*Cahier des charges*

Partie 2

**Analyse des enzymes de méthylation dans le cancer du sein métastatique**

Ekaterina Flin

17/02/2024

Table des matières

[Evaluer le risque de métastases dans les échantillons du cancer du sein 1](#_Toc159073479)

[Signature GEC 1](#_Toc159073480)

[Identifier les enzymes de méthylation activés dans le cancer métastatique 4](#_Toc159073481)

[Dans les données d’expression totale (8 datasets) 4](#_Toc159073482)

[Dans les données « single cell » 5](#_Toc159073483)

# Evaluer le risque de métastases dans les échantillons du cancer du sein

## Signature GEC

Cette partie a été déjà réalisée. La description et les figures sont données pour votre information.

Le risque de métastases peut être mesuré dans chaque échantillon du cancer du sein à l’aide de la signature moléculaire déjà connues, appelée GEC (Gene Expression Classifier) et basée sur l’expression de 5 gènes : DNMT3B, MCM10, EXO1, CENPF, CENPE.

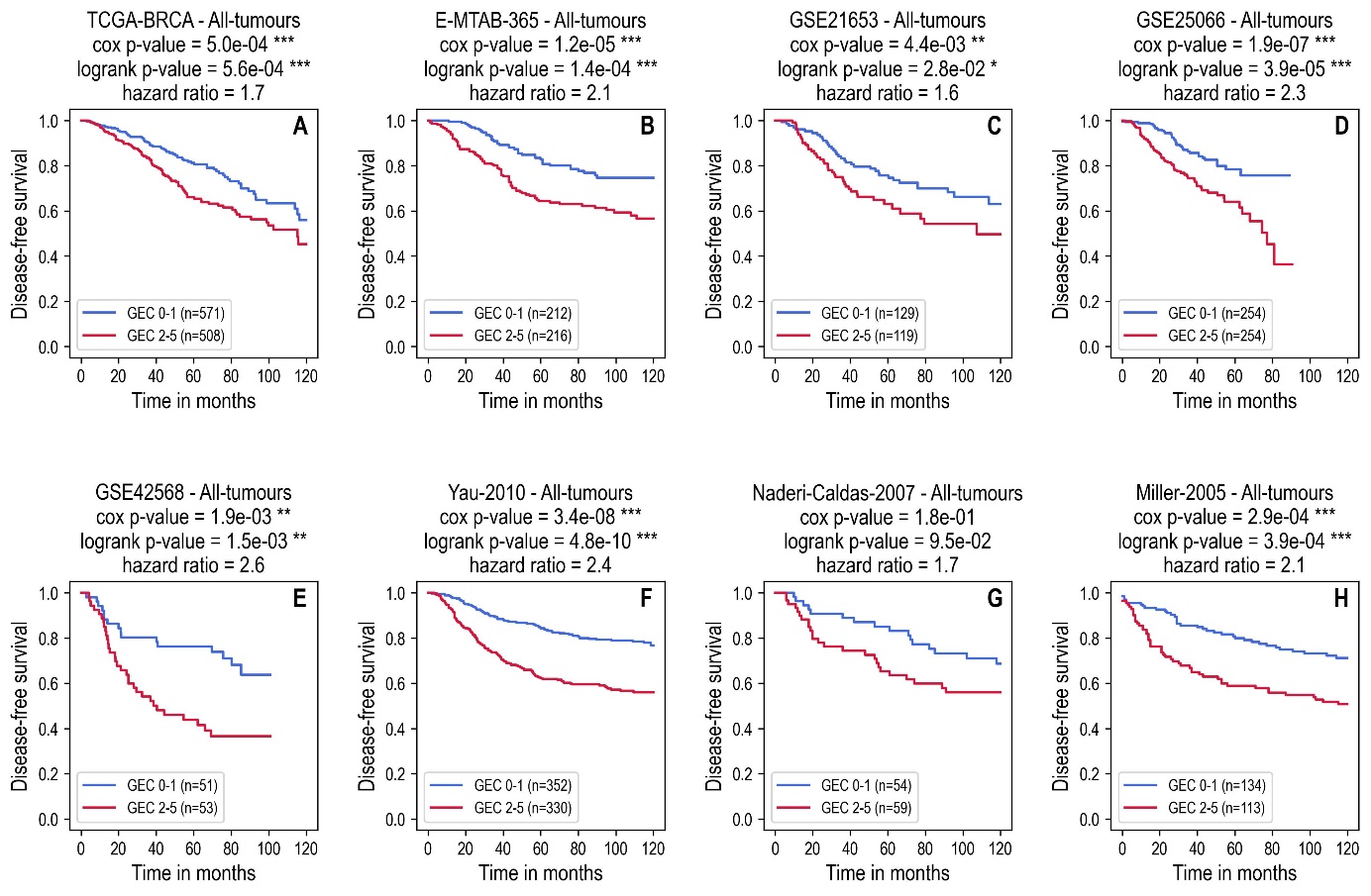
La signature GEC a été publiée dans l’article de *Jacquet et al. 2023 « Aberrant activation of five embryonic stem cell-specific genes robustly predicts a high risk of relapse in breast cancers »* (Jacquet et al. 2023 PMID 37592220).

Dans chaque échantillon du cancer du sein, il est possible de mesurer le nombre de gènes activés parmi ces 5 gènes (score GEC). Le score GEC peut varier de 0 (aucun gène activé, risque faible) à 5 (tous les gènes activés, risque fort).

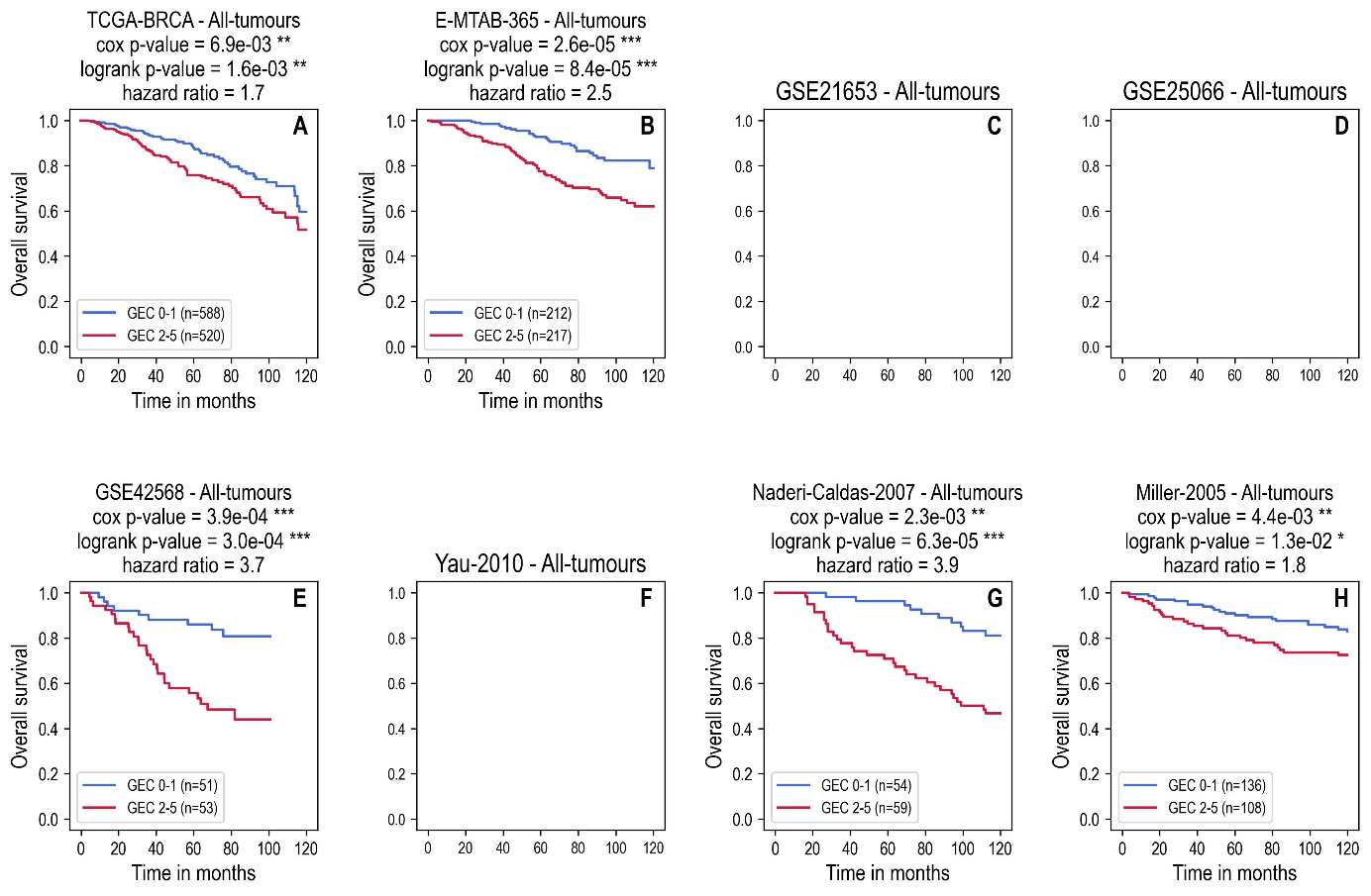
Le score GEC est déjà connus dans tous les échantillons de tous les datasets disponibles.

Le score GEC est significativement associé à la survie globale des patients (dans les datasets où les données de survie globale OS sont disponibles) ainsi qu’au risque de rechute de la maladie (données de survie DFS disponibles dans tous les datasets).

**Survie sans rechute**

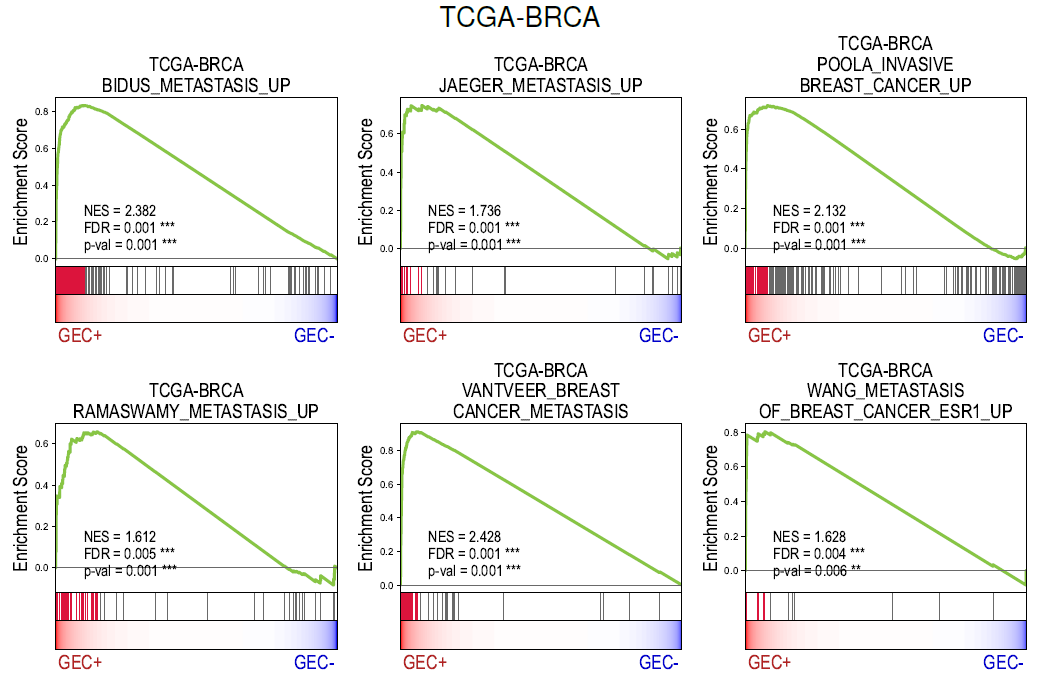


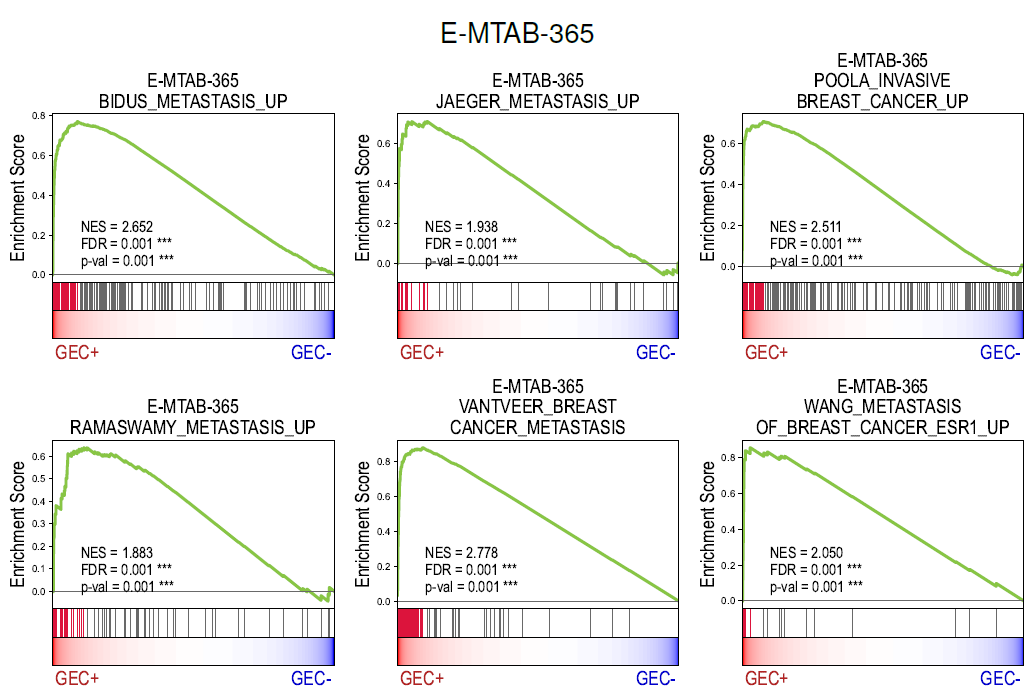
**Survie globale**



*case vide = données de survie OS absentes*

Le score GEC est également associé au risque de métastases (Jacquet et al. 2023 PMID 37592220). Notamment, les tumeurs GEC+ sont significativement enrichies en signatures de métastases (du cancer du sein et d’autres cancers solides) connues et publiées dans la littérature. Ci-dessous quelques résultats de l’analyse GSEA dans les datasets TCGA-BRCA et E-MTAB-365. Par exemple, les tumeurs GEC+ sont enrichies en signature de van’t Veer (van 't Veer et al. 2002, *Nature*, PMID 11823860), intégrée dans le test existant MammaPrint.





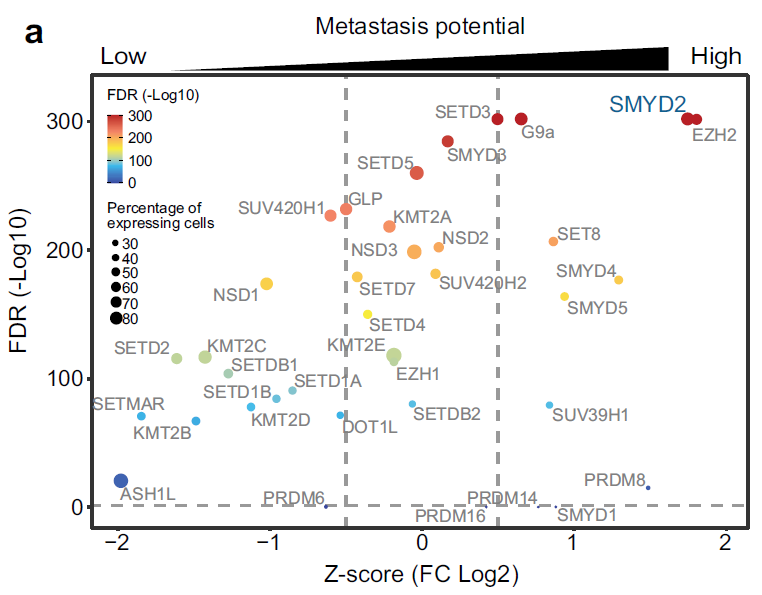
Conclusion : La signature GEC peut être utilisée comme une signature de risque de métastases dans le cancer du sein.

# Identifier les enzymes de méthylation activés dans le cancer métastatique

## Dans les données d’expression totale (8 datasets)

Réaliser une analyse différentielle entre les groupes GEC+ versus GEC- pour les enzymes de méthylations comme dans la première partie du travail (comme c’était pour les groupes cancer vs normal ou sous-types moléculaires).

Présenter les résultats dans un volcanoplot similaire à celui montré ci-dessous. Il y aura 8 figures, une par dataset.



*Figure issue d’un article publié de l’équipe de Nicolas Reynoird sur les KMT (Casanova et al. 2024 PMID 38296970). Les autres enzymes (KDM, KMB) n’avaient pas été inclus. Il serait intéressant de les analyser aussi.*

## Dans les données « single cell »

S’il reste du temps, faire la même chose dans les données single cell (1 nouveau dataset GSE118389 à ajouter, cancer du sein triple négatif / basal-like) dans les cellules tumorales uniquement.