limites de la technologie Visium

### Risque d’hétérogénéité cellulaire intra-spot

#### Séparation imparfaite des clusters (UMAP)

#### Trop faible spécificité des marqueurs génétiques pour les sous-types cellulaires

##### Marqueurs pour le moment trop patient-spécifique plutôt que cluster-spécifique

#### Dilution/Perturbation du signal CNA

##### Quelques cellules non cancéreuses au sein des spots annotés cancéreux

#### Difficulté pour caractériser des marqueurs génétiques et CNA spécifiques aux sous-types

## Perspectives du projet

### Analyse approfondie via snRNA-seq

#### Création d’un atlas MpBC spécifique

### Microdissection des sous-types spécifiques

#### Déterminants génétiques

##### Mutations drivers de chaque sous-type

##### Mutations clonales à l’origine de tous les MpBC

#### Déterminants non génétiques

##### Epigénétique & méthylome

##### Microenvironnement & interaction ligand-récepteur

### Caractérisation des marqueurs et voies de signalisation pour le diagnostique

# Conclusions

Références bibliographiques

Annexes