

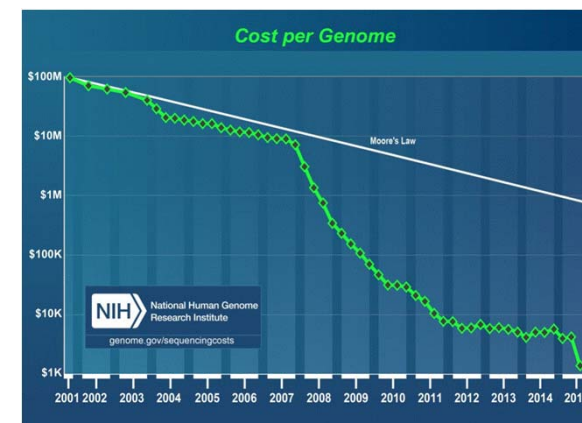
Thématiques émergentes développées dans le projet « Séquençage Occitanie Innovation » (SeqOccln) coordonné par l'INRAE

Clémence GENTHON

Le contexte du projet

Un domaine très évolutif

- Un projet porté par **GeT** (Génome et Transcriptome, US1426), plateforme nationale stratégique INRAE et membre du PIA **France Génomique** & par **Bioinfo Genotoul**, plateforme nationale stratégique INRAE et membre du PIA **Institut Français de Bioinformatique**
- Un contexte technologique **du séquençage haut débit** en évolution permanente depuis de nombreuses années, ouvrant la possibilité d'améliorer de manière significative la connaissance des génomes complexes
- La génomique, une **source majeure d'innovation** pour la recherche publique comme pour les entreprises
- La nécessité de **développer davantage l'expertise bio-informatique** et bio statistique indispensable à ces analyses



Les objectifs

Acquisition d'expertises avancées sur l'utilisation des séquenceurs 3^e génération

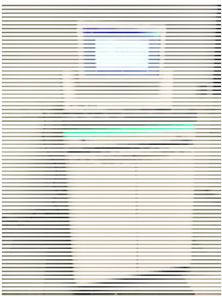
- Acquérir une expertise avancée sur les **nouvelles technologies de séquençage « Longs fragments »** et « **Molécules uniques** » à haut débit disponibles :

- Potentiels comparés des technologies
- Identification des combinaisons de technologies à mettre en œuvre en fonction des objectifs



... à trois niveaux d'étude :

- Axe 1 Le **génom**e : Connaissance de la variabilité du génome (ponctuelle ou structurale)
 - Axe 2 L'**épigénom**e : Etudes des marques épigénétiques de régulation de l'expression du génome
 - Axe 3 Les **métagénom**es : Connaissance fine des communautés au sein d'un métagénome
- Savoir conseiller les équipes publiques et privés pour l'utilisation des nouvelles technologies en fonction des objectifs recherchés



Les partenaires

Un partenariat monté pour renforcer le tryptique vertueux : *Plateforme – Equipe recherche – Partenaires privés.*

- **GeT-PlaGe, US 1426**
- **Genotoul Bioinfo, MIAT UR0875**
- **Sigenae, MIAT & GenPhySE UR0875 & UMR1388**
- **GenPhySE, UMR1388**
Génétique, Physiologie et Systèmes d'Elevage
- **GABI, UMR1313**
Génétique Animale et Biologie Integrative
- **MIAT, UR0875**
Mathématiques et Informatique appliquées Toulouse
- **GQE, UMR0320**
Génétique Quantitative et Evolution



Résultats attendus

Une expertise reconnue et une visibilité accrue des plateformes GeT et Bioinfo

- Une **expertise avancée sur l'utilisation des séquenceurs Longs fragments molécules uniques** pouvant être proposée :
 - Aux équipes académiques
 - ➔ Par **GeT-PlaGe** pour des projets de recherche (éventuellement avec partenariat Public Privé)
 - Aux entreprises privées
 - ➔ Par la Business Unit **GeT-IT** (INRAe Transfert) pour des prestations de routine



- La **production de jeux de données de référence distribués en open data** à l'issue du projet, contribuant à **créer une communauté de développeurs bioinfo** autour des plateformes GeT et Bioinfo

Le Génotypage Par Séquençage (GBS)

SeqOccin Axe n° 1 : l'analyse des variations des génomes

Le Génomage par séquençage (GBS)

Une alternative rentable et efficace aux technologies de génotypage par hybridation type puce ?

- Le génotypage à grande échelle est réalisé aujourd'hui à l'aide de **puces SNP**

Exemple : Affymetrix Axiom

Haut débit , Fiable mais peu modulable



Plateforme Gentyane, INRAE

- Le GBS est l'un des meilleurs candidats pour le génotypage de routine dans l'avenir
 - Découverte générale de marqueurs
 - Identification des haplotypes
 - Tests de pureté
 - Etudes d'association à l'échelle du génome (GWAS)
 - Sélection génomique (GS)
 - ...

Le Génomage par séquençage (GBS)

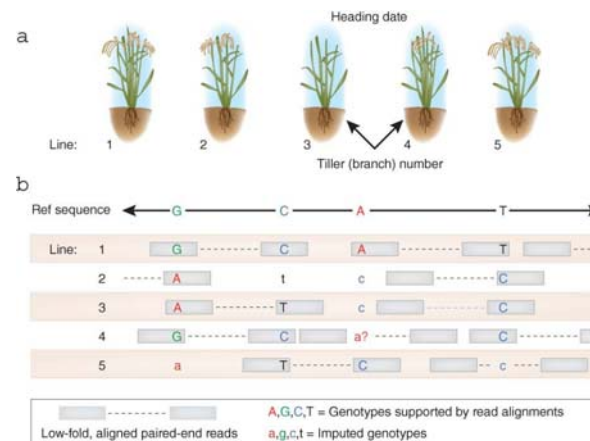
Les considérations à prendre en compte

● Génomome entier ou génome réduit ?

Exemple de méthodes utilisant des enzymes de restriction : RAD-Seq, ddRAD-Seq...

● Quelle profondeur ?

- En général , faible confiance si faible couverture
- Dans les cas extrêmes, 30X par locus et par individu est recommandé
- Cas particulier des **populations recombinantes avec génotypes parentaux de haute qualité** :
=> une couverture inférieure à 1 × peut suffire : **SKIM-Sequencing ou SPARSE Sequencing**

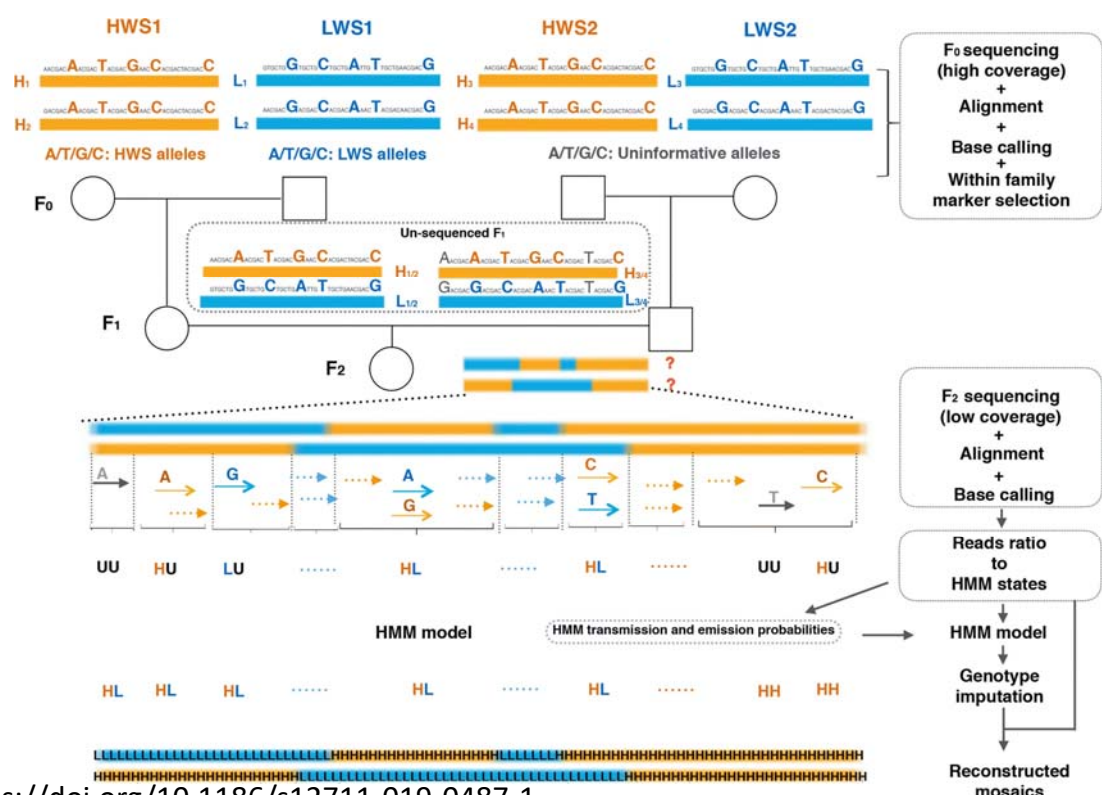


517 génomes de riz à 1X
Hatmap + GWAS
3,6 millions de SNP

SKIM Sequencing / SPARSE Sequencing

Objectif et Principe

- **Objectif** : Identifier quelle partie du génome a été héritée de chacun des individus parentaux
- **Design expérimental recommandé** : **fondeurs en 30X / descendance en 1X**
- L'**imputation** implique la déduction des génotypes manquants de manière computationnelle



SKIM Sequencing / SPARSE Sequencing

● Applications

- Construire des cartes d'haplotypes
- Estimer des taux de recombinaison
- Effectuer du GWAS
- Faire de la sélection génomique



● Projets pilote dans SeqOccln

- Maïs: Séquençage de 672 lignées en 0,8X (68 € par échantillon)
- Porc: Séquençage de 960 F2 en 0.8X (34 € par échantillon) + 12 parentaux en deep sequencing (1000 € par échantillon) => 3-4-5 M de SNP possibles

⇒ **Objectif : approcher le coût des puces de génotypage haut-débit**

L'apport des longs fragments pour la détection des marques de méthylation

SeqOccin Axe n° 2 : l'analyse des marques épigénétiques

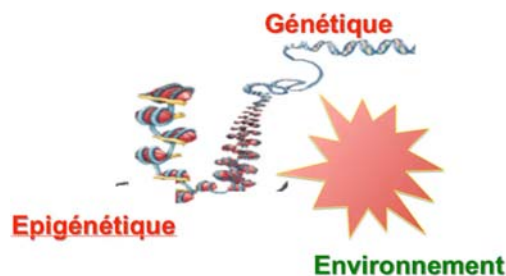
L'épigénétique pour mieux exprimer le potentiel génétique

Les facteurs d'élevage peuvent modifier l'expression des gènes et les performances

- **Epigénétique : Ensemble des changements héréditaires qui ne résultent pas d'une modification de la séquence des gènes**
 - Piste d'étude intéressante dans les productions animales



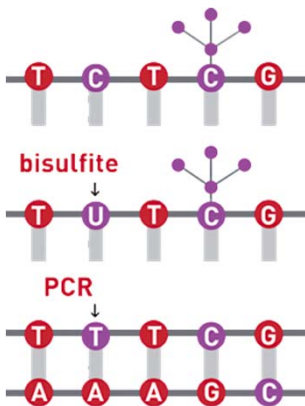
- **La recherche INRAE en lien avec la filière avicole**
 - Augmentation de la tolérance du poulet à la chaleur (ou au froid) en modifiant les paramètres d'incubation (métabolisme de vascularisation)
 - La nutrition précoce, dans l'oeuf ou au tout jeune âge, peut influencer certaines voies métaboliques et impacter la croissance ou la fertilité.
 - Une carence alimentaire chez la cane commune a un impact sur le poids de foie de ses descendants mulards



Méthodes actuelles pour détecter les methylations

- **Bisulfite conversion : 5mC, 5hmC, 4mC**
 - **WGBS** : Whole Genome Bisulfite Sequencing
 - **RRBS** : Reduced Representation Bisulfite Sequencing
 - **BSAS** : BiSulfite Amplicon Sequencing

⇒ **Le traitement au Bisulfite altère l'ADN**

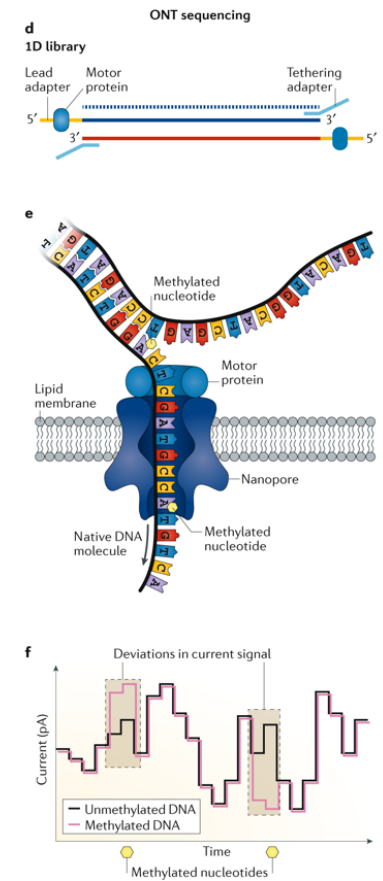


- **Digestion par une enzyme de restriction sensible à la méthylation** suivie d'un séquençage nouvelle génération (**MRE-Seq**) ou d'un pyroséquençage (**LUMA**): 4mC, 5mC, 6mA
- **Immunoprécipitation de l'ADN méthylé** suivi d'une puce à ADN (**MeDIP-Chip**) ou d'un séquençage (**MeDIP-Seq**): 5mC, 6mA



La détection des methylations avec Oxford Nanopore

Quel apport des longs fragments ?



● Sensibilité

Détecte les 6mA et 5mC

(Les algorithmes de détection des modifications sont en cours de développement)

● Couverture recommandée

~40X ou moins ?

Epigenetic modification		
	R9.4.1	R10.3
Called CpGs	> 94 %	> 94 %
Bisulphite correlation	0.94	0.96
Reproducibility	> 0.95	

PromethION, NA12878 @ 40 X depth
Megalodon with 5mC modified base model
Bisulphite correlation by Pearson coefficient, per site
ONT Poster: "Benchmarking nanopore methylation"

Projets pilote dans SeqOccin sur la Caille et le Porc

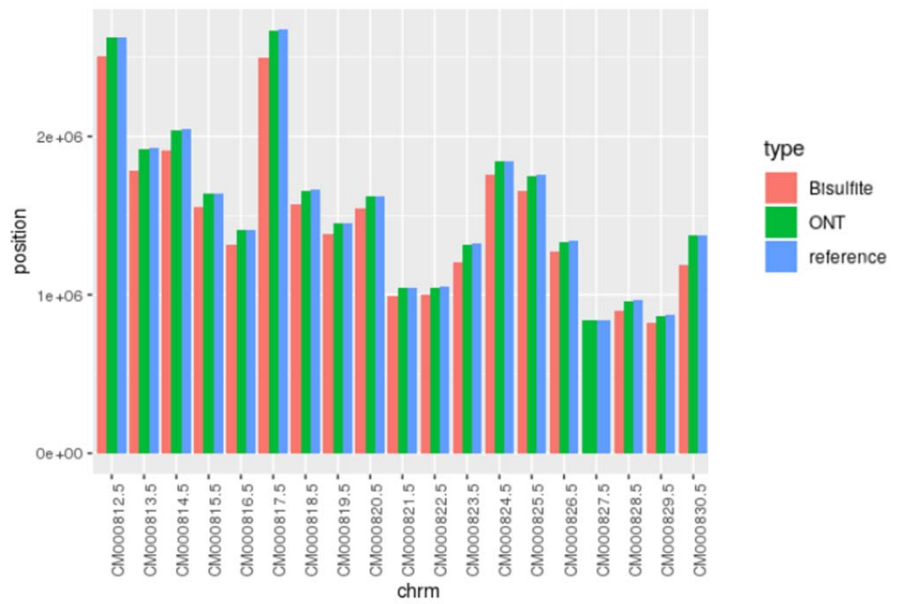
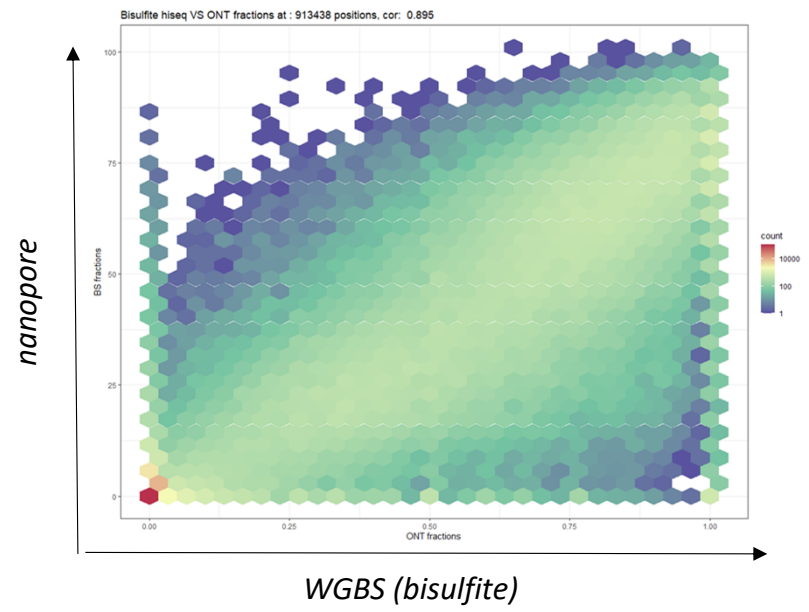
- Phase 1 : Développement de méthodes biologiques et bio-informatique pour identifier les marques de méthylation sur de l'ADN natif (5mC et 6mA)



Analyse sur le chromosome 1 de *C. Japonica*
Très bonne corrélation entre les 2 méthodes de détection (0,895)



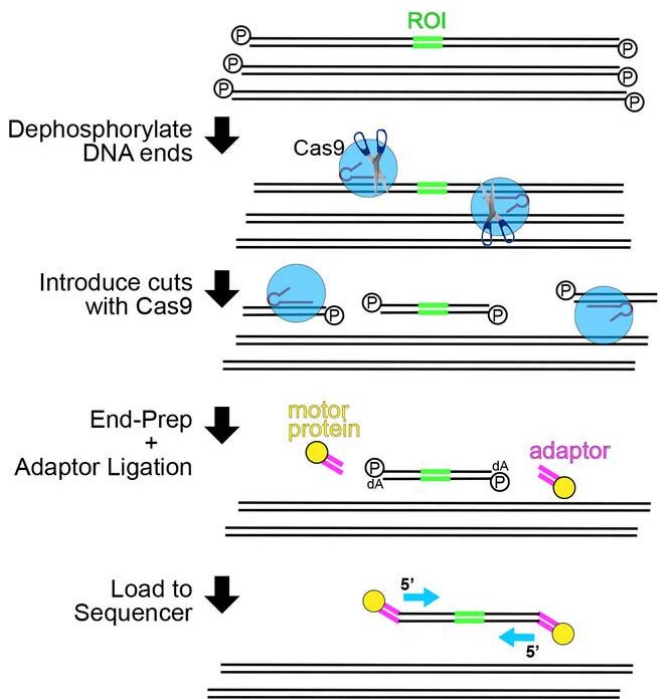
Whole genome CpG coverage – 8 % CpG detection improvement in ONT



Autres développements à venir

Phase 2 : Développement de méthodes d'analyses ciblées (Cas9)

(Gilpatrick et al 2019 doi: <https://doi.org/10.1101/604173>)



Oxford Nanopore - Flongle



⇒ **Obtention de l'information génétique ET épigénétique**

SeqOccln



INRAe Toulouse

Coordinatrice du projet :
Carole Iampietro
carole.iampietro@inrae.fr

Axe1 SeqOccln Assemblage et Variants

Clément Birbes

Arnaud Di-Franco

Andreea Dreau

Camille Eché

Thomas Faraut

Carole Iampietro

Christophe Klopp

Claire Kuchly

Matthias Zytnicki

Axe2 SeqOccln Epigénétique

Remy Felix Serres

Paul Terzian

Celine Vandecasteele

Christophe Klopp

Axe3 SeqOccln Métagénomique

Adrien Castinel

Jean Mainguy

Olivier Bouchez

Geraldine Pascal

Claire Hoede