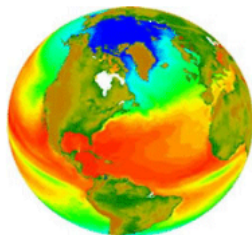


Contexte biologique



Impact de l'environnement sur la faune et la flore

→ Ressources alimentaires : volaille

Capacités d'adaptation des poulets : chaleur et alimentation

Quelle part des phénomènes épigénétiques dans la variabilité phénotypique observée chez les espèces d'élevage ?

Phénomènes épigénétiques

- Intervenant dans la régulation de l'expression des gènes
- Sensibles à l'environnement

→ Pas de modification sur la séquence d'ADN, transmis par mitose ou méiose

Méthylation de l'ADN : modification chimique des bases cytosines dans les dinucléotides CG



⇒ Méthylome

Projet ChickStress

Programme ANR

Mécanismes d'adaptation en réponse à un stress chaleur et alimentaire

Mesures de performances et physiologiques
+ données transcriptomiques et **épigénétiques**

Dispositif :

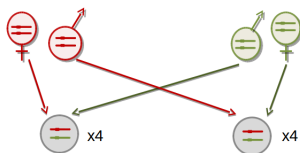
- Poule pondeuse : 2 lignées, 3 conditions, 48 animaux
- Condition contrôle et conditions de stress :
 - Alimentaire : aliments moins énergétiques
 - Chaleur : +10 degrés Celsius
- Tissus

⇒ Données ChickStress en préparation

Projet EpiBird

8 embryons de poulet analysés par des expériences de séquençage à haut-débit

- Croisement réciproque de 2 lignées : lignée R^- et lignée 6
→ Embryons mâles et femelles par sens de croisement



- Données initiales : transcriptome, génome, **méthylome**

⇒ Pipeline :

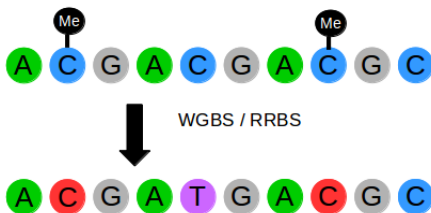
Données initiales de séquençage → taux de méthylation

Détection de différentiel

Techniques de séquençage haut-débit

EpiBird : WGBS (*Whole Genome Bisulfite Sequencing*)

→ 161 millions de lectures x8 = 80Go

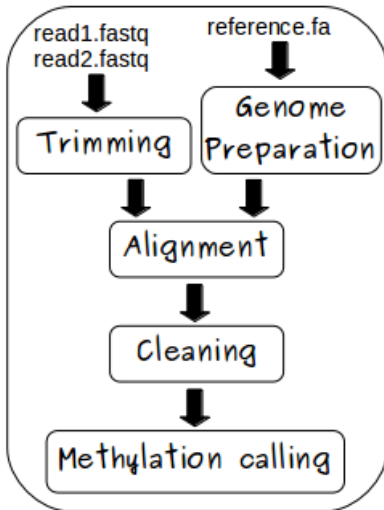


ChickStress : RRBS (*Reduced Representation Bisulfite Sequencing*)

+ Double digestion enzymatique + *sizing*

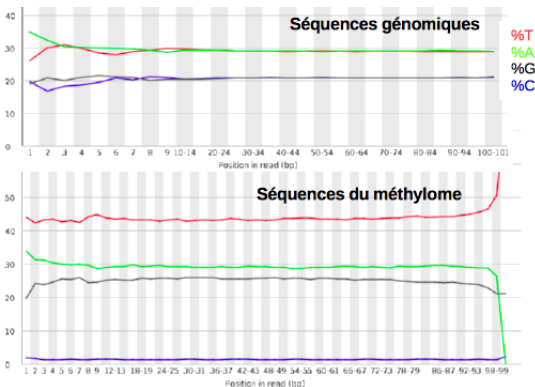
→ Détection de régions du génome riches en GC

Développement



Données initiales : Trimming

Éliminer les parties non informatives : adaptateurs moléculaires
Logiciel *Trim Galore!* (Krueger, 2012)

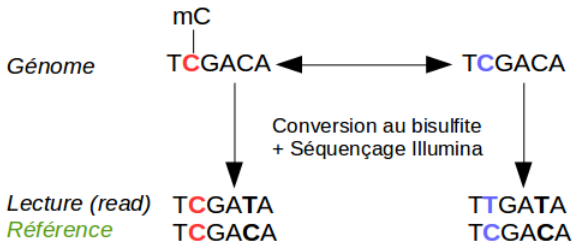


⇒ Majorité de C non méthylés présents dans le génome

Alignement

Alignement des lectures sur un génome de référence :
sous-séquence similaire

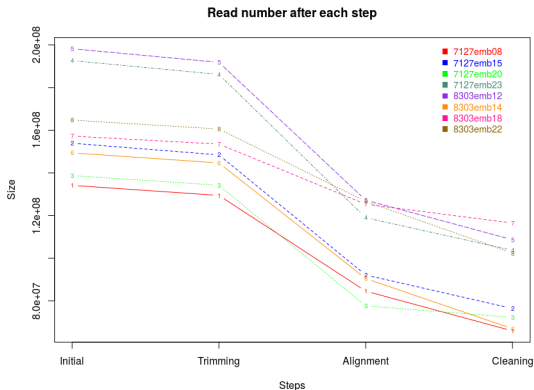
Logiciel *Bismark* (Krueger et al, 2012)



⇒ Reads alignés de façon unique avec le meilleur score
67.3% lectures alignées en moyenne

Cleaning

Éliminer les produits de certains biais expérimentaux
Outil *Samtools rmdup* (Li et al, 2009)



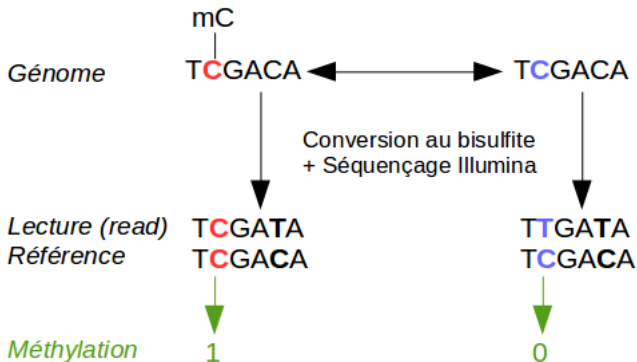
⇒ soit 45% de lectures éliminées

Taux de méthylation

Pour chaque C en contexte CG : nombre de C et de T

$$\frac{C}{C+T}$$

Package MethyKit (Akalin et al, 2012)



Taux de méthylation

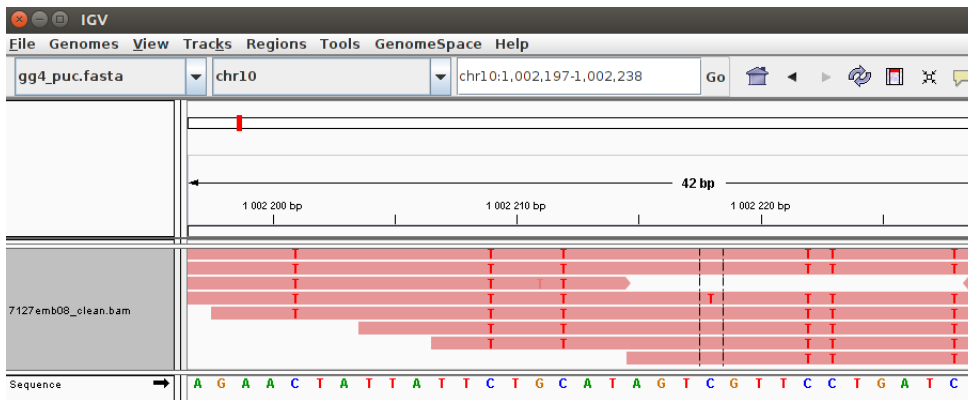
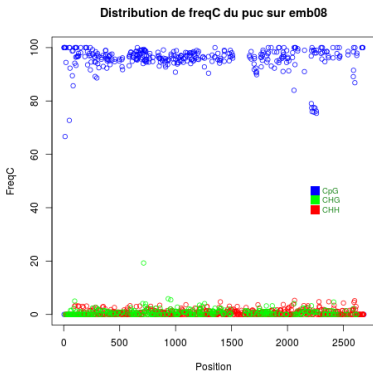


Figure: Fenêtre IGV du chromosome 10, position 1002217, embryon 08 : 6 C + 1 T soit 86% de méthylation

Contrôles internes

Vérification de la fiabilité du traitement au bisulfite :

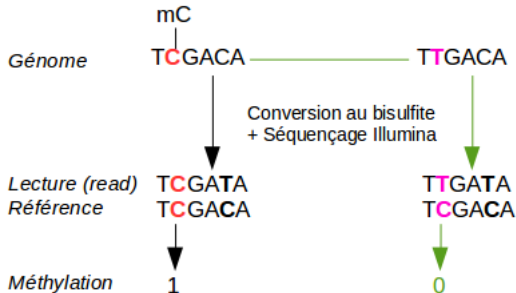
- Taux de conversion de C hors dinucléotides CG : taux = 99.33%
- Contrôle interne sur le pUC selon les 3 contextes possibles de C :



SNP (*Single Nucleotide Polymorphism*)

Élimination de SNP C/T : faussent le calcul de méthylation

Logiciel GATK (McKenna et al, 2010)



⇒ 3% des positions éliminées

Méthylation descriptive

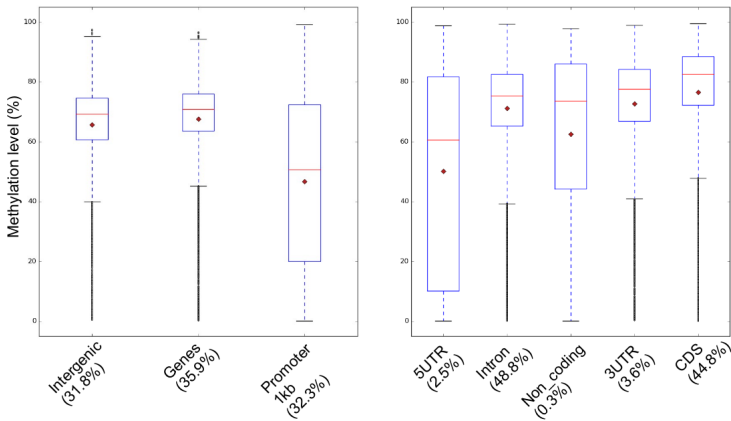
- Référence : *Gallus gallus* version 4
- 10'716'389 CG dans le génome de référence
- 8'745'437 positions de CG détectées communes (couverture = **1**)
→ 82%
- 1'416'683 positions de CG détectées communes (couverture = **5**)
→ 13%
- Moyenne de méthylation sur tout le génome (après filtre) : 67.5%
- Environ 0.7% de C hors CpG (issu de l'alignement)

Méthylation selon des intervalles génomiques

Intervalles de positions génomiques : catégories fonctionnelles du génome ou *features*

Projection sur le génome

Score de méthylation par intervalle



Lien méthylation / expression

Données de méthylation : Projection pour les promoteurs et taux de méthylation pour un embryon

Données d'expression (RNA-seq) : Comptage brut par exons

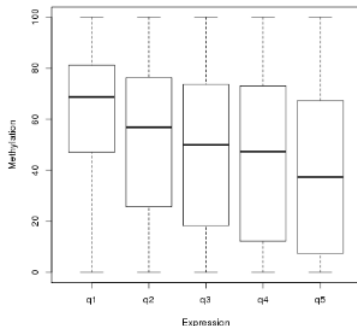


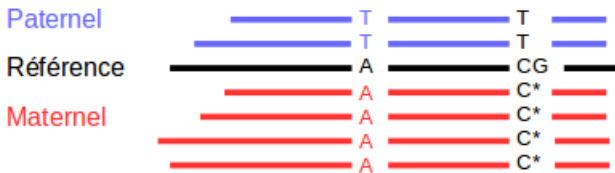
Figure: Distribution de la méthylation dans le promoteur selon les intervalles d'expression

Méthylation allèle spécifique

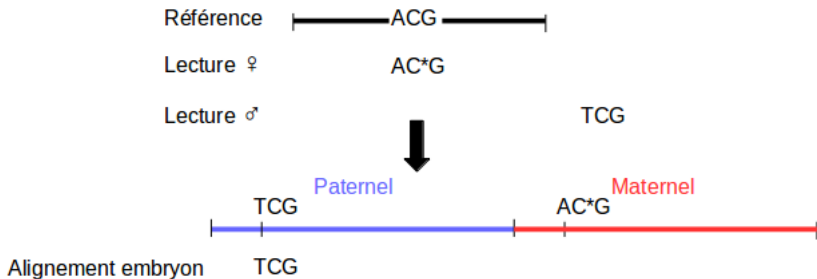
Plusieurs stratégies :

Éliminer les biais d'alignement de l'allèle de référence

- Alignement sur les haplotypes parentaux



Méthylation allèle spécifique



- Remplacer les SNP connus par des "N"

⇒ Lien avec l'expression allèle spécifique

Différentiel

DMC : C différentiellement méthylés

DMR : régions différentiellement méthylées

EpiBird : Sexe de l'embryon / sens de croisement

ChickStress : Stress environnementaux et contrôle

Logiciels utilisés :

- MethylKit : 25% de différence de méthylation, $qvalue = 0.01$
 - 250 DMC selon le sexe de l'embryon
 - 6628 DMC selon le sens de croisement
- eDMR : 20% de différence de méthylation, $qvalue = 0.05$,
3 CpG, 1 DMC
 - 10 DMR selon le sexe de l'embryon
 - 402 DMR selon le sens de croisement

Différentiel

→ Comment justifier de la pertinence des résultats statistiquement ? (Gaëlle Lefort, Nathalie Villa-Vialaneix)

- Normalisation et filtre
- DMC :
 - Test de Fisher ou régression logistique (methylKit)
 - Modèle bêta-binomial (DSS)
- DMR :
 - Fenêtre glissante (methylKit)
 - Modèle normal bimodal (eDMR)
 - À partir des DMC (DSS)

→ DSS (*Dispersion Shrinkage of Sequencing data*)

Conclusion

Implémentation :

- Étapes du pipeline : Jflow (WMS ou gestionnaire de workflow) (Mariette et al, 2015)
- Pipeline bioinformatique : Céline Noirot (MIAT)
Pipeline statistique : Gaëlle Lefort (MIAT)
→ Mise en ligne sur GitHub (groupe FAANG)

Perspectives :

- Régions génomiques d'intérêt (TSS, hotspots...)
- Différentiel de méthylation / Différentiel d'expression
- *Gallus gallus* version 5

Fin

Merci de votre attention !

Traitement au bisulfite (1)

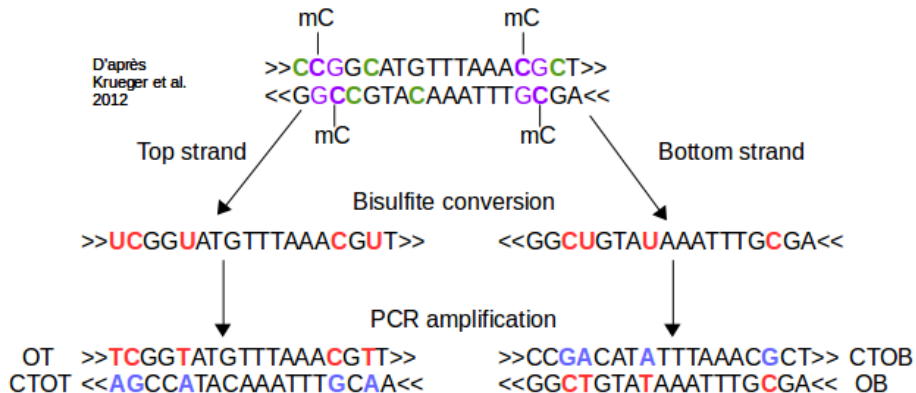


Figure: D'après Krueger et al, 2012 - Principe du traitement au bisulfite

Alignement (2)

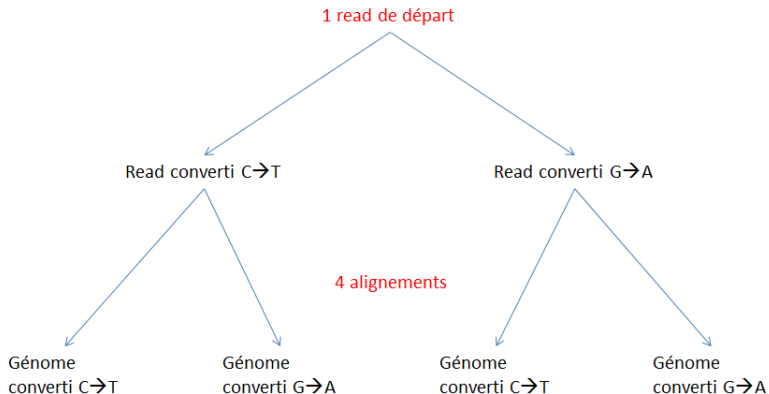


Figure: Principe de conversion multiple pour l'alignement par Bismark