

# Analyse différentielle des régions méthylées

## Revue de littérature

Jakobi Milan

TIMC-Imag

Vendredi 15 mars 2019



# Table des matières

identification  
DMRs

Jakobi Milan

DMRcate

Données  
Méthode

blabla

DMRcate

Données

Méthode

blabla

- ▶ 450k microarray avec 25% des probes sur des zones intercalaires.
- ▶ On utilise les M-values (  $= \log(\beta)$  )

- ▶ 450k microarray avec 25% des probes sur des zones intercalaires.
- ▶ On utilise les M-values (  $= \log(\beta)$  )
- ▶ Méthode extensible à toutes données génomiques : RRBS, WGBS...

# Présentation de la méthode

identification  
DMRs

Jakobi Milan

DMRcate

Données  
**Méthode**

blabla

3 types d'analyse proposées :

3 types d'analyse proposées :

1. Analyse entre deux groupes ( Traitement vs Contrôle)
2. Analyse de contraste.
3. Analyse de variabilité ( On identifie alors les VMRs).

# Pseudo-algorithme

identification  
DMRs

Jakobi Milan

Après avoir choisi le type d'analyse, les étapes de la procédure seront les suivantes :

DMRcate

Données  
Méthode

blabla

# Pseudo-algorithme

identification  
DMRs

Jakobi Milan

DMRcate

Données  
Méthode

blabla

Après avoir choisi le type d'analyse, les étapes de la procédure seront les suivantes :

1. On calcule  $Y_i$  nos statistiques de test.



Après avoir choisi le type d'analyse, les étapes de la procédure seront les suivantes :

1. On calcule  $Y_i$  nos statistiques de test.
2. On estime la distribution de nos statistiques de test  $Y_i$  par noyau gaussien ( par zones de taille  $\lambda$ ).

Après avoir choisi le type d'analyse, les étapes de la procédure seront les suivantes :

1. On calcule  $Y_i$  nos statistiques de test.
2. On estime la distribution de nos statistiques de test  $Y_i$  par noyau gaussien ( par zones de taille  $\lambda$ ).
3. On modélise, par méthode de Satterthwaite, nos statistiques de test "smoothées".

Après avoir choisi le type d'analyse, les étapes de la procédure seront les suivantes :

1. On calcule  $Y_i$  nos statistiques de test.
2. On estime la distribution de nos statistiques de test  $Y_i$  par noyau gaussien ( par zones de taille  $\lambda$ ).
3. On modélise, par méthode de Satterthwaite, nos statistiques de test "smoothées".
4. On en déduit les pvaleurs

Après avoir choisi le type d'analyse, les étapes de la procédure seront les suivantes :

1. On calcule  $Y_i$  nos statistiques de test.
2. On estime la distribution de nos statistiques de test  $Y_i$  par noyau gaussien ( par zones de taille  $\lambda$ ).
3. On modélise, par méthode de Satterthwaite, nos statistiques de test "smoothées".
4. On en déduit les pvaleurs
5. On fixe un seuil à partir duquel on exclue les variables dont la pvaleur est trop forte

Après avoir choisi le type d'analyse, les étapes de la procédure seront les suivantes :

1. On calcule  $Y_i$  nos statistiques de test.
2. On estime la distribution de nos statistiques de test  $Y_i$  par noyau gaussien ( par zones de taille  $\lambda$ ).
3. On modélise, par méthode de Satterthwaite, nos statistiques de test "smoothées".
4. On en déduit les pvaleurs
5. On fixe un seuil à partir duquel on exclue les variables dont la pvaleur est trop forte
6. On construit nos DMRs/ZMRs finales en regroupant les CpG sites qui sont au plus à  $\lambda$  nucléotides

Selon le type d'analyse, nos statistiques de test  $Y_i$  sont différentes : Pour les analyses entre deux groupes ou analyse de contraste, on a :

$$Y_i = \hat{t}^2$$

