



HAL
open science

Signatures moléculaires des gènes soumis à empreinte parentale chez le porc

Eva Jacomet

► **To cite this version:**

Eva Jacomet. Signatures moléculaires des gènes soumis à empreinte parentale chez le porc. Médecine vétérinaire et santé animale. 2023. dumas-04878862

HAL Id: dumas-04878862

<https://dumas.ccsd.cnrs.fr/dumas-04878862v1>

Submitted on 10 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ANNEE 2023 THESE : 2023 – TOU 3 – 4045

SIGNATURES MOLECULAIRES DES GENES SOUMIS A EMPREINTE PARENTALE CHEZ LE PORC

THESE D'EXERCICE pour obtenir le titre de

DOCTEUR VETERINAIRE

DIPLOME D'ETAT

*présentée et soutenue publiquement
devant l'Université Paul-Sabatier de Toulouse*

par

JACOMET Eva, Lucie

Directeur de thèse : M. Alain DUCOS

JURY

PRESIDENT :

M. Cédric LE CAIGNEC

Professeur à l'Université Paul Sabatier de TOULOUSE

ASSESEURS :

M. Alain DUCOS

Professeur à l'Ecole Nationale Vétérinaire de TOULOUSE

Mme Christelle CAMUS

Maître de Conférences à l'Ecole Nationale Vétérinaire de TOULOUSE

MEMBRES INVITES :

Mme Julie DEMARS

Chargée de recherche à l'INRAE, UMR GenPhySE

M. Jean-Noël HUBERT

Chercheur post-doctorant à l'INRAE, UMR GenPhySE

**Ministère de l'Agriculture et de l'Alimentation
ÉCOLE NATIONALE VÉTÉRINAIRE DE TOULOUSE**

Liste des directeurs/assesseurs de thèse de doctorat vétérinaire

Directeur : Professeur Pierre SANS

PROFESSEURS CLASSE EXCEPTIONNELLE

- M. **BERTAGNOLI Stéphane**, *Pathologie infectieuse*
- M. **BOUSQUET-MELOU Alain**, *Pharmacologie, thérapeutique*
- M. **BRUGERE Hubert**, *Hygiène et industrie des aliments d'origine animale*
- Mme **CHASTANT-MAILLARD Sylvie**, *Pathologie de la reproduction*
- M. **CONCORDET Didier**, *Mathématiques, statistiques, modélisation*
- M. **DELVERDIER Maxence**, *Anatomie pathologique*
- M. **ENJALBERT Francis**, *Alimentation*
- Mme **GAYRARD-TROY Véronique**, *Physiologie de la reproduction, endocrinologie*
- Mme **HAGEN-PICARD Nicole**, *Pathologie de la reproduction*
- M. **MEYER Gilles**, *Pathologie des ruminants*
- M. **SCHELCHER François**, *Pathologie médicale du bétail et des animaux de basse-cour*
- Mme **TRUMEL Catherine**, *Biologie médicale animale et comparée*

PROFESSEURS 1^{ère} CLASSE

- M. **BAILLY Jean-Denis**, *Hygiène et industrie des aliments*
- Mme **BOURGES-ABELLA Nathalie**, *Histologie, anatomie pathologique*
- Mme **CADIERGUES Marie-Christine**, *Dermatologie vétérinaire*
- M. **DUCOS Alain**, *Zootéchnie*
- M. **FOUCRAS Gilles**, *Pathologie des ruminants*
- M. **GUERIN Jean-Luc**, *Aviculture et pathologie aviaire*
- M. **JACQUIET Philippe**, *Parasitologie et maladies parasitaires*
- Mme **LACROUX Caroline**, *Anatomie pathologique, animaux d'élevage*
- Mme **LETRON-RAYMOND Isabelle**, *Anatomie pathologique*
- M. **LEFEBVRE Hervé**, *Physiologie et thérapeutique*
- M. **MAILLARD Renaud**, *Pathologie des ruminants*
- M. **MATHON Didier**, *Pathologie chirurgicale*

PROFESSEURS 2^{ème} CLASSE

- Mme **BOULLIER Séverine**, *Immunologie générale et médicale*
- M. **CORBIERE Fabien**, *Pathologie des ruminants*
- Mme **DIQUELOU Armelle**, *Pathologie médicale des équidés et des carnivores*
- M. **GUERRE Philippe**, *Pharmacie et toxicologie*
- Mme **MEYNADIER Annabelle**, *Alimentation animale*
- M. **MOGICATO Giovanni**, *Anatomie, imagerie médicale*
- Mme **PAUL Mathilde**, *Epidémiologie, gestion de la santé des élevages avicoles*
- M. **RABOISSON Didier**, *Médecine de population et économie de la santé animale*
- M. **VOLMER Romain**, *Microbiologie et infectiologie*

MAITRES DE CONFERENCES HORS CLASSE

- M. **BERGONIER Dominique**, *Pathologie de la reproduction*
- Mme **BIBBAL Delphine**, *Hygiène et industrie des denrées alimentaires d'origine animale*
- Mme **CAMUS Christelle**, *Biologie cellulaire et moléculaire*
- M. **JAEG Jean-Philippe**, *Pharmacie et toxicologie*
- M. **LYAZRHI Faouzi**, *Statistiques biologiques et mathématiques*
- Mme **PALIERNE Sophie**, *Chirurgie des animaux de compagnie*
- Mme **PRIYEMENKO Nathalie**, *Alimentation*

MAITRES DE CONFERENCES CLASSE NORMALE

- M. **ASIMUS Erik**, *Pathologie chirurgicale*
- Mme **BRET Lydie**, *Physique et chimie biologiques et médicales*
- Mme **BOHSIRA Emilie**, *Parasitologie, maladies parasitaires*
- M. **CARTIAUX Benjamin**, *Anatomie, imagerie médicale*
- M. **CONCHOU Fabrice**, *Imagerie médicale*
- Mme **DANIELS Hélène**, *Immunologie, bactériologie, pathologie infectieuse*
- Mme **DAVID Laure**, *Hygiène et industrie des aliments*
- M. **DIDIMO IMAZAKI Pedro**, *Hygiène et industrie des aliments*
- M. **DOUET Jean-Yves**, *Ophthalmologie vétérinaire et comparée*
- Mme **FERRAN Aude**, *Physiologie*
- Mme **GRANAT Fanny**, *Biologie médicale animale*
- Mme **JOURDAN Géraldine**, *Anesthésie, analgésie*
- M. **JOUSSERAND Nicolas**, *Médecine interne des animaux de compagnie*
- Mme **LALLEMAND Elodie**, *Chirurgie des équidés*
- Mme **LAVOUE Rachel**, *Médecine Interne*
- M. **LE LOC'H Guillaume**, *Médecine zoologique et santé de la faune sauvage*
- M. **LIENARD Emmanuel**, *Parasitologie et maladies parasitaires*
- Mme **MEYNAUD-COLLARD Patricia**, *Pathologie chirurgicale*
- Mme **MILA Hanna**, *Elevage des carnivores domestiques*
- M. **NOUVEL Laurent**, *Pathologie de la reproduction*
- M. **VERGNE Timothée**, *Santé publique vétérinaire, maladies animales réglementées*
- Mme **WARET-SZKUTA Agnès**, *Production et pathologie porcine*
- Mme **PALIERNE Sophie**, *Pathologie chirurgicale*
- M. **GAIDE Nicolas**, *Anatomie Pathologique*

INGENIEURS DE RECHERCHE

- M. **AUMANN Marcel**, *Urgences, soins intensifs*
- M. **AUVRAY Frédéric**, *Santé digestive, pathogénie et commensalisme des entérobactéries*
- M. **CASSARD Hervé**, *Pathologie des ruminants*
- M. **CROVILLE Guillaume**, *Virologie et génomique cliniques*
- Mme **DEBREUQUE Maud**, *Médecine interne des animaux de compagnie*
- Mme **DIDIER Caroline**, *Anesthésie, analgésie*
- Mme **DUPOUY GUIRAUTE Véronique**, *Innovations thérapeutiques et résistances*
- Mme **GAILLARD Elodie**, *Urgences, soins intensifs*
- Mme **GEFFRE Anne**, *Biologie médicale animale et comparée*
- Mme **GRISEZ Christelle**, *Parasitologie et maladies parasitaires*
- Mme **JEUNESSE Elisabeth**, *Bonnes pratiques de laboratoire*
- Mme **PRESSANTI Charline**, *Dermatologie vétérinaire*
- M. **RAMON PORTUGAL Félipe**, *Innovations thérapeutiques et résistances*
- M. **REYNOLDS Brice**, *Médecine interne des animaux de compagnie*
- Mme **ROUCH BUCK Pétra**, *Médecine préventive*
- M. **DELPONT Mattias**, *Clinique Aviaire*
- Mme **POUJADE Agnès**, *Anatomie pathologique Vétérinaire*
- Mme **LAYSSOL-LAMOUR Catherine**, *Imagerie Médicale*
- Mme **SAADA Chloé**, *Gestion intégrée de la santé des ruminants*

A Monsieur le Professeur Cédric Le Caignec

Professeur des Universités

Praticien Hospitalier

Génétique

Merci d'avoir accepté d'évaluer ce travail et de présider mon jury.

A Monsieur le Professeur Alain Ducos

Professeur de l'École Nationale Vétérinaire de Toulouse

Génétique

Merci d'avoir accepté d'être mon directeur de thèse. Merci pour vos précieux conseils pour la rédaction de ce manuscrit, et merci par ailleurs pour vos enseignements en Génétique.

A Madame le Docteur Christelle Camus

Maître de Conférence de l'École Nationale Vétérinaire de Toulouse

Biologie moléculaire

Merci d'avoir accepté d'être membre de mon jury, et merci par ailleurs pour vos enseignements en Biologie moléculaire.

A Madame le Docteur Julie Demars

Chargée de recherche à l'Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement

Génétique

Merci de m'avoir fait découvrir le monde de la recherche et de m'avoir intégrée dans votre équipe. Merci pour votre encadrement sans faille, et pour votre aide pour ce manuscrit.

A Monsieur le Docteur Jean-Noël Hubert

Post-doctorant à l'Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement

Génétique

Merci pour votre bienveillance et pour votre accompagnement dans la rédaction de ce manuscrit.

Liste des productions scientifiques associées à la thèse

Ce projet de thèse a été soutenu par le projet PIPETTE ANR-18-CE20-0018

Communications affichées dans un colloque national ou international (3)

Hubert JN, Iannuccelli N, Cabau C, **Jacomet E**, Billon Y, Serre RF, Vandecasteele C, Donnadiou C, Demars J. *A methyl-seq tool to capture genomic imprinted loci*. Genomic Imprinting – from Biology to Disease, 22-24 mar 2023, Cambridge, Royaume-Uni.

Hubert JN, Perret M, Iannuccelli N, **Jacomet E**, Cabau C, Demars J. *A bioinformatics pipeline to detect imprinting from methyl-seq trio data*. Genomic Imprinting – from Biology to Disease, 22-24 mar 2023, Cambridge, Royaume-Uni.

Hubert JN, Iannuccelli N, Cabau C, Serre RF, Vandecasteele C, **Jacomet E**, Brenot S, Billon Y, Hebrard W, Donnadiou C, Demars J. *Vers une carte de l'imprintome porcin : comparaison de deux méthodes de méthyl-seq*. Journées Scientifiques du Département de Génétique Animale, 13-15 sep 2022, Bordeaux, France.

Communication orale dans un colloque international (1)

Hubert JN, Iannuccelli N, Cabau C, Serre RF, Vandecasteele C, **Jacomet E**, Brenot S, Billon Y, Hebrard W, Donnadiou C, Demars J. *Targeting the pig imprintome: a comparative approach using two methyl-seq technologies*. Advances in Genome Biology and Technology Agriculture (AGBT-Ag), 4-6 avr 2022, San Diego, United States.

Article soumis dans une revue à comité de lecture (1)

Hubert JN, Iannuccelli N, Cabau C, **Jacomet E**, Billon Y, Serre RF, Vandecasteele C, Donnadiou C, Demars J. *A methyl-seq tool to capture genomic imprinted loci*. bioRxiv, 2023, <https://doi.org/10.1101/2023.02.21.529206>

Table des matières

PARTIE 1 : SYNTHÈSE BIBLIOGRAPHIQUE 14

I. L'ADAPTATION DE LA FILIÈRE PORCINE FRANÇAISE	14
1.1. DÉTERMINISME GÉNÉTIQUE DES CARACTÈRES	15
1.1.1. Caractères à déterminisme monogénique	15
1.1.2. Caractères à déterminisme polygénique	16
1.1.3. Du modèle monogénique au modèle omnigénique.....	17
1.2. L'ÉLEVAGE PORCIN EN FRANCE.....	19
1.2.1. Historique de la filière porcine française	19
1.2.2. Organisation de la filière porcine actuelle : un système pyramidal	21
1.2.2.1. L'étage de sélection	22
1.2.2.2. L'étage de multiplication	22
1.2.2.3. L'étage de production	23
1.2.2.4. Une rationalisation au niveau de la filière comme au niveau de l'élevage.....	23
1.2.3. Notions de génétique quantitative.....	24
1.2.3.1. Composantes influençant le phénotype d'un individu	24
1.2.3.2. Estimation de la valeur génétique additive	25
1.2.3.3. Estimation des paramètres génétiques	26
1.2.3.3.1. L'héritabilité.....	26
1.2.3.3.2. Les corrélations génétiques.....	28
1.2.4. Évolution des méthodes de sélection	29
1.2.4.1. Principes d'amélioration génétique.....	29
1.2.4.2. Sélection polygénique classique	30
1.2.4.3. BLUP-Modèle animal.....	30
1.2.4.4. Introduction à la sélection génomique (SG)	31
1.2.5. Évolution des objectifs de sélection.....	31
1.2.6. Vers de nouveaux objectifs de sélection	34
II. LES RESSOURCES OMIQUES COMME SUPPORT DE L'ÉVOLUTION DE LA SELECTION PORCINE	36
2.1. LA GÉNOMIQUE	36
2.1.1. Les années 1990 : les marqueurs génétiques	36
2.1.1.1. Historique de l'utilisation des marqueurs génétiques	36
2.1.1.2. Quelques applications des marqueurs génétiques.....	37
2.1.2. Les puces de génotypage.....	38
2.1.3. En quoi la sélection génomique (SG) représente-elle une révolution technologique ?	40
2.1.4. Le développement des techniques de séquençage en trois révolutions	41
2.1.4.1. 1 ^{ère} révolution : le séquençage Sanger (1977)	41
2.1.4.2. 2 ^{ème} révolution : le séquençage nouvelle génération (ou NGS : Next Generation Sequencing).....	42
2.1.4.3. 3 ^{ème} révolution : longues lectures et séquençage de la molécule d'ADN native.....	44
2.2. LES ÉLÉMENTS NON MENDELÉIENS MAIS TRANSMISSIBLES.....	46
2.2.1. Épigenétique.....	47
2.2.1.1. Définition de l'épigénétique.....	47
2.2.1.2. L'épigénétique transgénérationnelle	48
2.2.2. Microbiote	50
2.2.3. Génome mitochondrial.....	52

2.2.4. Héritage culturel.....	52
III. LES GENES SOUMIS A EMPREINTE PARENTALE	54
3.1. UNE EXPRESSION GÉNIQUE SINGULIÈRE.....	54
3.1.1. Découverte des gènes soumis à empreinte parentale	54
3.1.2. Cycle de l’empreinte	57
3.1.2.1. Étape 1 : l’établissement de l’empreinte parentale	58
3.1.2.2. Étape 2 : le maintien et la lecture.....	59
3.1.2.3. Étape 3 : l’effacement des marques d’empreinte	60
3.1.3. Mécanismes moléculaires impliqués dans la régulation des gènes soumis à empreinte parentale	61
3.1.3.1. Mécanisme insulateur	62
3.1.3.2. Mécanisme avec action de longs ARN non codants (ARNInc).....	63
3.1.3.3. Rôles des petits ARN non codants.....	64
3.1.3.4. L’empreinte non-canonique : un autre type de marque d’empreinte	65
3.1.4. Une expression tissu-dépendante des gènes soumis à empreinte	66
3.1.4.1. Le placenta, un organe clef pour les mécanismes d’empreinte parentale.....	66
3.1.4.2. Exemple du gène growth factor receptor bound protein 10 (Grb10).....	68
3.2. UNE INFORMATION TRANSGÉNÉRATIONNELLE JOUANT UN RÔLE DANS LES FONCTIONS DE DÉVELOPPEMENT ET CROISSANCE.....	69
3.2.1. Rôle des gènes soumis à empreinte parentale	69
3.2.1.1. Rôles dans le développement embryonnaire.....	69
3.2.1.2. Rôle sur le développement et la physiologie du système nerveux central.....	70
3.2.1.3. Rôle sur le développement de la glande mammaire	71
3.2.2. Connaissances comparées entre l’Homme, la souris et le porc	72
3.2.3. Lien entre empreinte parentale et santé chez l’Homme	73
3.2.3.1. Syndrome d’Angelman (SA)	74
3.2.3.2. Syndrome Prader-Willi (PWS)	75
3.3. LES THÉORIES SUR LE SUCCÈS ÉVOLUTIF DE L’EMPREINTE PARENTALE	76
3.3.1. Théorie de la prévention de la parthénogénèse	77
3.3.2. Théorie de la « bombe à retardement ovarienne »	77
3.3.3. Théorie de l’antagonisme sexuel	78
3.3.4. Théorie du conflit parental (ou « la bataille des sexes »).....	78
3.3.5. Théorie de la coadaptation foeto-maternelle.....	79

PARTIE 2 : ÉTUDE EXPÉRIMENTALE DES SIGNATURES MOLÉCULAIRES DE L’EMPREINTE PARENTALE SUR LA RÉGION *H19-IGF2/KCNQ1-CDKN1C* **81**

I. CONTEXTE.....	81
1.1. INTÉRÊT DU CROISEMENT RÉCIPROQUE POUR L’ÉTUDE DE L’EMPREINTE	81
1.2. DONNÉES NÉCESSAIRES À L’ÉVALUATION DES MÉCANISMES D’EMPREINTE.....	82
1.2.1. Les données génomiques	82
1.2.2. Les données épigénomiques.....	83
1.2.3. Les données transcriptomiques	84
1.3. CHOIX DE LA RÉGION D’ÉTUDE	85
II. MATÉRIELS ET MÉTHODES.....	86
2.1. LES ÉLÉMENTS DISPONIBLES AVANT MON ARRIVÉE	86
2.1.1. Les animaux et échantillons.....	86

2.1.2. Les données.....	87
2.2. ANALYSES BIOINFORMATIQUES	88
2.2.1. Séquençage de l'épigénome.....	88
2.2.1.1. Détermination des haplotypes.....	89
2.2.1.2. Identification des DMRs candidates	89
2.3. PROTOCOLES DE BIOLOGIE MOLÉCULAIRE.....	90
2.3.1. Design des amorces selon les expériences.....	90
2.3.1.1. Amorces de génotypage via la méthode PACE®	90
2.3.2. Conditions PCR selon les expériences.....	94
2.3.2.1. PCR pour la méthode PACE®.....	94
2.3.2.2. PCR pour le séquençage Sanger	94
2.4. LOGICIELS ET SITES UTILISÉS	95
III. RESULTATS	96
3.1. ORIGINE PARENTALE DES ALLÈLES	97
3.1.1. Sélection bioinformatique de variants informatifs.....	97
3.1.2. Validation des génotypes	98
3.2. IDENTIFICATION DE RÉGIONS HÉMI-MÉTHYLÉES	99
3.2.1. Analyse des données épigénomiques.....	99
3.2.2. Méthylation dépendante de l'origine parentale.....	101
3.2.2.1. Identification bioinformatique de DMR candidats	101
3.2.2.2. Validation de DMR parentale-spécifique	101
3.3. IDENTIFICATION DE L'EXPRESSION PARENTALE-SPÉCIFIQUE	103
3.3.1. Analyse des données transcriptomiques	104
3.3.2. Premières expériences de validation de l'expression de gènes soumis à empreinte.....	105
IV. DISCUSSION	106
4.1. MISE EN ÉVIDENCE DE DEUX RÉGIONS SOUMISES À EMPREINTE PARENTALE.....	106
4.2. UNE APPROCHE MULTIOMIQUE LOCALE POSANT DES BASES POUR LA CARACTERISATION DE L'IMPRINTOME PORCIN.....	108
CONCLUSION.....	110
BIBLIOGRAPHIE	111
ANNEXE	138

Liste des figures

FIGURE 1 : REPRÉSENTATION SCHÉMATIQUE DE LA DISTRIBUTION DE LA POPULATION POUR UN CARACTÈRE QUANTITATIF.	17
FIGURE 2 : UN CONTINUUM ALLANT DU DÉTERMINISME MONOGÉNIQUE AU DÉTERMINISME OMNIGÉNIQUE.....	18
FIGURE 3 : SYNTHÈSE DES GRANDES ÉTAPES DE LA STRUCTURATION DE LA FILIÈRE PORCINE FRANÇAISE.	21
FIGURE 4 : LA FILIÈRE PORCINE, UN SYSTÈME PYRAMIDAL À TROIS NIVEAUX (PINEL 2023).	22
FIGURE 5 : ORGANISATION DE LA CONDUITE EN 7 BANDES AVEC UN SEVRAGE À 28 JOURS.	23
FIGURE 6 : ESTIMATION DE L'HÉRITABILITÉ VIA LA RÉGRESSION PARENT-ENFANT.	27
FIGURE 7 : DÉSÉQUILIBRE DE LIAISON ET RECOMBINAISON LORS DE LA MÉIOSE.	28
FIGURE 8 : ÉTAPES D'UN SCHÉMA DE SÉLECTION.	29
FIGURE 9 : ÉVOLUTION DES OBJECTIFS DE SÉLECTION EN PRODUCTION VIANDE.	33
FIGURE 10 : ÉVOLUTION AU COURS DU TEMPS DES CARACTÈRES GMQ (A) ET PORCELETS NÉS VIFS PAR PORTÉES (B) DANS LES RACES LWF, LR, PI ET AU NIVEAU DE LA GTTT.	33
FIGURE 11 : LES ENJEUX AUXQUELS LE MONDE DE L'ÉLEVAGE DOIT FAIRE FACE.	35
FIGURE 12: PRINCIPES DES PUCES DE GÉNOTYPAGE.	39
FIGURE 13 : ÉVOLUTION DU COÛT DU SÉQUENÇAGE DU GÉNOME HUMAIN ENTRE 2001 ET 2021.	41
FIGURE 14 : PRINCIPE DU SÉQUENÇAGE SANGER.....	42
FIGURE 15 : PRINCIPE DE LA TECHNOLOGIE ILLUMINA.	43
FIGURE 16 : AVANTAGES DES LONGUES LECTURES POUR LE NGS.	44
FIGURE 17 : SCHÉMATISATION DE QUELQUES ÉTAPES CLÉS DE LA RÉVOLUTION TECHNOLOGIQUE DANS LE DOMAINE DE LA GÉNÉTIQUE.	46
FIGURE 18 : CONCEPT DE L'ÉPIGÉNÉTIQUE TRANSGÉNÉRATIONNELLE.	49
FIGURE 19 : SCHÉMATISATION DE L'EXPRESSION MONO-ALLÉLIQUE PARENT-DÉPENDANTE DES GÈNES SOUMIS À EMPREINTE.....	54
FIGURE 20 : EXPÉRIENCES DE TRANSFERT DE NOYAUX (SOLTER).	55
FIGURE 21 : EXPÉRIENCES DE TRANSLOCATION (CATTANACH).....	56
FIGURE 22 : CYCLE DE L'EMPREINTE PARENTALE CHEZ LA SOURIS.	58
FIGURE 23 : VARIATION DU NIVEAU DE MÉTHYLATION AU COURS DU CYCLE DE L'EMPREINTE.	59
FIGURE 24 : LES MÉCANISMES DE DÉMÉTHYLATION : LA DÉMÉTHYLATION PASSIVE VERSUS LA DÉMÉTHYLATION ACTIVE VIA L'ACTION DES ENZYMES TET.	61
FIGURE 25 : RÉGULATION ÉPIGÉNÉTIQUE BASÉE SUR LE MODÈLE INSULATEUR : EXEMPLE DES GÈNES IGF2 ET H19.	63
FIGURE 26 : RÉGULATION ÉPIGÉNÉTIQUE BASÉE SUR LE MODÈLE INSULATEUR : EXEMPLE DU CLUSTER IGF2R.	64
FIGURE 27 : COMPARAISON DES STRATÉGIES DE REPRODUCTION ET DE L'EXISTENCE DE L'EMPREINTE GÉNOMIQUE CHEZ DIFFÉRENTS GROUPES DE VERTÉBRÉS.	66
FIGURE 28 : COMPARAISON HISTOLOGIQUE SIMPLIFIÉE DES DIFFÉRENTES PLACENTATIONS EN FONCTION DES ESPÈCES.	67
FIGURE 29 : SCHÉMA RÉSUMANT LES RÔLES PHYSIOLOGIQUES DE LA PROTÉINE IGF2 DÉCRITS CHEZ L'HOMME ET MODÈLES DE SOURIS.	68
FIGURE 30 : RÉPARTITION DES GÈNES SOUMIS À EMPREINTE CONNUS CHEZ L'HOMME (A), LA SOURIS (B) ET LE PORC (C).	73
FIGURE 31 : SCHÉMA DE LA RÉGION 15Q11-2-Q13 IMPLIQUÉE DANS LES SYNDROMES D'ANGELMAN ET PRADER-WILLI.	74
FIGURE 32 : CROISEMENT RÉCIPROQUE DANS LE CADRE DE L'ÉTUDE.	81
FIGURE 33 : PRINCIPE DE LA CONVERSION AU BISULFITE DE L'ADN.....	83
FIGURE 34 : ORGANISATION DE LA PARTIE EXPÉRIMENTALE EN TROIS AXES PRINCIPAUX.	86
FIGURE 35 : LES TROIS FAMILLES UTILISÉES POUR CETTE ÉTUDE. EN ROUGE SOULIGNÉS : LES INDIVIDUS DONT LES DONNÉES DE SÉQUENÇAGE D'ADNG TRAITÉ SONT DISPONIBLES.	86
FIGURE 36 : SCHÉMA SIMPLIFIÉ DU PIPELINE UTILISÉ POUR LA DÉTECTION DES VARIANTS.	90
FIGURE 37 : PROGRAMMES PCR UTILISÉS POUR LA PACE® SUR ADNG.....	94
FIGURE 38 : PROGRAMME PCR CLASSIQUE UTILISÉ.	95
FIGURE 39 : FICHIER D'ANNOTATION DU GÉNOME PORCIN UTILISÉ GTF SUSSCROFA 11.1.109.....	96

FIGURE 40 : LOCALISATION DES 21 VARIANTS SÉLECTIONNÉS.	97
FIGURE 41 : GÉNOTYPAGE SUR ADNG À L'AIDE DE LA MÉTHODE PACE®.....	99
FIGURE 42 : VISUALISATION DES DONNÉES ÉPIGÉNOMIQUES DE L'INDIVIDU 128 À L'AIDE DE IGV.....	100
FIGURE 43 : COMPARAISON DES DÉTECTIONS BIOINFORMATIQUE ET VISUELLE DES RÉGIONS HÉMI-MÉTHYLÉES.	100
FIGURE 44 : EXEMPLE DU VARIANT 1 384 751 PB DISCRIMINANT LES CpG MÉTHYLÉES ET NON MÉTHYLÉES.	101
FIGURE 45 : ÉLECTROPHORÈSE SUR GEL D'AGAROSE DES PRODUITS PCR OBTENUS AVEC LES AMORCES DE MS-PCR.....	102
FIGURE 46 : CHROMATOGRAMME DU FRAGMENT MÉTHYLÉ ISSU DE MUSCLE POUR LE VARIANT 1 384 751 PB.....	103
FIGURE 47 : VISUALISATION DES DONNÉES TRANSCRIPTOMIQUES AVOISINANT 2 VARIANTS D'INTÉRÊT.	104
FIGURE 48 : GÉNOTYPAGE SUR ADNG ET ADNc DU VARIANT 1 645 662 PB L'AIDE DE LA MÉTHODE PACE®....	105
FIGURE 49 : SCHÉMATISATION DES RÉSULTATS OBTENUS CONCERNANT L'EMPREINTE PARENTALE DE LA RÉGION H19-IGF2/KCNQ1-CDKN1C CHEZ LE PORC (CHROMOSOME 2, ENTRE LES POSITIONS 1 381 838 BP ET 2 059 694 BP).....	106

Liste des tableaux

TABLEAU 1 : ESTIMATION DE L'HÉRITABILITÉ OBTENUE DANS UNE POPULATION NORVÉGIENNE DE LANDRACE NÉS ENTRE 1990 ET 2000.	27
TABLEAU 2 : EXEMPLES DE CORRÉLATIONS GÉNÉTIQUES ENTRE CARACTÈRES (HOLM ET AL. 2004).....	29
TABLEAU 3 : ÉCHANTILLONS ET DONNÉES DISPONIBLES.	88
TABLEAU 4 : AMORCES UTILISÉES POUR LA PACE® SUR ADNG.....	92
TABLEAU 5 : AMORCES UTILISÉES POUR LE SÉQUENÇAGE SANGER SUR ADNG POUR LE SNP N°9 (1 859 450 PB).	93
TABLEAU 6 : AMORCES UTILISÉES POUR LA PCR MÉTHYLATION-SPÉCIFIQUE.	93
TABLEAU 7 : RÉCAPITULATIF DES 21 VARIANTS SÉLECTIONNÉS SUR LA RÉGION D'ÉTUDE.	98

Liste des encadrés

ENCADRÉ 1 : LA PÉNÉTRANCE.	16
ENCADRÉ 2 : LE DÉSÉQUILIBRE DE LIAISON.	28
ENCADRÉ 3 : TECHNOLOGIE DE SÉQUENÇAGE À L'AIDE D'UN NANOPORE.	45
ENCADRÉ 4 : HÉRITAGES CULTUREL ET ÉCOLOGIQUE.	53
ENCADRÉ 5 : PRINCIPE DE LA PACE®.	82

Liste des abréviations

A : POTENTIEL GÉNÉTIQUE ADDITIF OU VALEUR GÉNÉTIQUE ADDITIVE
AB : AGRICULTURE BIOLOGIQUE
ADN : ACIDE DÉSOXYRIBONUCLÉIQUE
ADNC : ADN COMPLÉMENTAIRE
ADNG : ADN GÉNOMIQUE
AOP : GÈNE AQUAPORINE
ARN : ACIDE RIBONUCLÉIQUE
ARNLNC : ARN LONG NON CODANT
ARNr : ARN RIBOSOMIQUE
ASCL2 : GÈNE « ACHAETE-SCUTE COMPLEX HOMOLOG 2 »
AOC : APPELLATION D'ORIGINE CONTRÔLÉE
AOP : APPELLATION D'ORIGINE PROTÉGÉE
ASP : AGENBCE DE SÉLECTION PORCINE
BLAST : « BASIC LOCAL ALIGNMENT SEARCH TOOL »
BLUP : « BEST LINEAR UNBIASED PREDICTOR » OU MEILLEURE PRÉDICTION LIÉNAIRE NON BIAISÉE
CG : DINUCLÉOTIDE CYTOSINE-GUANINE
cM : CENTIMORGAN
CNAG : COMISSION NATIONALE D'AMÉLIORATION GÉNÉTIQUE
CPG : CYTOSINE-PHOSPHATE-GUANINE
D : POTENTIEL GÉNÉTIQUE DE DOMINANCE
ddNTP : DIDÉSOCYRIBONUCLÉOTIDE
DL : DÉSÉQUILIBRE DE LIAISON
DMR : « DIFFERENTIALLY METHYLATED REGION » OU RÉGION DIFFÉRENTIELLEMENT MÉTHYLÉE
DNMT : ADN MÉTHYLTRANSFÉRASE
dNTP : DÉSOXYRIBONUCLÉOTIDE
DOK7 : GÈNE DOCKING PROTEIN 7
E : CONDITIONS DE L'ENVIRONNEMENT
G : POTENTIEL GÉNÉTIQUE
G X E : INTERACTION ENTRE POTENTIEL GÉNÉTIQUE ET ENVIRONNEMENT
GDS : GROUPEMENT DE DÉFENCE SANITAIRE
GIS : GROUPEMENT D'INTÉRÊT SCIENTIFIQUE
GRB10 : GÈNE « GROWTH FACTOR RECEPTOR BOURND PROTEIN 10 »
GWAS : « GENOME WIDE ASSOCIATION STUDY »
H2 : HÉRITABILITÉ
HALN : GÈNE DE SENSIBILITÉ À L'HALOTHANE MUTÉ
I : POTENTIEL GÉNÉTIQUE D'INTERACTION
IA : INSÉMINATION ARTIFICIELLE
ICE : « IMPRINTED CONTROL ELEMENT »
ICR : « IMPRINTED CONTROL REGION »
IFIP OU ITP : INSTITUT FRANÇAIS DU PORC
IGF2 : GÈNE « INSULINE-LIKE GROWTH FACTOR 2 »
IGF2R : GÈNE « INSULINE-LIKE GROWTH FACTOR 2 RECEPTOR »
IGP : INDICATION GÉOGRAPHIQUE PROTÉGÉE
IGV : LOGICIEL « INTEGRATED GENOME VIEWER »
INRA : INSTITUT NATIONAL DE RECHERCHE POUR L'AGRICULTURE
IYD : GÈNE « IODOTYROSINE DEIODINASE »
KCNQ1 : GÈNE « POTASSIUM VOLTAGE-GATED CHANNEL SUBFAMILY Q MEMBER 1 »
KDM1B : LYSINE DÉMÉTHYLASE 1B
LDL : « LOW DENSITY LIPOPROTEIN » OU PROTÉINE DE FAIBLE DENSITÉ
LR : LANDRACE
LW : LARGE WHITE

MIARN : MICROARN
 MS-PCR : PCR METHYLATION-SPECIFIQUE
 NDN : GÈNE « NECDIN »
 MS : MEISHAN
 NHGRI : « NATIONAL HUMAN GENOME RESEARCH INSTITUTE » OU INSITUT NATIONAL AMÉRICAIN DE RECHERCHE SUR LE GÉNOME HUMAIN
 ONT : « OXFORD NANOPORE TECHNOLOGIES »
 OSBPL5 : GÈNE « OXYSTEROL BINDING PROTEIN LIKE 5 »
 OSP : ORGANISME DE SÉLECTION PORCINE
 P : PERFORMANCE
 PACE : PCR ALLÈLE-SPÉCIFIQUE
 PB : PAIRE DE BASE
 PCR : « POLYMÉRASE CHAIN REACTION » OU RÉACTION POLYMÉRASE EN CHAÎNE
 PEG 5 : GÈNE « PATERNALLY EXPRESSED GENE 5 »
 PEG 10 : GÈNE « PATERNALLY EXPRESSED GENE 10 »
 PGH : PROJET GÉNOME HUMAIN
 PI : PIÉTRAIN
 PIARN : ARN INTERAGISSANT AVEC LES PROTÉINES PIWI
 PIC : « PIG IMPROVEMENT COMPAGNY »
 PRKAG3 : GÈNE « PROTEIN KINASE AMP-ACTIVATED NON CATALYTIC SUBUNIT GAMMA 3 »
 PSE : « PALE, SOFT, EXSUDATIVE »
 PWS : PRADER-WILLI SYNDROME
 QTL : « QUANTITATIVE TRAIT LOCUS »
 RISC : « RNA-INDUCED SILENCING COMPLEX »
 RN : GÈNE « RENDEMENT NAPOLE »
 RYR1 : GÈNE « RYANODINE RECEPTOR 1 »
 SA : SYNDROME D'ANGLEMAN
 SAM : SÉLECTION ASSISTÉE PAR MARQUEURS
 SFMBT2 : GÈNE « SCM LIKE WITH FOUR MBT DOMAINS 2 »
 SG : SÉLECTION GÉNOMIQUE
 SIARN : ARN INTERFÉRENT
 SNOARN : ARN NUCLÉOLAIRE
 SNP : « SINGLE NUCLÉOTIDE POLYMORPHISM »
 SNRPN : « SMALL NUCLEAR RIBONUCLEOPROTEIN POPYPEPTIDE N »
 TET : « TEN ELEVEN TRANSLOCATION ENZYME », OU ENZYME DE TRANSLOCATION 10-11 »
 TSA : TROUBLE DU SPECTRE AUTISTIQUE
 TSSC4 : GÈNE « TUMOR SUPPRESSING SUBTRANSFERABLE CANDIDATE 4 »
 UBE3A : GÈNE « UBIQUITIN PROTEIN LIGASE E3A »
 UPRA : UNITÉ NATIONALE DE PROMOTION DES RACES AGRÉES
 ZBDF2 : GÈNE « ZINC FINGER DBF-TYPE CONTAINING 2 »
 ZFP57 : « ZINC FINGER PROTEIN 57 » OU PROTÉINE EN DOIGTS DE ZINC 57
 ZFP445 : « ZINC FINGER PROTEIN 445 » OU PROTÉINE EN DOIGTS DE ZINC 445
 5HMC : 5-HYDROXYMÉTHYLCYTOSINE
 5MC : 5-MÉTHYLCYTOSINE

PARTIE 1 : Synthèse bibliographique

I. L'ADAPTATION DE LA FILIÈRE PORCINE FRANÇAISE

La plupart des espèces sont diploïdes, ce qui signifie que chaque individu hérite d'une combinaison de segments de chromosomes parentaux composant un génome nucléaire unique, fruit de facteurs que l'on peut examiner à différentes échelles. L'échelle populationnelle est souvent la plus pertinente pour appréhender les grandes forces évolutives et leur impact sur façon dont la variation génétique est répartie. Classiquement, quatre grandes forces évolutives sont décrites (Allendorf et al. 2022) :

(i) La mutation (pouvant survenir au sein de la lignée somatique et donc influencer uniquement l'individu porteur ou bien au sein de la lignée germinale ce qui implique une transmissibilité à la descendance), qui représente la source première de variation en produisant de nouveaux allèles,

(ii) La sélection (naturelle ou artificielle), qui désigne le processus de tri des individus selon la valeur adaptative relative de leur génotype, et qui peut aboutir à la perte de certains allèles,

(iii) La migration, qui désigne l'apport d'allèles issus d'autres populations, et qui tend à augmenter la variation génétique disponible en même temps qu'à homogénéiser génétiquement les populations connectées,

(iv) La dérive, qui traduit l'effet aléatoire lié à la taille finie de la population, et qui peut provoquer la perte d'allèles.

La variation génétique constitue une base dans laquelle la sélection va pouvoir puiser pour retenir les variants d'intérêt. Dans le cas des populations domestiquées soumises à une sélection artificielle, les programmes de sélection peuvent conduire à une baisse soutenue de la diversité génétique, puisqu'ils consistent à ne faire reproduire qu'une fraction des individus à chaque génération. Or, un appauvrissement de la variation génétique disponible peut contribuer à l'augmentation de la proportion d'allèles récessifs délétères dans la population et à diminuer le nombre de variants associés aux caractères non sélectionnés, d'où l'intérêt d'adopter des pratiques permettant de limiter la perte de diversité génétique au cours du temps tout en sélectionnant (Doublet et al. 2019).

C'est donc la présence de variation génétique au sein d'une population, sous l'effet de forces évolutives comme la mutation et la migration, qui rend possible la sélection de certains caractères d'intérêts présentant des modèles de déterminisme variés.

1.1. Déterminisme génétique des caractères

1.1.1. Caractères à déterminisme monogénique

Un caractère à déterminisme monogénique est classiquement défini comme étant gouverné par un seul gène. Dans les maladies à déterminisme monogénique, une seule mutation pathogène peut donc être causale bien qu'il puisse exister de l'hétérogénéité allélique (différentes mutations au sein d'un même gène peuvent avoir le même effet).

Un exemple de déterminisme monogénique est fourni par le gène de sensibilité à l'halothane chez le porc. Une mutation récessive (notée Halⁿ) touchant le gène *ryanodine receptor 1 (RYR1)* est en effet à l'origine de plusieurs effets zootechniques (Fujii et al. 1991). Les individus homozygotes pour l'allèle muté (Halⁿⁿ) vont ainsi développer une hyperthermie maligne suite à une anesthésie gazeuse à l'halothane, en raison d'un hypermétabolisme des muscles squelettiques. De plus, un lien entre ce génotype et le phénotype « viande PSE » (Pale, Soft and Exsudative) a été mis en évidence. Cet état de la viande PSE est une conséquence d'une réponse exacerbée au stress généré par le transport et la zone d'attente avant abattage, et est un des problèmes de qualité de viande les plus fréquents (Salas, Mingala 2017). Chez le porc, un autre gène impliqué dans la glycolyse musculaire a un effet majeur sur la qualité de la viande. Il s'agit du gène Protein Kinase AMP-Activated Non-Catalytic Subunit Gamma 3 (PRKAG3), initialement nommé Rendement Napole (RN), qui est associé au phénotype « viande acide » (Milan et al. 2000). La viande des sujets porteur de l'allèle muté RN- (allèle dominant) présente une perte de poids importante à la cuisson ainsi qu'une acidité élevée (en raison d'une augmentation de la teneur du muscle en glycogène, conduisant au moment de l'abattage à la production d'acide lactique en grande quantité), contrairement à la viande des sujets homozygotes rn+ (allèle récessif) (Salas, Mingala 2017).

Dans ces cas, un seul gène est responsable du phénotype, ce qui simplifie la contre-sélection des individus porteurs des allèles défavorables. Par exemple pour le caractère « sensibilité à l'halothane », un test se basant sur la détection de la mutation impliquée dans ce phénotype a pu être commercialisé. Cependant le lien entre génotype et phénotype pour un caractère monogénique peut s'avérer difficile à mettre en évidence en cas de pénétrance incomplète (Encadré 1) par exemple ou de pléiotropie.

Encadré 1 : La pénétrance.

La pénétrance est la proportion d'individus possédant un génotype donné qui exprime le phénotype. Par exemple, dans le cas d'un locus à deux allèles A et a (avec a récessif responsable d'une affection héréditaire par exemple) : on dira que la pénétrance est de 0,8 (donc incomplète) si 80% seulement des individus de génotype aa ont le phénotype auquel le génotype aa prédispose (Kingdom, Wright 2022). Par exemple, chez les bovins le caractère culard, bien connu des éleveurs et se traduisant par une hypertrophie musculaire, est lié à la présence d'un allèle récessif mh du gène codant pour la myostatine. Tous les individus homozygotes mh/mh, ne présenteront pas le phénotype culard (Chelh et al. 2009). La pénétrance incomplète peut s'expliquer par la présence de variants situés dans des régions régulatrices, par des phénomènes épigénétiques, ou encore par des facteurs environnementaux (Kingdom, Wright 2022).

1.1.2. Caractères à déterminisme polygénique

Les caractères à déterminisme polygénique (encore appelés caractères quantitatifs ou caractères complexes) sont des caractères influencés par plusieurs gènes et par le milieu. Chaque effet individuel d'un allèle muté est à lui seul insuffisant pour induire un phénotype quantifiable. En revanche, la présence simultanée de plusieurs allèles mutés chez un même individu conduit à l'apparition du phénotype.

Comprendre l'architecture génétique d'un caractère contribue à l'amélioration des programmes de sélection, mais nécessite d'autant plus d'informations que le caractère à appréhender est complexe. Pour établir ces connaissances, des approches de cartographie génétique ont largement été utilisées, permettant de mettre en évidence les régions du génome impliquées dans la variation des caractères quantitatifs ; ces régions étant nommées « Quantitative Trait Locus » (QTL). En effet, l'utilisation de marqueurs génétiques dont on connaît l'ordre et la position les uns par rapport aux autres permet de localiser dans le génome un QTL affectant un caractère d'intérêt. L'identification des QTL responsables d'un phénotype a été rendue possible grâce (i) aux développements technologiques permettant d'exploiter et d'analyser un grand nombre de marqueurs simultanément et (ii) au progrès des méthodes de calculs et notamment grâce aux outils bioinformatiques. Les caractères complexes sont notamment étudiés via des études d'association à l'échelle du génome (Genome-Wide Association Study, GWAS) dont le but est de mettre en évidence les variants génomiques associés à un caractère particulier.

Néanmoins, si les méthodes d'analyse des données ne semblent plus être un frein à l'étude de caractères polygéniques, pour certains de ces caractères la limite principale est la quantité de données disponibles comme c'est le cas pour la cryptorchidie chez le porc, qui est une anomalie

congénitale à déterminisme polygénique, mais dont l'étude du déterminisme exact est compliquée par un manque de données (Riquet et al. 2016).

Une comparaison des caractères à déterminisme monogénique, digénique et polygénique est proposée dans la Figure 1. La variabilité au sein de la population pour un caractère monogénique (3 états possibles) est distribuée différemment de la variabilité d'un caractère digénique et a fortiori d'un caractère polygénique. En effet, si l'équilibre de Hardy-Weinberg est vérifié dans la population étudiée, les caractères à déterminisme polygénique vont présenter une distribution continue, tandis que les caractères à déterminisme monogénique vont présenter une distribution discrète.

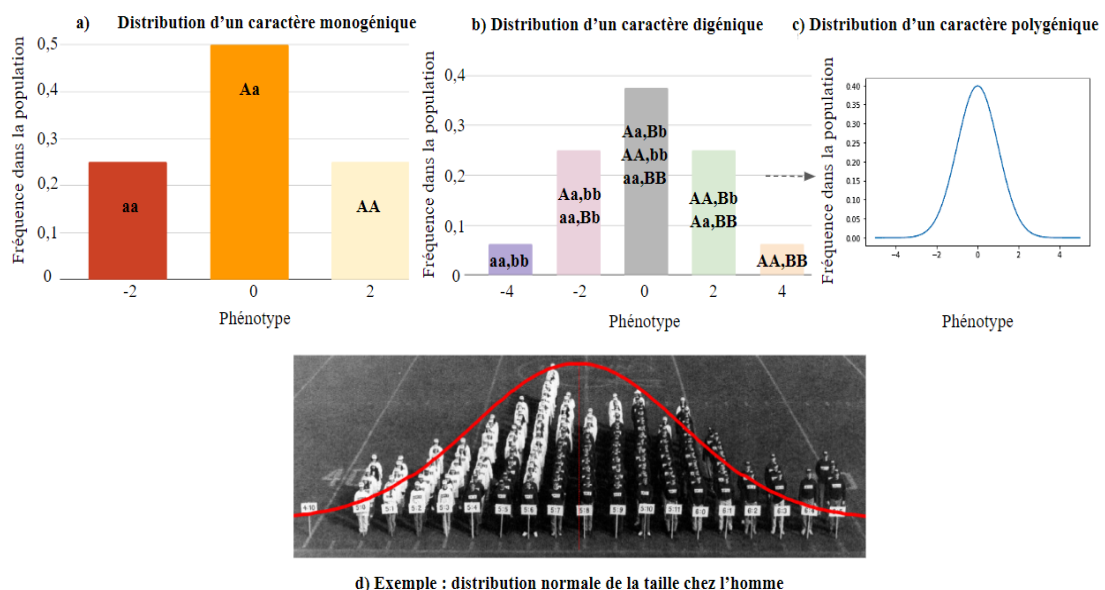


Figure 1 : Représentation schématique de la distribution de la population pour un caractère quantitatif.

a), b) et c) : représentation de la distribution d'un caractère ne présentant pas de relation de dominance entre les allèles dans une population à l'équilibre vérifiant les lois de Hardy-Weinberg selon que son déterminisme soit (a) monogénique (un locus bi-allélique a avec effet(A) = +1 et effet (a) = -1), (b) digénique (deux loci bi-alléliques a et β avec effet(A) = effet (B) +1 et effet (a) = effet(b) = -1) ou (c) polygénique, (d) : Exemple imagé de la distribution normale de la taille chez l'Homme, les femmes étant habillées en blanc et les hommes en noir.

1.1.3. Du modèle monogénique au modèle omnigénique

La frontière entre caractères monogéniques et polygéniques reste floue et est remise en question. Par exemple, chez l'Homme la maladie de l'artère coronaire peut être le résultat d'un déterminisme monogénique ou polygénique selon les patients et selon le phénotype considéré (Jacob, Hegele 2022). Cette maladie aussi appelée hypercholestérolémie familiale est le plus souvent causée par des mutations délétères dans les gènes codants pour le récepteur des lipoprotéines de basse densité, de l'apolipoprotéine B et de la protéine convertase

subtilisine/kexine de type 9. Ces mutations nuisent considérablement à l'élimination des particules de lipoprotéines de basse densité (LDL) de la circulation par les récepteurs, entraînant une exposition à des taux plasmatiques élevés de cholestérol LDL (LDL-C) et, par conséquent, un risque accru de maladie cardiovasculaire athéroscléreuse. Cependant, seule une partie des patients atteints diagnostiqués cliniquement est porteuse d'un variant délétère au sein d'un de ces gènes, donc seule une partie des patients présente un déterminisme monogénique de cette maladie. En effet, il a été reconnu que le phénotype d'hypercholestérolémie familiale peut être polygénique (D'Erasmus et al. 2021; Jacob, Hegele 2022). Ainsi, l'appréciation du déterminisme nécessite une connaissance de plus en plus fine du phénotype.

Le modèle dit « oligogénique », illustré par l'exemple de la coloration du pelage au niveau des extrémités chez le lapin, est un modèle de déterminisme intermédiaire. En effet, il existe des interactions épistatiques impliquant des gènes clefs ayant un effet fort et d'autres gènes ayant de faibles effets (Demars et al. 2022). Un gène dicte si la pigmentation se produit, un autre est responsable de la localisation du dépôt de pigment et des interactions entre quatre autres gènes semblent impliquées dans la gamme de coloration. Ces résultats soulignent l'importance de considérer les allèles impliqués comme une combinaison spécifique et non individuellement. En théorie, tous les gènes du génome pourraient contribuer à un caractère donné, ce qui a conduit à la proposition du modèle à déterminisme omnigénique, qui reste cependant un concept débattu. Ainsi, la représentation des différents types de déterminisme génétique pourrait être vue comme un continuum allant du modèle monogénique au modèle omnigénique (Figure 2).

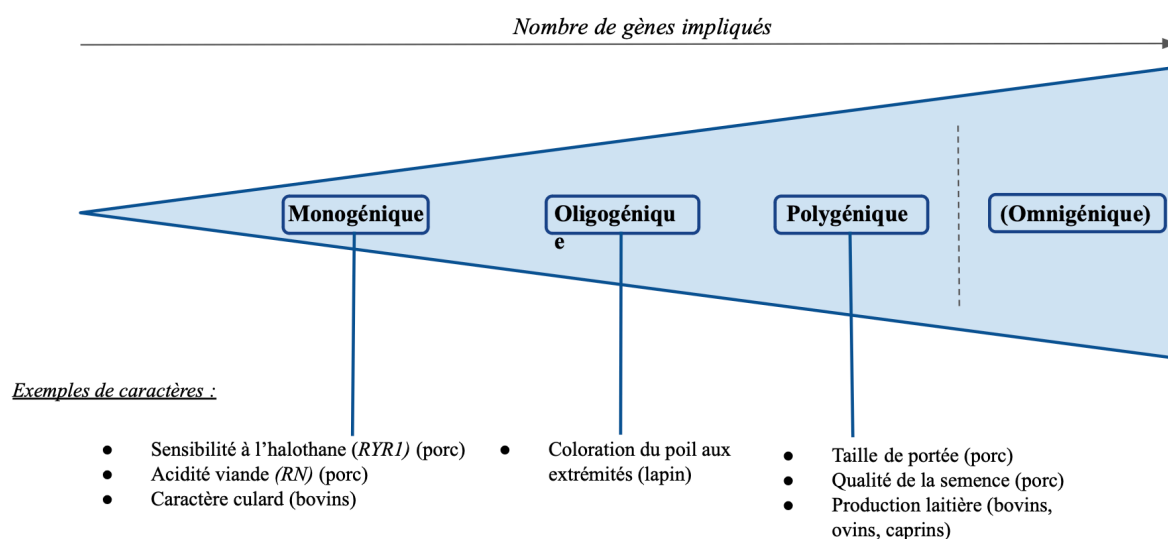


Figure 2 : Un continuum allant du déterminisme monogénique au déterminisme omnigénique.

1.2. L'élevage porcin en France

L'amélioration génétique des populations animales d'élevage a pour but d'obtenir, génération après génération, une descendance présentant des phénotypes meilleurs que ceux de la génération précédente pour certains caractères choisis. Ces phénotypes sélectionnés répondent aux attentes des éleveurs et des autres acteurs de la filière (transformateurs, distributeurs, consommateurs) (Saugère 1989).

Deux stratégies d'amélioration génétique sont utilisées de longue date dans la filière porcine :

(i) La sélection, qui consiste à mettre à la reproduction des individus mâles et femelles qui présentent des caractéristiques particulières (exemple : absence d'un allèle néfaste) ou, plus globalement, qui ont les valeurs génétiques additives les plus favorables pour les caractères pris en compte dans les objectifs de sélection ;

(ii) Le croisement, permettant de bénéficier de l'effet d'hétérosis (supériorité des animaux croisés par rapport à la moyenne des races parentales, due à l'effet non additif de certains gènes) et de l'effet de complémentarité entre races. Par exemple, le croisement est utilisé dans la filière porcine pour obtenir des truies reproductrices hybrides via un croisement des races Large White et Landrace. Dans ce croisement, la race Large White présente l'avantage de produire une viande de bonne qualité et de présenter une bonne productivité numérique ; tandis que la race Landrace est plutôt choisie pour ses qualités maternelles et sa précocité sexuelle (GDS Mayenne 2019).

Nous nous focaliserons dans ce manuscrit sur la sélection.

1.2.1. Historique de la filière porcine française

Les caractéristiques de la filière porcine française actuelle (Roguet 2023) sont le produit d'une évolution qui a été initiée dès les années 1950 (Figure 3).

La période des Trente Glorieuses s'est accompagnée d'une augmentation du niveau de vie moyen de la population française, ce qui a entraîné un changement de comportement, notamment alimentaire, avec une consommation de produits animaux en forte croissance (Hebel 2013). En parallèle, durant les années 1960, la filière porcine a connu une crise liée à la mise en place du marché commun agricole. La concurrence s'est traduite par la disparition de nombreux élevages et par un déficit commercial. Un plan de relance a alors été mis en place et

a permis de restructurer la filière (Ministère de l'agriculture de l'alimentation de la pêche et des affaires rurales 2004).

En 1966, la loi sur l'élevage a participé à la restructuration de l'élevage porcin français en appuyant la mise en place de structures visant à l'amélioration génétique du porc en France. Durant cette période, plusieurs structures d'encadrement de la sélection ont été créées : le comité consultatif de la Commission Nationale d'Amélioration Génétique (CNAG), l'Unité nationale de sélection et de Promotion des Races Agréées (UPRA porcine) en charge de la gestion des livres généalogiques ainsi que des orientations et des modes de sélection notamment, le secteur sélection de L'Institut Français du Porc (IFIP) en charge du contrôle des performances et du contrôle des schémas de sélection notamment, et le département de Génétique Animale de l'Institut National de Recherche pour l'Agriculture (INRA) (BIDANEL et al. 2020). Cette loi a également contribué au développement des dispositifs de contrôle de performances, produisant les données indispensables aux activités de sélection (Ollivier et al. 1986).

Dans le même temps, une organisation pyramidale de l'élevage porcin classique a vu le jour. Cette organisation présente trois niveaux : les sélectionneurs, les multiplicateurs et les producteurs. Si les sélectionneurs étaient au départ des éleveurs indépendants, des groupements de producteurs, tout comme des entreprises privées, se sont progressivement organisés et développés aux dépens des éleveurs indépendants. En 1994, l'Agence de la Sélection Porcine (ASP) a été créée afin d'assurer des missions de promotion de la génétique porcine française (anciennement accomplies par l'UPRA porcine disparue au profit de l'ASP), et des missions d'expertise. L'arrêté du 7 Mars 1994 a modifié assez drastiquement l'organisation de la sélection porcine en donnant naissance aux Organisations de Sélection Porcine (OSP) dont les rôles sont les suivants : sélection de populations porcines, production de reproducteurs hybrides, tenue de livres généalogiques, tenue de registres zootechniques.

Enfin, les évolutions récentes portent sur la structuration des activités de Recherche et Développement de la filière avec la création de structures dont le but est d'unir les compétences et les connaissances. Par exemple, en 2003, les OSP et l'IFIP se sont unis sous la forme d'une association nommée d'abord Bioporc et aujourd'hui Alliance R&D, visant à la mise en place de technologies de génétique moléculaire, de Sélection Assistée par Marqueurs (SAM), puis de la Sélection Génomique (SG) dans les schémas de sélection porcins français (Hiboost 2018; Bidanel et al. 2020). Ces événements historiques de la filière porcine française ont laissé en héritage le système pyramidal qui est toujours d'actualité.

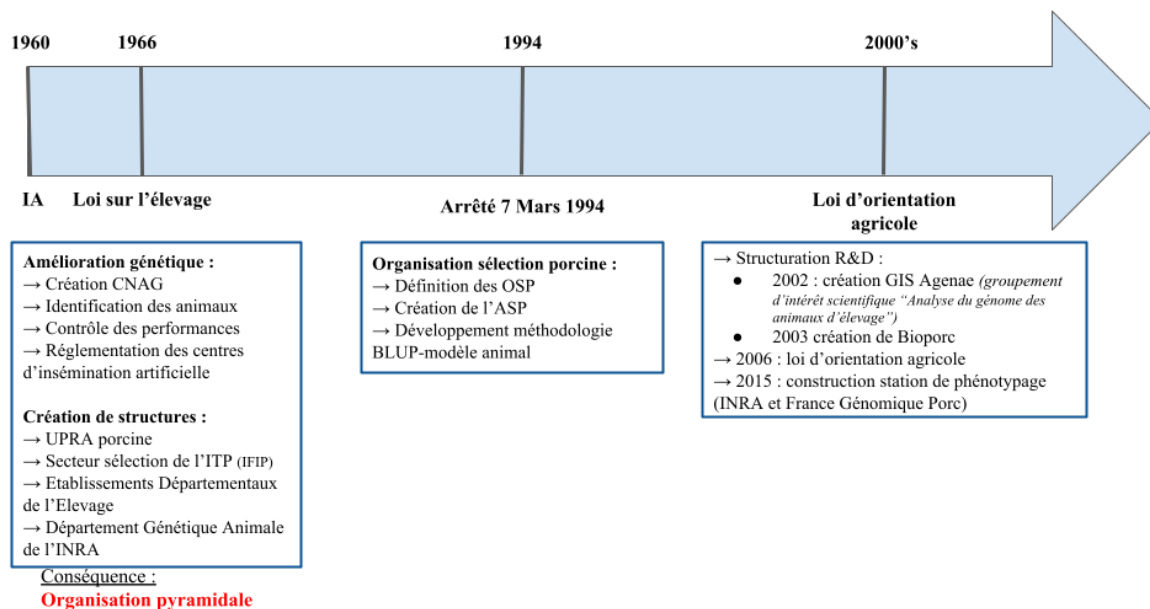


Figure 3 : Synthèse des grandes étapes de la structuration de la filière porcine française.

Données extraites de (Bidanel et al. 2020). IA : Insémination Artificielle ; CNAG : Commission Nationale d'Amélioration Génétique ; UPRA : Unité nationale de sélection et de Promotion des Races Agréées ; ITP (IFIP) : Institut français du Porc ; INRA : Institut National de Recherche pour l'Agriculture ; OSP : Organisation de Sélection Porcine ; ASP : Agence de Sélection Porcine ; BLUP : « Best Linear Unbiased Predictor » ou Meilleure Prédiction Linéaire non Biaisée ; GIS : Groupement d'Intérêt Scientifique.

1.2.2. Organisation de la filière porcine actuelle : un système pyramidal

Actuellement, l'élevage porcin conventionnel français est toujours organisé selon le système pyramidal à trois niveaux : les sélectionneurs, les multiplicateurs, et les producteurs (Figure 4).

Les races (ou variétés de races sélectionnées de manière indépendante par les OSP) communément sélectionnées en élevage conventionnel sont les suivantes : Large White (sélectionnées en voie mâle et en voie femelle), Landrace (voie femelle) et Piétrain (race présentant des qualités bouchères exceptionnelles notamment en termes de musculature, utilisée sur la voie mâle) (GDS Mayenne 2019). D'autres races ou lignées (Duroc, lignées composites sino-européennes, par exemple) sont également utilisées sur les voies mâle ou femelle, mais représentent une part beaucoup plus faible de la génétique présente chez les porcs charcutiers, produits ultimes de la filière (Bidanel et al. 2020).

Cette organisation pyramidale de la filière a permis la généralisation de l'usage des truies hybrides en production puisque le croisement est assuré ainsi par les multiplicateurs (Bidanel et al. 2020).

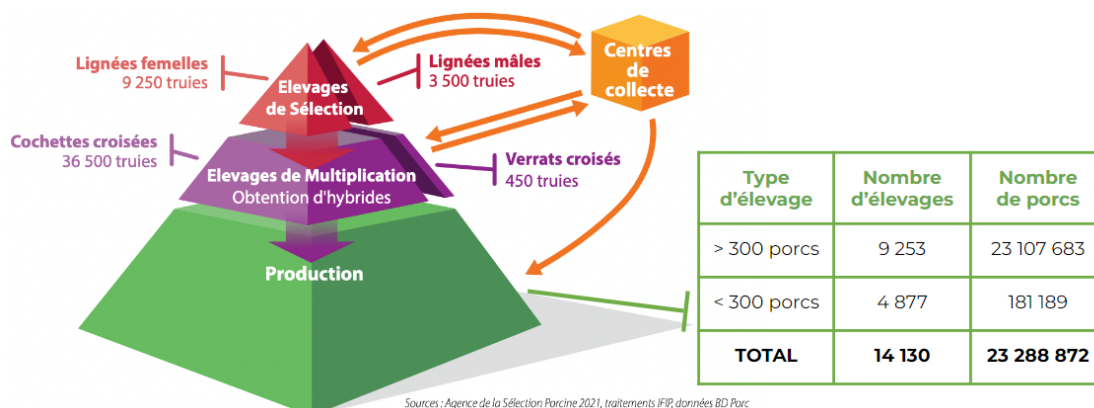


Figure 4 : La filière porcine, un système pyramidal à trois niveaux (Pinel 2023).

Les chiffres présents sur cette figure sont des données de l'année 2021. En 2021, les 9 253 sites (65%) ayant produit plus de 300 porcs (porcelets + porcs à l'engraissement) dans l'année réalisaient 99,2% de la production de porcs charcutiers.

1.2.2.1. L'étage de sélection

La génétique porcine est un outil utilisé principalement par les acteurs situés au sommet de la pyramide, c'est-à-dire les OSP qui sont chargées de choisir la meilleure génétique au sein de lignées pures et qui vont par la suite diffuser cette génétique aux étages en aval. Cette organisation de l'élevage porcin a notamment été rendue possible par la démocratisation de l'insémination artificielle à la fin des années 1980, qui permet aux sélectionneurs de diffuser à bien plus grande échelle leur génétique (Le Roux 2002; Bidanel et al. 2020).

Actuellement, trois OSP « historiques » dominent le marché de la génétique porcine en France : Choice Genetics, Axiom et Nucléus, aujourd'hui aux côtés d'autres OSP comme Topigs-Norsvin, Hypor, ou Pig Improvement Company (PIC) (Ministère de l'Agriculture et de la Souveraineté alimentaire 2022). Ces OSP sélectionnent désormais exclusivement des populations « autonomes » (c'est-à-dire privées), même si certaines ont contribué dans le passé à la sélection de populations collectives inscrites sur des livres généalogiques, résultant d'accords sur l'orientation de la sélection (Bidanel et al. 2020).

1.2.2.2. L'étage de multiplication

Les multiplicateurs produisent des truies hybrides (truies Large White x Landrace par exemple) qui seront utilisées pour la production. L'utilisation de truies hybrides a connu une réelle expansion puisque ces dernières représentaient 36% des truies en 1964, 67% en 1981 et plus de 90 % en 1994) (Bidanel et al. 2020).

1.2.2.3. L'étage de production

L'étage de production regroupe des naisseurs, des engraisseurs et des naisseurs-engraisseurs. A cet étage, les éleveurs sont chargés de produire des porcelets puis des porcs charcutiers qui seront engraisés jusqu'à environ 6 mois et 110 kg de poids vif en moyenne. Les porcelets sont généralement les produits d'une truie Large White x Landrace et d'un verrat Piétrain (ou plus rarement Large White x Piétrain). En 2021, 65% des exploitations (soit 9253 exploitations) ont produit plus de 300 porcs (porcelets et porcs à l'engraissement). Ces exploitations représentent 99,2% de la production de porcs charcutiers. En terme de nombre d'exploitations, la majorité d'entre elles sont engraisseurs, cependant la majorité des porcs charcutiers (57,6%) est produite par les exploitations de type naisseurs-engraisseurs (Pinel 2023). Par ailleurs, concernant la répartition des truies, en 2020 sur les 8448 exploitations agricoles de 100 porcs ou 20 truies et plus, 44% sont naisseurs-engraisseurs et détiennent 79% des truies (Roguet 2023).

1.2.2.4. Une rationalisation au niveau de la filière comme au niveau de l'élevage

Le marché mondial très concurrentiel dans lequel évoluent les éleveurs porcins français les oblige à réduire à (ou à minima maîtriser) leurs coûts de production. Cette réduction des coûts a incité les éleveurs à rationaliser et optimiser l'élevage des animaux, notamment par la mise en place de la conduite en bande (Hebert et al. 2007). Une bande regroupe des individus au même stade (stade physiologique pour les truies, âge ou poids pour le porc en engraissement), et chaque bande est séparée par un écart d'âge constant. Cette méthode d'élevage permet une organisation et une industrialisation du travail dans un but de gain de temps et de maîtrise sanitaire. La pratique la plus répandue depuis les années 80 est la conduite en 7 bandes avec un sevrage à 28 jours (Figure 5).

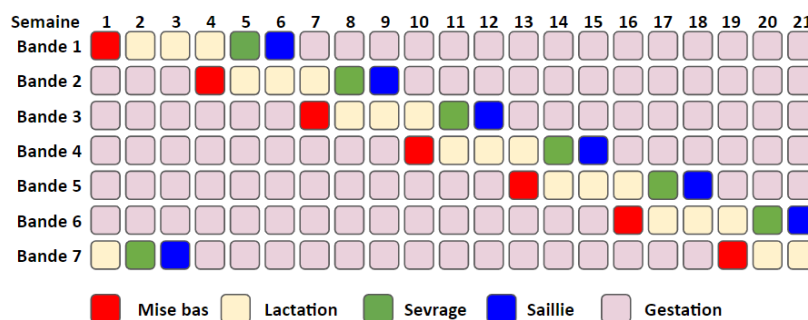


Figure 5 : Organisation de la conduite en 7 bandes avec un sevrage à 28 jours.

Cycle de production de la truie : durée du cycle sexuel : 21 jours ; durée de gestation : 114-117 jours selon les populations ; durée d'allaitement : 21 à 28 jours en élevage conventionnel, au moins 40 jours en agriculture biologique ; intervalle sevrage-œstrus : 5 jours ; durée du cycle de production : 147 jours avec un sevrage à 28 jours ou 140 jours avec un sevrage à 21 jours.

Une autre solution pour faire face à ce contexte économique serait d'optimiser la valeur des animaux élevés, et c'est vers cela que tendent de plus en plus les programmes d'amélioration génétique, dont le but est de fournir aux éleveurs des animaux « optimisés ». Les OSP qui pilotent cette amélioration génétique organisent leurs activités de création de progrès génétique via des méthodes et objectifs de sélection en perpétuelle évolution.

1.2.3. Notions de génétique quantitative

1.2.3.1. Composantes influençant le phénotype d'un individu

Le phénotype d'un individu résulte de l'effet et de l'interaction entre deux composantes : (i) les facteurs environnementaux et (ii) les nombreux gènes contrôlant ces caractères. Ceci peut se traduire par l'équation (1) (Ducrocq 1992; Minvielle 1992; Falconer, Mackay 1996) :

$$(1) : P = G + E + G \times E$$

Avec P la performance, G le potentiel génétique, E les conditions de l'environnement et $G \times E$ l'interaction entre le potentiel génétique et l'environnement.

Jusqu'à peu, le potentiel génétique était considéré comme le seul terme de l'équation se transmettant à la descendance, et a donc été utilisé afin d'estimer les performances d'un individu descendant à partir de celles d'individus parents. Cependant, les connaissances actuelles permettent d'affirmer que le potentiel génétique n'est pas le seul terme de l'équation transmissible à la descendance. Ces éléments feront l'objet de développements spécifiques au Chapitre II.

Le potentiel génétique d'un individu (pour un caractère quelconque) se décompose à son tour en une somme d'effets :

$$(2) : G = A + D + I$$

Avec :

A (potentiel génétique additif ou valeur génétique additive) : \sum effets moyens des gènes présents chez un individu, hérités pour moitié du père et pour moitié de la mère. Dans le cadre de la sélection, seule la valeur génétique additive (A) est utilisée (les « index » calculés et utilisés dans les schémas de sélection sont des prédictions de valeurs génétiques additives). Un individu reproducteur transmettant à chacun de ses descendants la moitié de son génome,

transmet aussi, en espérance, la moitié de sa valeur génétique additive, ce qui constitue le fondement de la sélection ;

D (potentiel génétique de dominance) : \sum effets produits par les interactions entre allèles présents au sein d'un même locus (somme pour les différents loci). Le potentiel génétique de dominance dépend des relations de dominance entre allèles pour les différents loci.

I (potentiel génétique d'interaction) : \sum effets produits par les interactions entre gènes présents à des loci différents.

En combinant les équations (1) et (2) il vient :

$$(3) : P = A + D + I + E + (A + D + I) \times E$$

Décomposer ainsi la performance permet de mieux cibler les leviers d'amélioration génétique d'une population. En effet, comme annoncé dans l'introduction, les deux grandes stratégies pour améliorer une population sont : la sélection (qui vise, pour chaque caractère de l'objectif de sélection, à améliorer la valeur génétique additive moyenne des individus d'une population, génération après génération) et le croisement (qui vise, en accouplant des reproducteurs de races différentes, à reconstituer chez les descendants des associations alléliques originales se traduisant par de forts effets de dominance). Nous nous focaliserons ici sur la sélection, qui s'intéresse à la valeur génétique additive A. Pour plus de précision sur la génétique quantitative, il est possible de se référer aux documents suivants : (Ducrocq 1992; Minvielle 1992).

1.2.3.2. Estimation de la valeur génétique additive

Il a été démontré que la valeur génétique additive A était une bonne approximation du potentiel génétique d'un individu (Hill, Goddard, Visscher 2008). De plus, il s'agit du paramètre le plus facilement estimable parmi les trois affectant le potentiel génétique. C'est pourquoi, dans les modèles actuellement utilisés, G a été approximé par A.

Ainsi, en espérance, et en négligeant les effets d'interaction spécifiques à chaque descendant, on a :

$$(4) : E(G_{descendant}) = \frac{1}{2} (A_{père} + A_{mère})$$

$$A_{descendant} = \frac{1}{2} (A_{père} + A_{mère}) + \omega$$

avec ω = aléa de méiose et $E(\omega) = 0$

S'il est possible d'estimer la valeur génétique additive des parents, il sera possible d'en déduire celle des descendants, qui est une approximation de leur potentiel génétique (Hill, Goddard, Visscher 2008). Il faut donc prédire la valeur génétique additive des candidats à la sélection, qui est aussi appelée « valeur/index de sélection ».

Cependant, comme mentionné précédemment, il n'y a pas que la valeur génétique additive qui contribue aux variations de la valeur phénotypique. Il ne faut en effet pas perdre de vue le fait que cette approche reste une approximation de la valeur génétique d'un individu. Des modèles plus complexes peuvent être proposés dans le but de prendre en compte les autres termes de la décomposition de la valeur phénotypique. Par exemple, les interactions épistatiques affectent la réponse à la sélection en modulant la variance génétique additive utilisée en sélection (Forneris et al. 2017). De façon générale, il est important de savoir dans quelle mesure la variance génétique additive intervient dans la variation phénotypique observée. C'est ce que nous indique un autre paramètre génétique nommé « hérabilité ».

1.2.3.3. Estimation des paramètres génétiques

1.2.3.3.1. L'hérabilité

L'hérabilité mesure la contribution des gènes à la variation des caractères. Elle peut être définie au sens large ou étroit. L'hérabilité au sens large (H^2) est la proportion de la variance phénotypique pouvant être attribuée à la variance des valeurs génotypiques (Baril, Bergonzini 1994; Lynch, Walsh 1998). Avec les notations introduites précédemment, H^2 s'écrit :

$$(4) : H^2 = \frac{\text{var}(G)}{\text{var}(P)}$$

L'hérabilité au sens étroit (h^2) concerne la part attribuable aux seules différences de valeurs génétiques additives entre les individus d'une population (Minvielle 1992). h^2 s'écrit alors :

$$(5) : h^2 = \frac{\text{var}(A)}{\text{var}(P)}$$

Lorsque le terme d'hérabilité est employé sans précision particulière, il désigne généralement l'hérabilité au sens étroit. Défini comme un rapport de variances, l'hérabilité est un paramètre compris entre 0 et 1, qui est estimé pour un caractère et une population donnés, et qui est susceptible d'évoluer dans le temps. Plus l'hérabilité est proche de 1, plus la valeur phénotypique des descendants s'approche de celle des parents (Figure 6). Il s'agit d'un

indicateur de l'efficacité avec laquelle on pourra sélectionner un caractère, et donc des moyens à mettre en œuvre pour sélectionner un caractère.

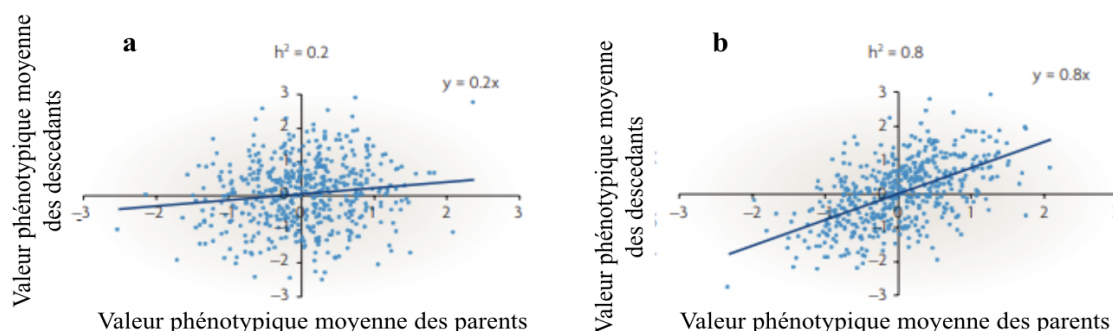


Figure 6 : Estimation de l'héritabilité via la régression parent-enfant.

Extrait de (Visscher, Hill, Wray 2008). La figure montre la valeur phénotypique de descendants en fonction de la valeur phénotypique moyenne de leurs parents. La pente de la droite de régression est une estimation de l'héritabilité pour des caractères présentant une héritabilité de 0,2 (a) et 0,8 (b) et une variance phénotypique de 1. La variance des observations autour de la droite de régression est de 0,98 (a) et 0,68 (b), ce qui démontre que la valeur phénotypique moyenne des parents est un meilleur prédicteur de la valeur phénotypique des descendants lorsque l'héritabilité est grande.

Comme l'illustre le Tableau 1, les caractères de reproduction sont en moyenne peu héréditaires (à l'exception de l'âge à la mise à la reproduction), tandis que les caractères de production sont moyennement héréditaires, et que les caractères de qualité de la viande sont fortement héréditaires (à l'exception du score de qualité).

	Caractère	Héritabilité
Reproduction	Age à la mise à la reproduction	0,37 ± 0,1
	Nombre de porcelets nés vivants dans la première portée	0,10 ± 0,01
	Nombre de porcelets nés vivants dans la seconde portée	
	Intervalle sevrage - premier service après la 1 ^{ère} portée	0,06 ± 0,02
	Intervalle sevrage - premier service après la 2 ^{ème} portée	0,03 ± 0,02
Production	Age ajusté à 100 kg de poids vif	0,30 ± 0,03
	Consommation d'aliments pour une croissance de 25 à 100 kg de poids vif sur les verrats nourris ad libitum	0,22 ± 0,03
Qualité de la viande	Teneur en viande maigre	0,58 ± 0,04
	Épaisseur du lard dorsal	0,44 ± 0,03
	Score : teneur et répartition des matières grasses dans le muscle	0,23 ± 0,03

Tableau 1 : Estimation de l'héritabilité obtenue dans une population norvégienne de Landrace nés entre 1990 et 2000.

Extrait de (Holm et al. 2004).

Pour décider de sélectionner un caractère, il ne suffit pas de prendre en compte son héritabilité. Il faut aussi s'assurer que l'on ne risque pas de sélectionner en même temps un caractère défavorable. On peut apprécier les conséquences d'une sélection conjointe sur plusieurs caractères génétiquement liés en estimant les corrélations génétiques entre ces caractères.

1.2.3.3.2. Les corrélations génétiques

La corrélation entre deux caractères mesure l'intensité et la direction de leur relation, qui dépend notamment de deux phénomènes génétiques. Le premier phénomène est la pléiotropie : un même gène peut intervenir et réguler plusieurs caractères. Le deuxième phénomène est l'existence du déséquilibre de liaison (DL) (Encadré 2), qui rend compte de la transmission non aléatoire de groupes d'allèles (haplotypes) (Figure 7).

Encadré 2 : Le déséquilibre de liaison.

Le déséquilibre de liaison désigne des associations préférentielles entre allèles de deux (plusieurs) loci différents. Lors de la méiose, deux allèles de deux loci différents (on parlera « d'haplotype ») ont tendance à être transmis ensemble s'ils sont à proximité sur le même chromosome (séparés en cas de crossing-over)

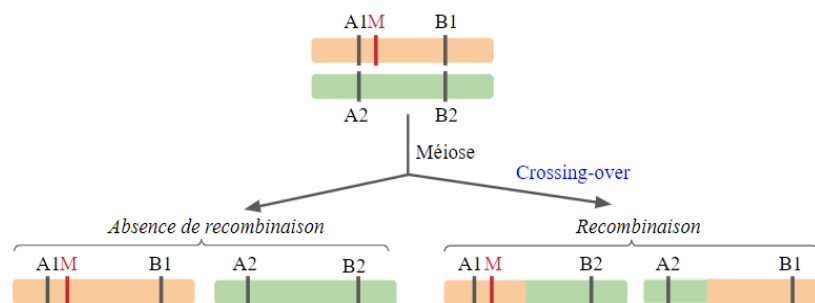


Figure 7 : Déséquilibre de liaison et recombinaison lors de la méiose.

Inspiré de (Boichard et al. 1998a). Soient A1, A2 deux allèles d'un locus A et B1, B2 deux allèles d'un locus B. L'allèle (ou marqueur) M est proche de A1 et en déséquilibre de liaison avec A1. A1 et M sont donc transmis ensemble même lors de la méiose, contrairement à l'allèle B1 qui est sur le même chromosome que A1 et M avant recombinaison mais qui n'est plus sur le même chromosome après recombinaison.

La corrélation génétique permet donc de mesurer les conséquences d'une éventuelle sélection d'un caractère sur les autres caractères qui lui sont liés. Par exemple (Tableau 2), la corrélation entre les caractères « âge à 100 kg de poids vif » et « nombre de porcelets nés vivants dans la première portée » est assez élevée (0,60). Ainsi, le fait de sélectionner pour diminuer l'âge à 100 kg de poids vif implique le risque de diminuer le nombre de porcelets nés vivants dans la première portée. Par ailleurs, la corrélation entre la teneur en viande maigre et le nombre de

porcelets nés vivants dans la première portée est négative, ce qui signifie que la sélection de l'un va entraîner une contre-sélection de l'autre. Ainsi la connaissance de la corrélation génétique entre caractères implique la nécessité de faire des compromis entre les caractères.

Caractère	Age à 100 kg de poids vif	Teneur en viande maigre
Nombre de porcelets nés vivants dans la première portée	0,60 ± 0,05	-0,12 ± 0,07
Intervalle sevrage – premier service après la 1 ^{ère} portée	0,16 ± 0,07	-0,09 ± 0,10

Tableau 2 : Exemples de corrélations génétiques entre caractères (Holm et al. 2004)

1.2.4. Évolution des méthodes de sélection

1.2.4.1. Principes d'amélioration génétique

En élevage, la sélection des meilleurs reproducteurs pour un caractère est organisée selon le schéma présenté sur la Figure 8. Tout d'abord, les objectifs de sélection doivent être définis et varient notamment en fonction des attentes de la société. Puis, il s'agit de vérifier si les caractères ciblés sont effectivement améliorables. Cette vérification nécessite d'estimer différents paramètres génétiques pour le caractère considéré comme l'héritabilité ou les corrélations entre le caractère à améliorer et les autres caractères d'intérêt pour une filière donnée. Si ces caractères sont effectivement améliorables, il faut ensuite mettre en place des dispositifs de contrôle des performances afin de pouvoir réaliser l'évaluation génétique des candidats à la sélection (permettant d'obtenir un index).

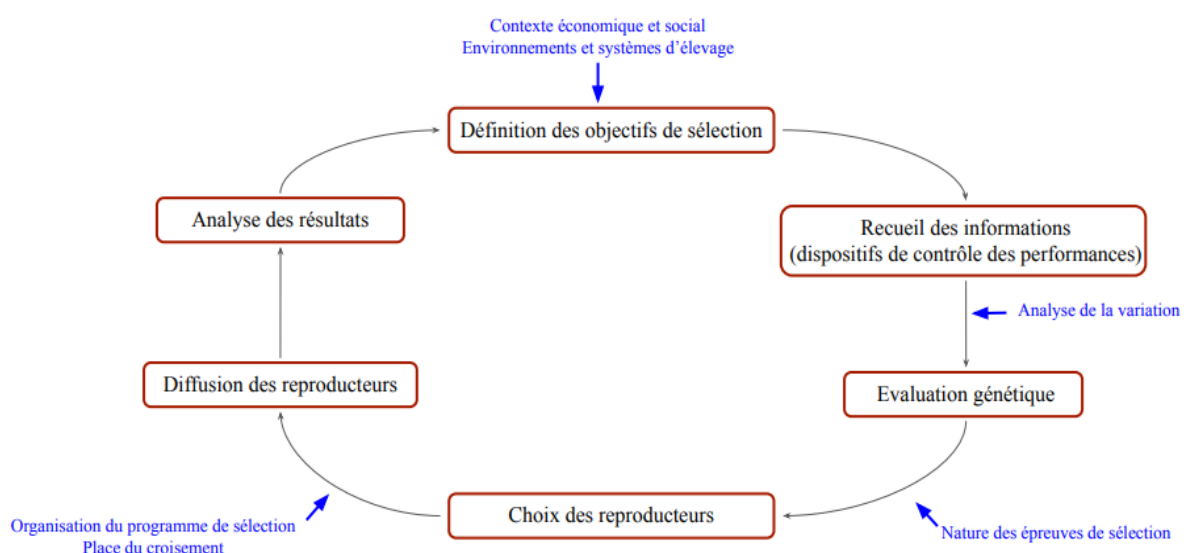


Figure 8 : Étapes d'un schéma de sélection.

Extrait de (Verrier, Milan, Rogel-Gaillard 2020).

Le choix des reproducteurs est effectué par troncature dans la population en choisissant les individus aux meilleurs index pour le caractère sélectionné, permettant ensuite de diffuser la génétique de ces reproducteurs (via l'insémination artificielle notamment). Le gain génétique va dépendre de la pression de sélection appliquée, de la précision avec laquelle les index sont calculés, et de l'héritabilité du caractère.

1.2.4.2. Sélection polygénique classique

Les méthodes de sélection ont évolué avec les progrès scientifiques et techniques dans les domaines de la génétique quantitative et moléculaire, de l'algorithmique et de l'informatique. En effet, entre 1966 et 1994 l'évaluation génétique des reproducteurs s'est tout d'abord appuyée sur la théorie des indices de sélection « classiques ». Jusqu'en 1985, ces indices ne prenaient en compte que les performances propres des individus candidats à la sélection, ou celles de leurs descendants. C'est à partir du milieu des années 1980 que les performances des collatéraux (germains et/ou demi-germains) ont également été prises en compte dans le modèle, permettant de calculer des indices de sélection qualifiés de « combinés ». Cependant, ces indices ne permettaient alors qu'une comparaison intra-bande ou intra-élevage (Ollivier 1986; Bidanel et al. 2020).

1.2.4.3. BLUP-Modèle animal

Les méthodes de sélection ont connu un nouvel élan à partir de 1994 avec le déploiement de la méthodologie « BLUP » (Best Linear Unbiased Predictor ou Meilleure Prédiction Linéaire non Biaisée) associée au Modèle animal (Henderson 1963; 1975). Avec cette nouvelle méthode, il a été possible d'estimer les indices de sélection d'un individu en utilisant les données de performances d'individus collatéraux et les données de généalogie alors qu'aucune donnée de performance n'était connue au sein de sa bande. On est ainsi passé d'une sélection intra-bande ou intra-élevage à une sélection à l'échelle de la population. Cependant, l'obtention d'indices fiables pour un candidat à la sélection via la méthode BLUP-Modèle animal nécessitait de disposer d'un grand nombre de descendants de cet individu, et donc ne peut être réalisée qu'après plusieurs années de vie de l'animal candidat (Phocas 2012). L'efficacité de la sélection a été sensiblement améliorée, notamment pour les caractères les moins héréditaires (prolificité des truies notamment) qui ont pris, à compter de cette date, une importance majeure dans les objectifs de sélection.

Cette révolution a notamment été permise par trois facteurs : la démocratisation de l'insémination artificielle (permettant de créer les indispensables connexions entre bandes, entre élevages), la mise en place d'un système efficace de centralisation rapide des données généalogiques et phénotypiques et l'accroissement des capacités de calcul permise par les progrès informatiques (Bidanel et al. 2020).

1.2.4.4. Introduction à la sélection génomique (SG)

A partir de la fin des années 2010, il est devenu possible de prédire les valeurs génétiques additives des individus candidats à la sélection de façon très précoce grâce à la mise en place de dispositifs d'évaluation génomique. La sélection en filière porcine est passée d'une sélection basée sur des critères phénotypiques et de généalogie à une sélection basée sur des données de génotypage moléculaire. En effet, aujourd'hui, l'information génétique d'un individu peut suffire à elle seule à calculer ses index (dans la mesure où une population de référence existe), c'est-à-dire qu'il est possible d'estimer les index d'un individu dès sa naissance, via un échantillon biologique. La SG sera abordée de façon plus détaillée dans la partie II.

1.2.5. Évolution des objectifs de sélection

Des objectifs de sélection sont établis pour chaque population animale sélectionnée. Ils consistent à faire des choix sur les caractéristiques des individus que l'on souhaite obtenir. Formellement, ces objectifs sont des combinaisons linéaires de caractères, les pondérations correspondant à l'importance économique relative et/ou aux progrès génétiques souhaités pour chacun d'eux (les différents caractères étant liés/corrélés entre eux de façon positive ou négative : ainsi, l'effort de sélection consenti pour un caractère a des conséquences sur l'ensemble des autres).

Les objectifs de sélection concernaient dans un premier temps des caractères de production comme la croissance ou la composition de la carcasse (Figure 9). Puis, dans les années 1980, l'efficacité alimentaire a été intégrée dans les objectifs de sélection, ainsi que les qualités technologiques de la viande (celles-ci ayant été dégradées en raison des critères de sélection précédemment choisis). Le caractère « qualités technologiques de la viande » a été pris en compte via une combinaison linéaire de plusieurs facteurs physico-chimiques impliqués dans la qualité de la viande (pH, couleur, pouvoir de rétention d'eau...). Au même moment, un

deuxième objectif était l'éradication des individus porteurs de l'allèle de sensibilité à l'halothane.

Puis c'est à partir des années 1990 que des objectifs de prolificité (nombre de porcelets nés totaux par exemple) ont été ajoutés au modèle. En parallèle, une séparation progressive des lignées mâles et femelles s'est mise en place dans la race Large White avec des objectifs différents (lignée Large White « femelle » sélectionnée pour les aptitudes maternelles et lignée Large White « mâle » sélectionnée pour les aptitudes bouchères). En raison de la concentration des programmes de sélection sur les caractères de prolificité, les taux de mortinatalité et de mortalité naissance-sevrage se sont accrus, ce qui a poussé les sélectionneurs à modifier les objectifs de sélection pour remplacer le nombre de porcelets nés totaux par le nombre de porcelets nés vivants et le nombre de tétines fonctionnelles dans les critères de sélection des races Large White et Landrace dans les années 2000.

Par la suite, de 2010 à 2014, d'autres critères ont été ajoutés afin d'enrichir ce modèle : le nombre de porcelets sevrés ainsi que le poids moyen à la naissance (Bidanel et al. 2020). Plus récemment, les sélectionneurs se sont intéressés à des critères comportementaux. Par exemple, le taux de survie des porcelets est parfois réduit en raison d'un taux d'écrasement important dans l'espèce porcine. Or, il existe des facteurs génétiques liés à la truie qui peuvent être sélectionnés (Brenaut 2021; Canario 2021).

Les objectifs de sélection évoluent également en fonction des enjeux sociétaux et environnementaux. Par exemple, le bien-être animal est une notion au cœur des attentes actuelles de la société, et la génétique pourrait s'avérer être un levier utile pour s'orienter vers de nouvelles pratiques d'élevage. Par exemple, la castration des porcelets est un sujet controversé. En effet, les hormones sexuelles des verrats sont à l'origine d'une odeur indésirable de la viande qui peut incommoder certains consommateurs. Dans ce contexte, et jusqu'à une période très récente, la majorité des élevages porcins pratiquaient une castration chirurgicale des porcelets mâles (en général durant la première semaine de vie). Pour des raisons économiques, cet acte était le plus souvent réalisé sans prise en charge de la douleur, ce qui soulève des questionnements éthiques. Un arrêté du 24 février 2020 interdit désormais cette castration chirurgicale sans anesthésie et analgésie. Des alternatives à la castration (dont le coût a augmenté avec la nécessité de prendre en charge la douleur des porcelets) ont vu le jour et se sont développées : l'élevage de porcs mâles entiers et la vaccination notamment. La génétique pourrait être un levier dans la mesure où il serait possible d'imaginer une sélection sur ce

caractère d'odeur indésirable (Parois et al. 2018). Les critères de sélection actuels relèvent également de plus en plus des enjeux environnementaux, c'est pourquoi les recherches se tournent de plus en plus vers des critères de robustesse des individus dans l'optique d'une transition agroécologique des systèmes d'élevages (Revilla et al. 2019; Ducos et al. 2021).

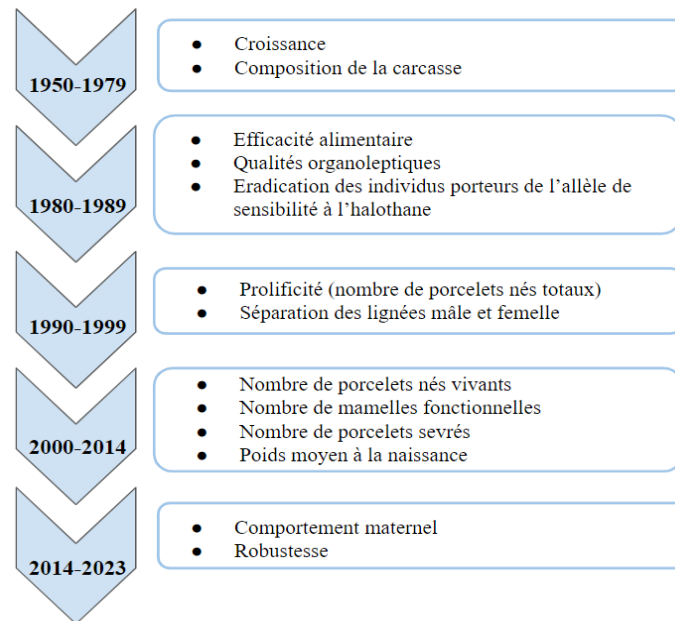


Figure 9 : Évolution des objectifs de sélection en production viande.

Inspiré de (Roy, Ducos, Phocas 2019).

La sélection, au travers des choix qui sont faits concernant les objectifs de sélection, a contribué à faire évoluer les phénotypes des populations porcines utilisées en élevage conventionnel. Par exemple, tandis que l'indice de consommation (rapport entre consommation alimentaire et gain de poids de l'individu) a diminué, le pourcentage de viande maigre et le gain moyen quotidien se sont accrus (Figure 10).

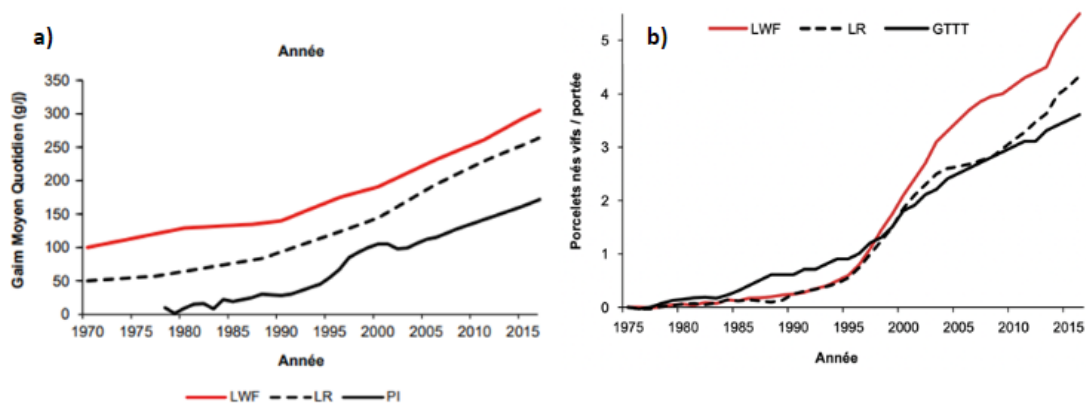


Figure 10 : Évolution au cours du temps des caractères GMQ (a) et porcelets nés vifs par portées (b) dans les races LWF, LR, PI et au niveau de la GTT.

(Bidanel et al. 2020), GMQ = Gain Moyen Quotidien ; LWF = Large White Femelle ; LR = Landrace ; PI = Piétrain ; GTT = Gestion technique des troupeaux de truies.

1.2.6. Vers de nouveaux objectifs de sélection

Durant les cinquante dernières années, la sélection en filière porcine a été réalisée dans le but premier d'optimiser la productivité, dans des environnements très contrôlés (élevage hors sol, conduite en bande suivant un calendrier au jour près) conduisant à une standardisation des animaux utilisés et donc à une perte de diversité génétique intra-population et une dégradation de certains caractères (robustesse, qualités maternelles...) (Roy, Ducos, Phocas 2019). L'utilisation du croisement en filière porcine permet cependant de réduire l'impact de la perte de diversité des races pures.

En France, les consommateurs prennent peu à peu conscience des enjeux environnementaux et éthiques qui concernent l'élevage, et le système porcin conventionnel se retrouve de plus en plus en rupture avec ces enjeux. L'évolution des systèmes d'élevage intègre de plus en plus les attentes des consommateurs et les sujets clés sont, par ordre d'importance : les conditions de vie des animaux et le bien-être animal de manière générale (notamment le développement de l'élevage plein air), la sécurité sanitaire et la qualité des produits, la compétitivité des élevages et la protection de l'environnement (Delanoue 2017). En 2015, un autre sondage visant un public de jeunes adultes suggérait déjà que le bien-être animal était considéré comme l'enjeu majeur pour l'élevage, devant les enjeux environnementaux (Roguet, Delanoue 2015). Par ailleurs, l'intérêt porté par la société au bien-être animal a contribué à influencer le législateur sur certains textes de lois relatifs au bien-être des porcs en élevages : notamment avec l'interdiction de la castration à vif des porcelets prononcée dans l'arrêté du 24 Février 2020 évoqué précédemment. De la même façon, la Commission Européenne a voté le 10 Juin 2021 l'interdiction de l'élevage en cage d'ici 2027 (Parlement Européen 2021). Ainsi, le défi pour les filières de productions animales, et pour la filière porcine en particulier, consiste à concilier les enjeux économiques, environnementaux et sociétaux (Figure 11). Les enjeux environnementaux concernent notamment la limitation des rejets polluants et de l'empreinte carbone de l'élevage. Les enjeux économiques relèvent en premier lieu de la rentabilité des élevages, qui doivent permettre aux éleveurs de dégager des ressources suffisantes pour vivre de la vente de leurs produits, mais concernent aussi l'accessibilité des produits animaux auprès de la population, de manière à permettre d'assurer une alimentation variée à chacun. Par ailleurs, les enjeux sociaux regroupent, comme évoqué précédemment, le bien-être animal qui est une réelle préoccupation des consommateurs, mais également le bien-être au travail des éleveurs qui font face à des évolutions multiples, ainsi que le renouvellement des générations d'éleveurs qui est une problématique majeure (Depoudent, Ponet, Bernard 2023).

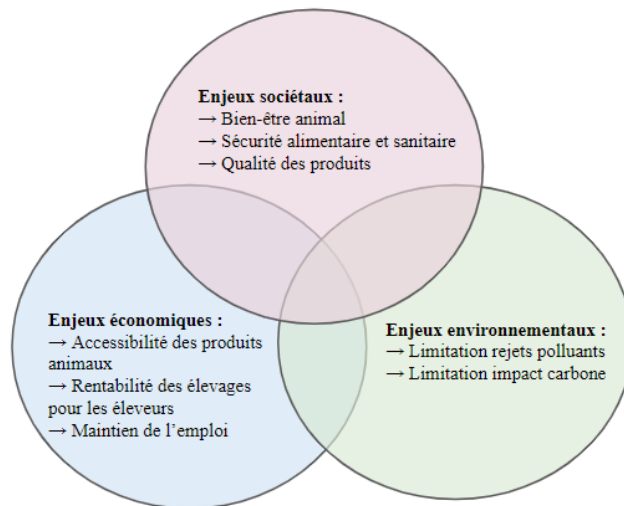


Figure 11 : Les enjeux auxquels le monde de l'élevage doit faire face.

Une diversification des systèmes d'élevage est observée avec l'émergence de systèmes alternatifs comme les élevages labellisés (Label rouge, Indication Géographique Protégée (IGP), Appellation d'Origine Contrôlée (AOC), Appellation d'Origine Protégée (AOP), Agriculture Biologique (AB) ...), les élevages de races locales (porc basque, porc de Bayeux, porc blanc de l'Ouest...). Cependant, ces systèmes alternatifs restent bien moins nombreux que les élevages conventionnels. A titre d'exemple, les produits porcins labellisés AB et ceux labellisés Label Rouge représentaient respectivement 1,2% et 3,7% de la production porcine française en 2021. Leur part de marché pourrait diminuer prochainement comme en témoigne la crise de l'agriculture biologique qui touche également le porc labellisé AB, conséquence du désengagement des consommateurs face à l'inflation (Kressmann 2022).

Les nouveaux systèmes d'élevage sont des systèmes aux environnements moins contrôlés que ne l'est le système conventionnel ce qui implique la nécessité de repenser l'organisation de l'élevage porcin. En effet, ces nouveaux systèmes d'élevage nécessitent des animaux qui soient beaucoup plus résilients. La résilience constitue l'aptitude d'un animal ou d'un système à absorber et/ou répondre à des perturbations, en particulier à des perturbations soudaines et imprévisibles (par exemple de nature nutritionnelle ou sanitaire), en mobilisant des capacités tampon (réallocation temporaire de ressources), d'adaptation ou de transformation (Roy, Ducos, Phocas 2019). De plus, les enjeux d'aujourd'hui et de demain concernent la compétition pour les ressources, le bien-être animal ainsi que l'impact environnemental des élevages. Dans ce contexte, des caractères comme le rejet de gaz à effet de serre sont étudiés afin d'envisager d'orienter la sélection vers ces caractères (Verrier, Milan, Rogel-Gaillard 2020).

L'élevage porcin français, conventionnel (très majoritaire) et alternatif (en développement), se trouve actuellement à un tournant. Afin que l'élevage porcin puisse prendre ce virage de transition agroécologique et corresponde plus aux attentes actuelles de la société, la prise en considération de la génétique et d'autres composantes pourrait être un réel outil d'amélioration (David et al. 2019; Ducos et al. 2021). Bien qu'il semble utopique de penser concilier parfaitement les trois enjeux sociétaux, économiques et environnementaux, l'objectif pourrait être d'essayer de trouver un équilibre acceptable entre ceux-ci.

II. LES RESSOURCES GÉNÉTIQUES COMME SUPPORT DE L'ÉVOLUTION DE LA SÉLECTION PORCINE

2.1. La génomique

2.1.1. Les années 1990 : les marqueurs génétiques

2.1.1.1. Historique de l'utilisation des marqueurs génétiques

Comme évoqué précédemment, les caractères à déterminisme polygénique (constituant la majorité des caractères d'intérêt en génétique animale) sont étudiés via l'utilisation de marqueurs génétiques permettant de cibler des QTL d'intérêts. La SAM a ainsi été proposée en 1990 par Lande et Thompson (Lande, Thompson 1990). L'utilisation de marqueurs pour détecter des QTL est rendue possible par le DL liant le marqueur et le QTL, qui constitue la clé de voûte de la SAM. La SAM a permis l'estimation de la valeur génétique des animaux en utilisant un « score moléculaire », permettant de diminuer l'importance des relations de parenté fondatrices du modèle infinitésimal (exemple BLUP-Modèle animal). En effet, si jusqu'à présent le lien entre génotype et phénotype était réalisé via l'utilisation de données de généalogie et de performances des apparentés, ce lien peut désormais se faire via l'utilisation de marqueurs moléculaires.

L'utilisation des marqueurs est certes indispensable à l'étude des caractères à déterminisme polygénique, cependant elle est également utile pour sélectionner des caractères à déterminisme monogénique dont la mutation causale est encore inconnue. Par exemple, l'éradication de l'allèle de sensibilité à l'halothane de la population de porc Landrace à partir du milieu des années 1980 a constitué l'une des premières applications de SAM chez les animaux d'élevage. Le test à l'halothane ne permettait d'identifier que les homozygotes sensibles. Or, deux marqueurs étaient en fort DL avec le locus HAL dans cette population. Une sélection basée sur

les allèles de ces deux marqueurs a ainsi permis de détecter aussi les hétérozygotes porteurs, et donc de réduire très fortement la fréquence de l'allèle responsable de la sensibilité à l'halothane dans la population Landrace en quelques années (Saugère 1989). La découverte de la mutation causale au début des années 1990 (Fujii et al. 1991) et la mise au point d'un test moléculaire ont ensuite permis d'éliminer les derniers animaux porteurs (Bidanel et al. 2020). Si la SAM a été mise en place dès le début des années 2000 dans quelques races bovines laitières, ce n'est qu'à partir des années 2010 et grâce au développement des puces de génotypage que la SAM a commencé à intéresser la filière porcine (Boichard et al. 2002; Tribout 2011).

L'utilisation des marqueurs n'a pu se développer que grâce aux avancées de la biologie moléculaire. Ces outils ont notamment permis la réalisation de cartes génétiques (Verrier, Milan, Rogel-Gaillard 2020). Les deux principales avancées historiques dans ce domaine, à cette époque, sont la découverte des microsatellites et le développement de la « Polymerase Chain Reaction » (PCR). Jusqu'en 1970 la principale limite à la constitution des cartes génétiques était le manque de marqueurs polymorphes. Bien qu'en 1980 le polymorphisme de longueur des fragments de restriction ait été découvert, leur utilisation à grande échelle s'est révélée complexe car leur analyse était lente et difficilement automatisable. En 1989, la découverte des microsatellites pouvant jouer le rôle de marqueurs a donc été une véritable révolution puisqu'elle a répondu à un réel besoin (Bidanel et al. 1996; Boichard et al. 1998b). Les microsatellites sont des séquences d'ADN constituées de répétitions en tandem de di- ou tri-nucléotides. Ils présentent l'avantage d'être très nombreux, bien répartis sur le génome, et surtout de présenter un polymorphisme important lié à une variation du nombre de répétitions du motif nucléotidique (Boichard et al. 1998b). Ces nouveaux marqueurs vont permettre un enrichissement considérable des cartes génétiques des espèces d'élevage (Moore et al. 1991; Beckman, Weber 1992). Par exemple, en 1993 la carte génétique porcine contenait 30 marqueurs polymorphes et 20 marqueurs microsatellites ont alors été ajoutés augmentant nettement le nombre total de marqueurs (Ellegren et al. 1993).

2.1.1.2. Quelques applications des marqueurs génétiques

La cartographie génétique consiste à établir à partir de données familiales l'ordre et la distance génétique d'une série de loci positionnés physiquement sur le même chromosome. La distance génétique (exprimée en centimorgan (cM)) est proportionnelle au taux de recombinaison existant entre marqueurs polymorphes portés par le même groupe de liaison, au moins pour des distances faibles (1 cM correspond à 1% de recombinaison) (Vaiman 2000). Les cartes

génétiqes ont été très utilisées en sélection mais également pour la publication des premiers génomes, puis leur utilisation a diminué avec la révolution du séquençage. En plus de l'établissement des cartes génétiques, les marqueurs ont pu être utilisés à d'autres fins. Par exemple, ils ont permis d'identifier des gènes majeurs impliqués dans un caractère, l'objectif étant de trouver un marqueur toujours plus proche de ceux-ci (Boichard et al. 1998b).

Ainsi, grâce aux progrès de la biologie moléculaire, la SAM a pu être mise en place. Elle présente l'avantage de rendre possible une sélection précoce des candidats à la reproduction (Chevalet, Boichard 1992). Cependant, les principales limites de la SAM sont liées aux événements de recombinaison qui diminuent le pouvoir prédictif des marqueurs au cours du temps, et dans la nécessité d'un ré-étalonnage régulier qui engendre un certain coût (Elsen 2011).

2.1.2. Les puces de génotypage

Le développement du séquençage a permis la mise en évidence de centaines de milliers puis de millions de marqueurs SNP (« Single Nucleotide Polymorphism »). Le premier génome eucaryote entièrement séquencé est celui de la levure dans le courant des années 90, en combinant les données acquises par un consortium de laboratoires. Le génome de l'Homme a été le premier grand génome complexe séquencé dans le cadre du Projet Génome Humain (PGH), qui a été lancé en 1989 et s'est achevé en 2004, 3 ans après la première publication d'une séquence brute en 2001 obtenue à partir de la méthode de séquençage Sanger (Bidanel, Boichard, Chevalet 2008; Zhang et al. 2011). Puis les génomes des espèces modèles et des espèces d'intérêt agronomique ont été séquencés. Le premier génome d'une espèce agronomique à avoir été rendu public est celui de la poule en 2004 (Roy, Ducos, Phocas 2019). Le génome du porc a lui été séquencé en 2012 (Groenen et al. 2012).

Les SNP sont des polymorphismes qui impliquent le changement d'un nucléotide (Adénine (A), Guanine (G), Thymin (T), Cytosine (C)) en un locus donné. Ils sont de manière générale bi-alléliques (m, M). Trois génotypes sont donc possibles pour chaque SNP, les deux génotypes homozygotes (mm et MM) et le génotype hétérozygote (mM). Les SNP constituent une source d'information riche puisqu'ils apparaissent en moyenne tous les 100 à 1000 nucléotides dans l'ADN chez les mammifères : on peut trouver plusieurs millions de SNP dans le génome d'une même espèce. Leur connaissance permet de caractériser l'état du segment de chromosome où ils se trouvent (Robert-Granié, Legarra, Ducrocq 2011).

Les marqueurs SNP ont alors représenté un grand intérêt pour les entreprises de biotechnologies car il devenait possible de génotyper simultanément plusieurs marqueurs. Les puces à ADN ont alors été développées dans les années 1990 (Schena et al. 1995) avec au début des puces permettant de génotyper 96 SNP en même temps, puis plus récemment plusieurs milliers voire plusieurs millions comme certaines puces utilisées chez l'Homme (Kelly et al. 2023). Dans l'espèce porcine, plusieurs puces à ADN commerciales existent permettant de génotypage d'environ 60 000 marqueurs ou de 600 000 marqueurs. Deux types de puces existent : celles reposant sur le principe d'hybridation et celles reposant sur le principe de la PCR allèle-spécifique (Figure 12) (LaFramboise 2009).

La possibilité de génotyper simultanément des milliers de marqueurs polymorphes de type SNP a conduit à la proposition du concept de la SG par Meuwissen et ses collègues en 2001 (Meuwissen, Hayes, Goddard 2001; Robert-Granié, Legarra, Ducrocq 2011; Roy, Ducos, Phocas 2019).

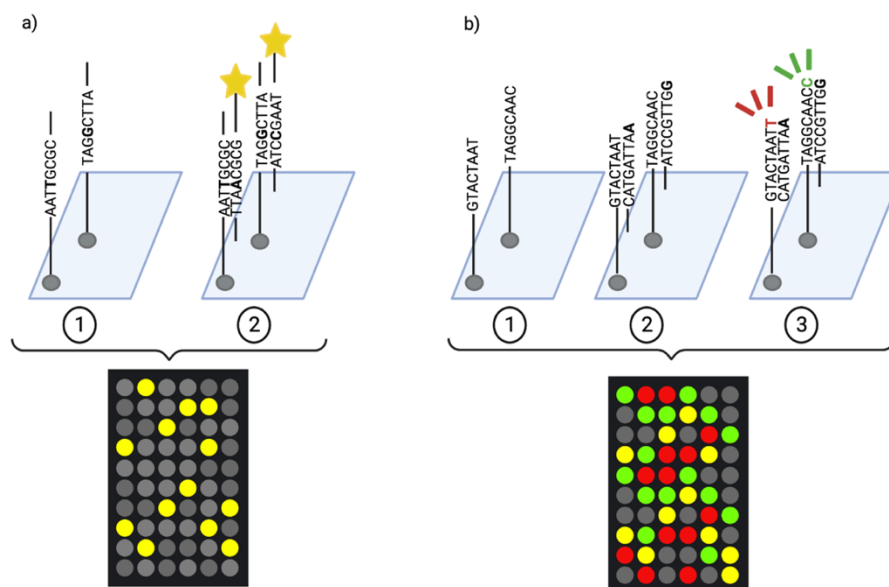


Figure 12: Principes des puces de génotypage.

a) Puce reposant sur le principe d'hybridation : des fragments d'ADN (ou sondes) sont localisés en positions connues sur la puce (a1) ; l'ADN de l'individu (ou cible) à génotyper est fragmenté et marqué par un fluorochrome puis déposé sur la puce. Cet ADN cible est ensuite déposé sur la puce : en cas d'hybridation sonde/cible, un signal lumineux est émis (a2). b) Puce reposant sur le principe de la PCR allèle-spécifique (Illumina [sans date]) : b.1) Sondes en position connue sur la puce, positionnées sur des billes ; b.2) Ajout de l'ADN cible et hybridation sur les sondes correspondantes ; b.3) Ajout de nucléotides marqués (A,T,G,C) et d'ADN polymérase permettant la synthèse du nucléotide manquant, l'ajout de ce nucléotide émet un signal lumineux dont la couleur permet d'identifier ce dernier (la couleur jaune étant la somme de l'émission de vert et de rouge).

2.1.3. En quoi la sélection génomique (SG) représente-elle une révolution technologique ?

La SG a tout d'abord été déployée en France pour l'espèce bovine en 2012 (Boichard et al. 1998b). La différence majeure entre la SAM et la SG repose sur le choix des marqueurs : dans la SAM le choix des marqueurs est biaisé puisque ceux-ci sont en nombre limité et choisis a priori, tandis que dans la SG aucune hypothèse sur le nombre et la localisation des marqueurs n'est faite et toutes les données sont utilisées sans a priori. Le principe de la SG est le même que celui de la SAM. Il s'agit d'exploiter le DL existant entre marqueurs moléculaires et des gènes causaux affectant les caractères à sélectionner. Des équations permettent de prédire les valeurs génétiques à partir des données de performances (phénotype) et de marqueurs (génotype) observés chez les mêmes animaux d'une population de référence. Ces équations peuvent ensuite être appliquées à d'autres animaux sans performances connues afin de prédire la valeur génétique des candidats à la sélection dès leur naissance (en pratique dès que l'on peut extraire l'ADN).

La SG a constitué une véritable révolution technologique puisqu'elle a permis d'augmenter considérablement le nombre de marqueurs pris en compte dans les modèles. Par exemple, à partir de 2008 les puces bovines permettent de génotyper les individus pour 50 000 marqueurs SNP. Ces puces ont permis un développement de la SG à grande échelle dans les principales races bovines laitières en France (Boichard et al. 2012). Concernant la filière porcine, une puce contenant 60 000 SNP a été proposée en 2009. Cependant, la SG ne s'est pas tout de suite mise en place car elle nécessitait une réorganisation complète des méthodes de sélection au sein de la filière (Tribout 2011). Le développement de la SG chez le porc a été facilité par la mise au point de la méthode « Single Step genomic BLUP » (SS-gBLUP) ou BLUP génomique (Legarra et al. 2014; Bidanel et al. 2020). Une évaluation génomique pour les caractères de reproduction a été mise en place en race Landrace en 2016 et en race Large White (lignée femelle) en 2017 (Bouquet 2017; Bidanel et al. 2020).

Grâce à la SAM, la valeur génétique d'un individu peut être estimée dès sa naissance, à partir d'un échantillon biologique. La SG a permis quant à elle d'augmenter considérablement le nombre de marqueurs pris en compte, améliorant la précision de l'évaluation génétique. La SG a également permis d'envisager l'analyse de nouveaux caractères peu héréditaires ou difficilement mesurables (caractères de santé et de bien-être animal, qualité de viande...) (Robert-Granié, Legarra, Ducrocq 2011).

L'utilisation de la SG dans les filières de production animale a notamment été rendue possible par la diminution drastique des coûts de génotypage et de séquençage (Figure 13). Le séquençage du génome humain, à titre d'exemple, a représenté un coût d'environ 3 milliards de dollars (Lander et al. 2001; van Dijk et al. 2018). En 2004, l'Institut national américain de recherche sur le génome humain (NHGRI) a lancé un programme visant à ramener le coût du séquençage du génome entier à 1 000 dollars en 10 ans (Schloss 2008; van Dijk et al. 2018). Ce projet a participé à l'impulsion du développement de nouvelles techniques de séquençage plus rapides et moins coûteuses. En effet, 10 ans plus tard, le coût du séquençage a drastiquement diminué (d'un facteur 100 000) et a effectivement atteint l'objectif visé de 1 000 dollars en 2014 (van Dijk et al. 2014).

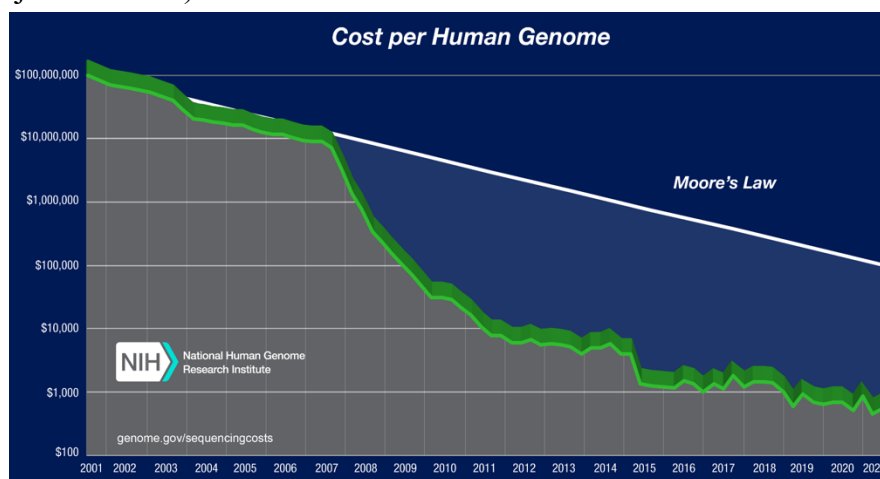


Figure 13 : Évolution du coût du séquençage du génome humain entre 2001 et 2021.

(National Human Genome Research Institute 2023)

2.1.4. Le développement des techniques de séquençage en trois révolutions

2.1.4.1. 1^{ère} révolution : le séquençage Sanger (1977)

Il s'agit d'une méthode de séquençage de l'ADN basée sur l'incorporation sélective de didésoxyribonucléotides (ddNTP, nucléotide sans groupement hydroxyle en position 3') de terminaison de chaîne par l'ADN polymérase (Figure 14). Le mélange comprend l'échantillon d'ADN à séquencer, des amorces spécifiques, de l'ADN polymérase, des désoxyribonucléotides (dNTP : dATP, dCTP, dGTP et dTTP) et un des ddNTP (ddATP, ddCTP, ddGTP et ddTTP). Une réaction d'élongation à partir de l'amorce va ainsi être initiée, jusqu'à l'incorporation aléatoire d'un ddNTP. L'incorporation de ce ddNTP empêche la formation de la liaison phospho-diester reliant les nucléotides d'une séquence et empêche donc la poursuite de l'élongation de l'ADN par l'ADN polymérase. En d'autres termes, un seul type de ddNTP est incorporé par réaction. Des fragments d'ADN de tailles variables sont ainsi

obtenus et peuvent être séparés en fonction de leur taille par électrophorèse. Par exemple, si le ddATP est incorporé, plusieurs fragments de tailles variables seront obtenus, tous se terminant par le nucléotide A (adénine), permettant de déterminer la position des nucléotides A (Sanger, Nicklen, Coulson 1977). Néanmoins, la technologie Sanger présente l'inconvénient de ne pas pouvoir séquencer plus de 2 000 pb simultanément derrière l'amorce (Metzker 2010).

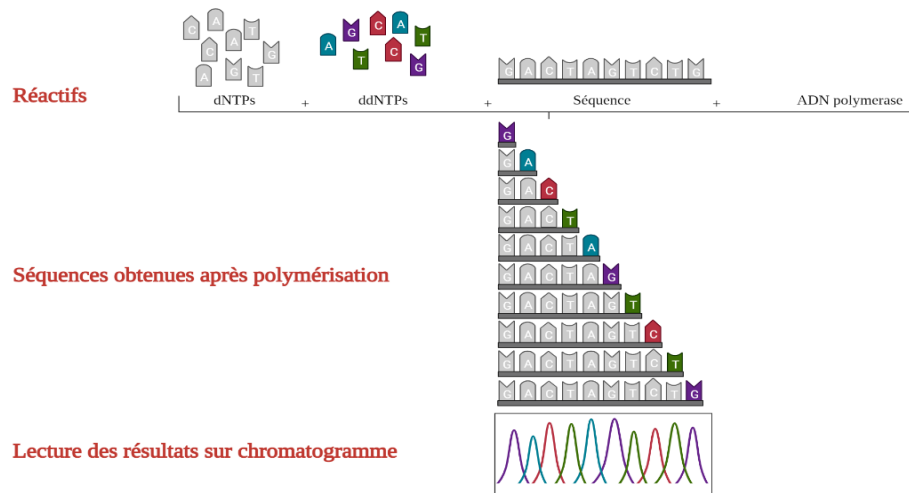


Figure 14 : Principe du séquençage Sanger.

2.1.4.2. 2^{ème} révolution : le séquençage nouvelle génération (ou NGS : Next Generation Sequencing)

Une nouvelle approche a été utilisée de façon transitoire et a constitué les prémices des technologies de séquençage dites de « nouvelle génération ». Cette nouvelle approche est désignée par les termes de « shotgun sequencing » et a été développée sous l'impulsion du Projet Génome Humain. Dans cette approche, l'ADN est segmenté sous action enzymatique ou mécanique pour former de petits fragments qui seront clonés et séquencés au sein de vecteurs. Cette technologie a constitué un socle pour le développement des technologies de séquençage nouvelle génération (Zhang et al. 2011).

Depuis, plusieurs méthodes de séquençage NGS ont vu le jour, s'appuyant sur des technologies d'hybridation ou de synthèse. Celles-ci sont abordées dans la revue de Sarah Goodwin et ses collègues, ou encore dans la revue de Slatko et ses collègues. (Goodwin, McPherson, McCombie 2016; Slatko, Gardner, Ausubel 2018). Alors que le séquençage Sanger permettait de séquencer avec une très bonne qualité des fragments allant jusqu'à 2 000 pb en une expérience, le séquençage nouvelle génération permet lui de séquencer le génome entier en une expérience. Le séquençage nouvelle génération présente également l'avantage d'être plus rapide et que le séquençage Sanger.

Nous détaillerons ici l'une des technologies les plus utilisées : la technologie basée sur le séquençage de lectures courtes proposée par Illumina® reposant sur le séquençage par synthèse. Tout d'abord, l'ADNg est fragmenté par réaction enzymatique ou mécanique en petits fragments. Suite à cette fragmentation, une étape de ligation intervient pour lier des adaptateurs aux extrémités de chacun des petits fragments d'ADNg. Ceci constitue une banque NGS. Les deux fonctions principales des adaptateurs sont : (i) l'accrochage puis l'amplification sur une cellule de flux (de l'anglais « flow cell » : plaque en verre où se déroulent les réactions de séquençage et la visualisation des nucléotides incorporés) et (ii) l'hybridation d'une amorce qui sert de point de départ pour le séquençage. La technologie Illumina® repose sur deux grandes étapes : l'amplification en pont puis le séquençage (Figure 15) (van Dijk, Thermes 2021).

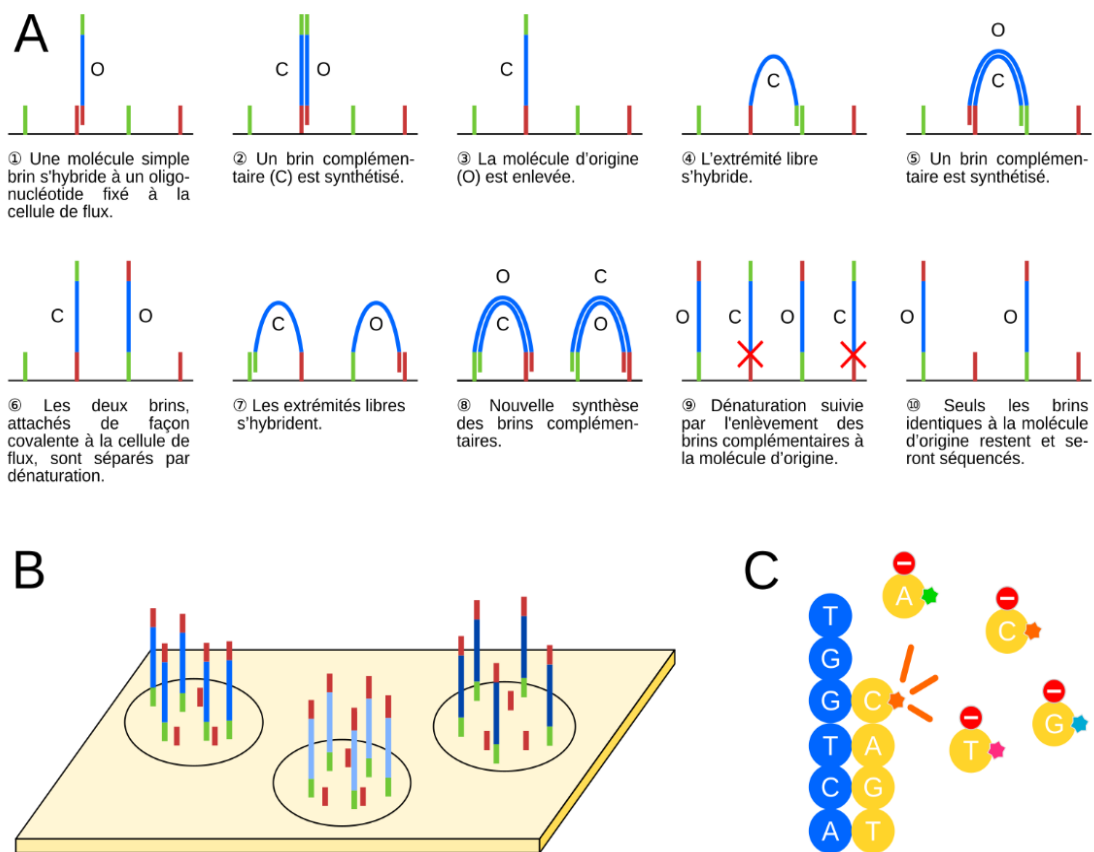


Figure 15 : Principe de la technologie Illumina.

(van Dijk, Thermes 2021). (A) Amplification en pont des clusters sur la cellule de flux (trait noir). Les adaptateurs et les oligonucléotides complémentaires présents sur la cellule de flux sont indiqués en rouge et vert, les fragments d'ADN à séquencer en bleu. La molécule s'étant initialement fixée à la cellule de flux, et les brins dont la séquence est identique, sont notés O ; les brins complémentaires sont notés C. La croix rouge représente le clivage. (B) Différentes molécules d'ADN sont amplifiées puis séquencées en même temps sur une même cellule de flux. (C) Séquençage : une polymérase synthétise un brin complémentaire (jaune) au brin situé sur la cellule de flux (bleu) en incorporant des nucléotides qui portent des groupes fluorescents (bleu, rose, orange, ou vert) et un groupe « stop » qui bloque la polymérisation. S'ensuit une étape d'imagerie afin d'identifier le nucléotide incorporé. Ensuite, les groupes fluorescents et « stop » sont enlevés et le processus se répète.

Par la suite, un traitement bioinformatique va être appliqué en fonction du but recherché : (i) pour construire un génome de référence, un assemblage de novo est nécessaire tandis que pour (ii) identifier des variants ou remaniements, un alignement sur le génome de référence doit être réalisé (les lectures courtes sont mises bout à bout afin de reconstituer la séquence complète). Au fil des années, Illumina a agrandi sa gamme de séquenceurs et propose aujourd’hui des machines qui peuvent générer de 4 millions à 1,2 milliard de séquences par run, avec une longueur de lecture allant de deux fois 25 à deux fois 300 pb (Illumina 2023).

2.1.4.3. 3^{ème} révolution : longues lectures et séquençage de la molécule d’ADN native

L’une des principales limites des technologies de NGS est qu’elles sont basées sur le séquençage de courtes lectures comprises entre 50 pb et 300 pb. Or, les génomes contiennent souvent de nombreuses séquences répétées qui sont plus longues que les lectures NGS, ce qui peut entraîner des défauts d’assemblage et un manque d’information dans certaines régions du génome. L’utilisation de longues lectures permet de faciliter l’assemblage des lectures (Figure 16a). Le séquençage longues lecture est aussi appelé séquençage de troisième génération (van Dijk et al. 2018).

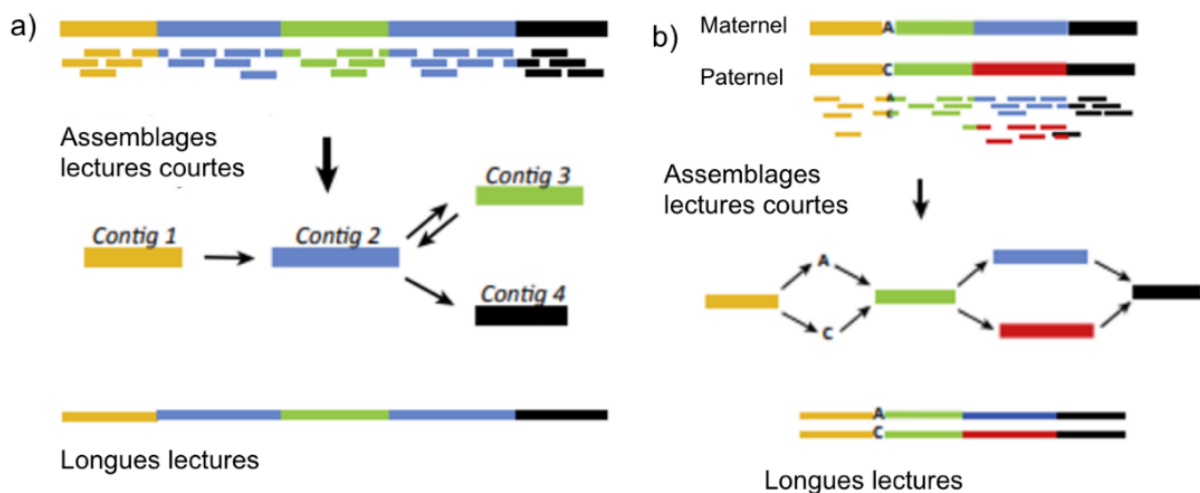


Figure 16 : Avantages des longues lectures pour le NGS.

(van Dijk et al. 2018). a) Exemple montrant la résolution d’une région répétée du génome par assemblage à lecture courte ou lecture longue. Les courtes lectures tombant dans les répétitions (lectures « bleues ») ne peuvent être alignées de façon unique au contraire des lectures longues. b) Reconstruction des haplotypes. Le SNP (A ou C) et les variations plus importantes (rouge ou bleu) entre les allèles des deux haplotypes qui sont trop éloignés pour être couverts par une seule lecture seront difficile à mettre en phase. Cela conduira à des trajectoires ambiguës qui entraînent des assemblages fragmentés.

Le séquençage longue lecture consiste à séquencer la molécule d'ADN native, ce qui permet d'obtenir directement les informations d'haplotype et de phase et donc de traverser des régions d'homozygotie qui, dans le cas des lectures courtes, représentent un obstacle à la détermination des haplotypes (van Dijk et al. 2018) (Figure 16b). Cependant, cette technologie n'est pas infaillible puisque l'obstacle des grandes régions d'homozygotie peut perdurer lorsque la taille de celle-ci dépasse la taille des lectures, même dans le cas des longues lectures. De plus, une des limites des longues lectures réside dans la qualité du séquençage : en effet, le taux d'erreur est plus élevé dans le cas des longues lectures par rapport aux lectures courtes qui ont donc l'avantage de présenter une meilleure qualité. Différentes méthodes de séquençage longues lectures existent, un focus sur l'utilisation d'un nanopore est présentée dans l'Encadré 3.

Encadré 3 : Technologie de séquençage à l'aide d'un nanopore.

L'idée de séquencer à l'aide d'un nanopore n'est pas neuve puisqu'elle date des années 1980 (Deamer, Akeson, Branton 2016). La technologie de séquençage à l'aide d'un nanopore a été utilisée par Oxford Nanopore Technologies (ONT) qui a fourni le premier séquenceur nanopore (MinION) en 2014 (Deamer, Akeson, Branton 2016; Jain et al. 2016). La technologie repose sur un pore protéique d'échelle nanométrique (d'où le nom « nanopore »). Dans une solution électrolytique, une tension constante est appliquée pour produire un courant ionique : ainsi les molécules d'ADN ou ARN chargées négativement sont entraînées à travers le nanopore vers le côté chargé positivement. Le passage de la séquence dans le nanopore entraîne des modifications électriques, chaque nucléotide provoque un signal électrique spécifique ce qui permet, après analyse des données électriques via des algorithmes, de déterminer la séquence nucléotidique (Deamer, Akeson, Branton 2016).

Pour conclure, durant les vingt dernières années une véritable révolution technologique s'est opérée pour l'étude du génome des espèces et du génotype des individus (Figure 17). Le développement d'outils technologiques de plus en plus performants s'est également accompagné d'une baisse des coûts permettant leur diffusion pour des applications comme l'appui à la sélection animale. Cette révolution technologique est toujours en cours, et maintenant que la séquence génomique humaine est considérée comme complète (Nurk et al. 2022), de plus en plus d'études s'intéressent à des variants de natures différentes, non plus « contenus dans » mais « portés par » la séquence, comme c'est le cas des modifications épigénétiques. L'objectif d'ensemble reste le même : accroître notre compréhension des mécanismes moléculaires impliqués dans les caractères complexes.

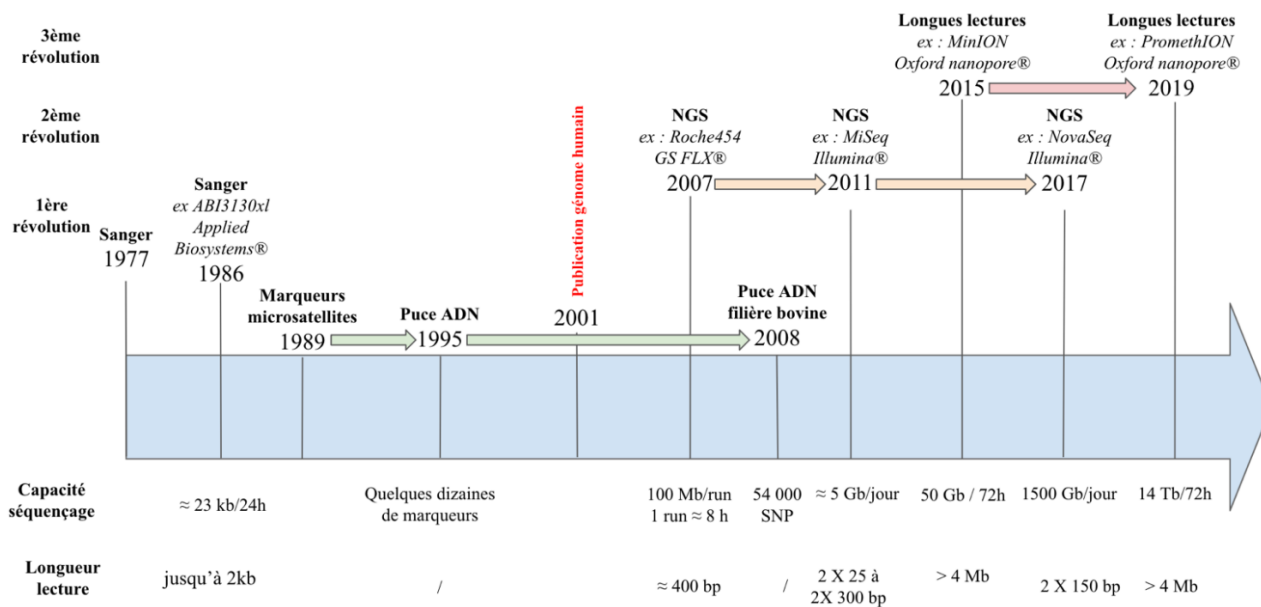


Figure 17 : Schématisation de quelques étapes clés de la révolution technologique dans le domaine de la génétique.

(Laver et al. 2015; Genotoul 2017; Jordan 2017; Roy, Ducos, Phocas 2019; van Dijk, Thermes 2021; Illumina 2021; Oxford Nanopore Technologies).

2.2. Les éléments non mendéliens mais transmissibles

Comme décrit précédemment, l'estimation des performances (P) d'un individu par son seul potentiel génétique (G) s'est révélée très efficace pour la sélection. Cette approximation revient à ne considérer que les éléments transmissibles de façon mendélienne. Or, de nombreux éléments non mendéliens mais transmissibles entre générations influencent le phénotype d'un individu. En vue d'une connaissance et une prédiction toujours plus précise des caractères d'intérêt en élevage, les généticiens s'intéressent de plus en plus à d'autres facteurs contribuant à l'hérédité mais non transmis à partir du matériel génétique lui-même (c'est-à-dire non transmis par la séquence nucléotidique du génome nucléaire). Cette hérédité fait partie du concept « d'hérédité manquante » (Manolio et al. 2009). Elle peut avoir pour support un matériel biologique autre que la séquence nucléotidique (épigénétique (Trerotola et al. 2015; Gayon 2016), microbiote (Sandoval-Motta et al. 2017), génome mitochondrial) ou bien aucun vecteur transmissible d'information a priori (héritage culturel) (David et al. 2019).

2.2.1. Épigenétique

2.2.1.1. Définition de l'épigenétique

Le terme « épigenétique » a été introduit pour la première fois par Waddington en 1942, qui a suggéré l'interaction de facteurs externes avec le génome, et que cette interaction pourrait affecter le développement des individus (Waddington 2009). L'épigenétique a été reconnue comme modulatrice de l'expression des gènes (Hyde, Friso, Choi 2020). En effet, les mécanismes épigenétiques vont participer à la régulation de l'expression des gènes : des marques épigenétiques ont pour effet de modifier l'accessibilité de l'ADN au complexe d'initiation de la transcription, sans pour autant altérer la séquence nucléotidique. Ainsi, l'épigenétique concerne les modifications moléculaires portées par l'ADN qui peuvent réguler l'activité des gènes et qui sont indépendantes de la séquence d'ADN et mitotiquement stables (Li 2021). Selon les définitions, les modifications épigenétiques peuvent comprendre : (1) la méthylation des cytosines au sein des dinucléotides Cytosine-Guanine (CG), appelés CpG (enchaînement de dinucléotides cytosine-guanine, séparés par un phosphate), chez les eucaryotes ; (2) les modifications post-traductionnelles des protéines histones et le remodelage de la chromatine ; et (3) les mécanismes basés sur les petits ARN non codants (Gibney, Nolan 2010). Les marques épigenétiques ont la particularité d'être transmissibles au cours de la mitose voire de la méiose (épigenétique transgénérationnelle), dynamiques et réversibles, contrairement aux modifications de la séquence (mutations).

Les marques épigenétiques sont le reflet de l'impact du milieu sur l'expression génique, elles évoluent au cours du temps. Le milieu correspond à l'environnement dans lequel vit (ou a vécu) l'individu étudié, et comprend notamment l'alimentation ou encore l'exposition à des substances toxiques (Reamon-Buettner, Mutschler, Borlak 2008; Tiffon 2018). L'effet de l'environnement est illustré par les différences phénotypiques observées entre jumeaux monozygotes. En effet, des jumeaux monozygotes partagent la même séquence d'ADN (Li et al. 2014), donc le même génome, aux quelques mutations somatiques post-zygotiques près. Or, ces jumeaux présentent généralement des phénotypes (légèrement) différents. Cette variation phénotypique peut s'expliquer par une expression génique différente entre les jumeaux, en raison de différences dans l'apposition des marques épigenétiques, pouvant résulter de l'effet d'environnements différents (Fraga et al. 2005; Castillo-Fernandez, Spector, Bell 2014). Dans le cas de pathologie comme le cancer du sein, il a été montré qu'une seule des jumelles monozygotiques pouvait développer la maladie en raison de profils épigenétiques différents entre jumelles. Ces jumelles ne diffèrent pas par leur information génétique mais par le taux de

méthylation au niveau du gène *docking protein 7 (DOK7)*. Ainsi, l'hyperméthylation de *DOK7* peut constituer un biomarqueur pouvant permettre d'identifier précocement certains cancers, puisqu'elle survient plusieurs années avant le diagnostic histologique (Heyn et al. 2013). Cet exemple témoigne de l'effet de l'environnement d'un individu sur son épigénome, et donc sur son expression génique et sur son phénotype (Kaminsky et al. 2009).

L'impact de l'environnement sur la mise en place de marques épigénétiques conduisant à une expression différentielle et donc à des phénotypes différents peut aussi être illustré par l'exemple du gène *agouti* chez la souris. Ce gène contribue à déterminer la couleur de la robe. Or, certaines souris possédant toutes le même allèle *Avy* (et donc la même séquence nucléotidique) présentent des pelages de couleurs différentes. Ceci est rendu possible par une expression différente du gène *agouti* en fonction du degré de méthylation d'un élément transposable localisé en amont du gène dans une séquence régulatrice. Lorsque l'élément transposable n'est pas méthylé, le gène *agouti* est fortement exprimé, et exprimé de façon ectopique. Les mélanocytes ne synthétisent alors que de la phéomélanine et le pelage est jaune. En revanche, lorsque l'élément transposable est fortement méthylé, le gène *agouti* est exprimé normalement, les deux types de mélanine (eumélanine et phéomélanine) sont synthétisés et le pelage est brun (*agouti*). Ces modifications épigénétiques sont notamment influencées par l'alimentation de la mère au cours de la gestation, ce qui témoigne de l'influence de l'environnement précoce sur l'épigénome et par conséquent sur l'expression génique et sur le phénotype des individus (Martin, Cropley, Suter 2008).

Un autre exemple illustrant les conséquences des marques épigénétiques sur le phénotype est l'exemple des abeilles mellifères (*Apis mellifera*). En effet, en fonction de l'alimentation reçue par les larves (gelée royale ou non), celles-ci deviendront des reines ou des ouvrières, or ces deux types d'abeilles possèdent des morphologies et comportements très différents (Rasmussen, Amdam 2015).

2.2.1.2. L'épigénétique transgénérationnelle

De nombreuses études se sont intéressées aux effets des modifications épigénétiques chez l'Homme (Battram et al. 2022) et chez les plantes (Cortijo et al. 2014), mais peu ont été réalisées chez des espèces d'intérêt agronomique (Ibeagha-Awemu, Zhao 2015; Khatib 2021). Or, l'épigénétique peut fortement contribuer à l'héritabilité manquante des caractères et donc à affiner les modèles de prédiction. En effet, parmi les modifications épigénétiques, certaines sont

transmissibles à la descendance suivante (épigénétique intergénérationnelle) et sur plusieurs générations successives (épigénétique transgénérationnelle). L'épigénétique transgénérationnelle correspond à la transmission d'information épigénétique via la lignée germinale, en l'absence de toute exposition environnementale directe (Figure 18) (Skinner 2011). L'épigénétique transgénérationnelle permettant une hérédité non mendélienne a tout d'abord été découverte chez les plantes (Brink 1956), puis pour la première fois chez un mammifère en 2006, avec des expériences utilisant la souris comme modèle (Rassoulzadegan et al. 2006). Les mécanismes à l'origine de la transmission entre générations de l'information épigénétique font l'objet de recherches qui permettront de mieux cerner le rôle joué par l'information épigénétique dans le phénotype (Fitz-James, Cavalli 2022). L'empreinte génomique parentale est un cas particulier de transmission d'information épigénétique transgénérationnelle ; ce mécanisme particulier sera détaillé dans le Chapitre III.

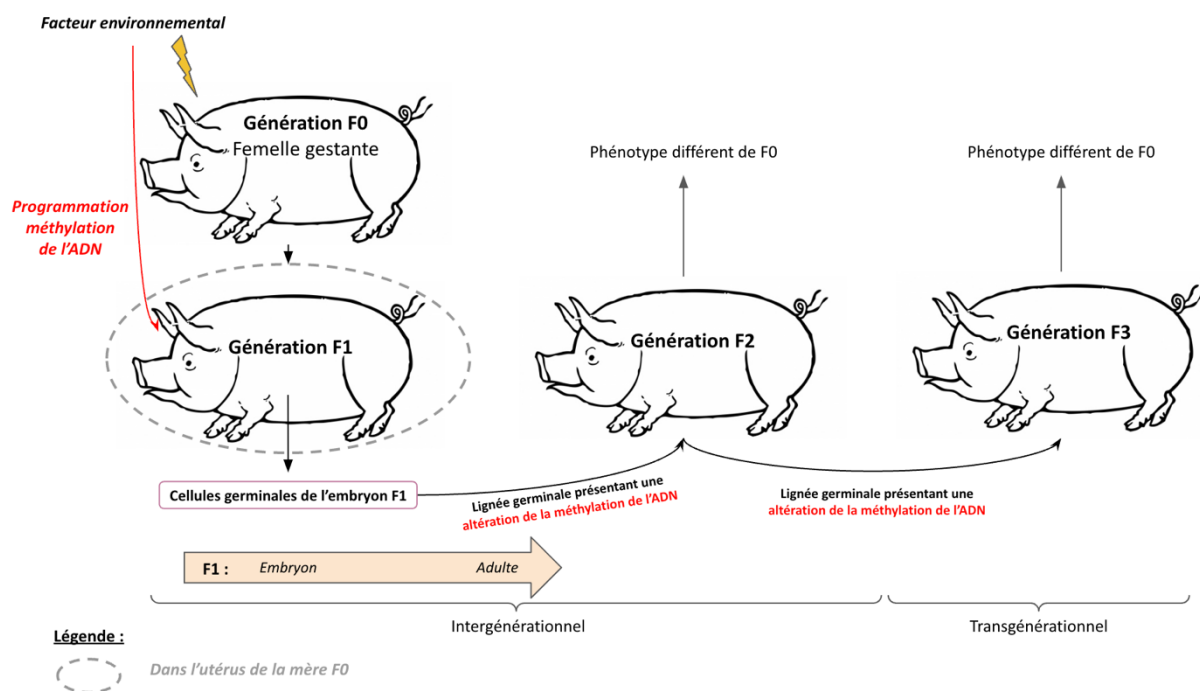


Figure 18 : Concept de l'épigénétique transgénérationnelle.

Inspiré de (Skinner 2011). Lorsqu'un individu (F0) est exposé à une agression environnementale, l'épigénome des cellules somatiques et germinales peut être affecté. De plus, s'il s'agit d'une truie gestante, le fœtus (F1) et ses cellules germinales, qui donneront lieu à une prochaine génération (F2) sont directement touchés. Ainsi, si ces marques épigénétiques persistent chez les générations suivantes (F3 et au-delà si l'exposition a eu lieu chez une truie gestante, F2 et au-delà si l'exposition a eu lieu sur un verrat ou une truie non gestante), il y aura une transmission transgénérationnelle.

Les paramètres de l'environnement qui semblent exercer une influence sur l'épigénome chez des espèces d'intérêt agronomique sont l'alimentation (exemple de l'impact de la nutrition maternelle chez les ovins (Pankey et al. 2021)), la présence de substances épigénotoxiques (exemple de l'effet de la génistéine (perturbateur endocrinien) chez les caillies (Leroux et al. 2017)), ou encore les paramètres environnementaux (exemple de l'influence de la chaleur chez les poulets (Loyau et al. 2016; Pitel et al. 2019)). L'espèce porcine a constitué une espèce d'intérêt agronomique pionnière concernant l'étude de l'épigénétique transgénérationnelle. En effet, le régime alimentaire (standard ou enrichi en micronutriments méthylants) reçu par les verrats de la génération F0, influence les caractéristiques de la carcasse des descendants F2 mais également leur profil d'expression génique ainsi que leur profil de méthylation au niveau du gène *iodotyrosine deiodinase 1 (IYD)* (Braunschweig et al. 2012).

Les données épigénétiques pourraient également servir dans le futur à étudier les facteurs de stress chez les populations animales, en utilisant les marques épigénétiques comme biomarqueurs (Nery da Silva et al. 2021). Les données sur l'épigénome des populations d'élevage intéressent de plus en plus les sélectionneurs puisqu'elles peuvent contribuer à affiner les modèles d'estimation des performances (Goddard, Whitelaw 2014). En raison de cet intérêt croissant pour l'épigénome, des technologies permettant d'évaluer l'état de méthylation ont peu à peu été développées. Ces approches, basées sur la conversion des cytosines par traitement chimique ou enzymatique pour la grande majorité, peuvent s'intéresser à quelques régions d'intérêt ou à tout le génome (Li 2021; Mehrmohamadi et al. 2021; Halabian et al. 2021). A titre d'exemple, des puces ciblant des milliers de dinucléotides CpG sont actuellement mises au point pour des applications agrigénomiques (Arneson et al. 2022). Cependant, la plus grande difficulté qui subsiste pour évaluer l'impact de l'épigénétique sur un phénotype est de discriminer entre l'effet des polymorphismes génétiques et des marques épigénétiques.

2.2.2. Microbiote

Le microbiote désigne l'ensemble des microorganismes vivant en symbiose avec leur hôte. Différents types de microbiotes existent (digestif, respiratoire, cutané...), le terme microbiote (singulier) sera utilisé pour désigner l'ensemble de ces différents types de microbiotes. Parmi eux, le microbiote digestif joue des rôles cruciaux pour l'hôte qui sont bien documentés chez les espèces d'élevage : intervention dans le développement de l'épithélium intestinal,

interaction avec le système immunitaire, régulation du métabolisme, et communication avec le système nerveux (lien entre le microbiote et le comportement) (Calenge et al. 2014). Le microbiote est un domaine de recherche très important chez l'Homme puisque son implication dans de nombreuses pathologies humaines est avérée (Chen, Zhou, Wang 2021). L'ensemble des génomes des microorganismes constituant le microbiote est appelé microbiome. L'espèce porcine est la première espèce pour laquelle un catalogue du microbiome intestinal a été constitué (Xiao et al. 2016). Il s'agit de la troisième espèce à bénéficier d'un tel catalogue après l'Homme et la souris. Le catalogue porcine a récemment été étendu (Chen et al. 2021).

Chez le porc, des études se penchent sur le lien entre le microbiome et les paramètres de production, notamment sur le lien avec le paramètre d'efficacité alimentaire qui pourrait être amélioré via le contrôle du microbiome (Maltecca, Bergamaschi, Tiezzi 2020). Une meilleure connaissance du microbiote porcine offre également une perspective d'amélioration des paramètres de santé, comme la résistance à certaines infections comme les salmonelloses (Argüello et al. 2019). En effet, il est possible que le recours à la transplantation fécale, déjà utilisée chez l'Homme et en développement chez les carnivores domestiques, se déploie aussi dans l'espèce porcine (Barko et al. 2018; Canibe, O'Dea, Abraham 2019).

Le microbiome présente la particularité d'être transmissible verticalement (entre générations) mais également horizontalement (entre les populations) (Estellé 2019). Il est possible que dans le futur le microbiome soit pris en compte dans les modèles de prédiction des phénotypes, et ce pour deux raisons : le microbiome porcine présente une influence directe sur la croissance, et évolue avec son hôte (Larzul et al. 2021; Rogel-Gaillard et al. 2021). La sélection de demain pourrait optimiser les phénotypes par une action conjointe sur l'hôte, son environnement et son microbiote, et ainsi tendre vers des approches holo-omiques et vers la sélection d'holobiontes (entité vivante naturelle constituée d'un organisme supérieur, c'est-à-dire pluricellulaire, appelé hôte, et de son microbiote, c'est-à-dire de la cohorte de microorganismes qui lui est étroitement associée) (Estellé 2019; INRAE 2020; Zhou et al. 2022). Aujourd'hui le domaine de la « métagénomique » est en plein essor car les nouvelles technologies de séquençage permettent de séquencer un microbiome entièrement. En revanche, malgré les possibilités offertes par ces approches prometteuses, des limites telle que les défauts d'annotation des nombreuses espèces de microorganismes doivent être levées (Breitwieser, Lu, Salzberg 2019; Shi et al. 2022).

2.2.3. Génome mitochondrial

Le génome mitochondrial est un génome à hérédité maternelle stricte chez les mammifères, représentant un peu moins de 16,7 kb chez le porc (taille assez standard chez les mammifères), soit moins de 0.001% de la taille du génome nucléaire. Le génome mitochondrial a été séquencé dès 1998 (Ursing, Arnason 1998). Il code pour 13 des 90 sous-unités de la chaîne de transfert d'électrons, qui génèrent de l'ATP par phosphorylation oxydative qui est essentielle pour les cellules ayant des besoins énergétiques élevés, telles que les neurones, les cellules cardiaques et les cellules musculaires squelettiques (Moyes, Battersby, s 1998).

Chez l'Homme, les haplotypes de l'ADN mitochondrial (ADNmt) ont été associés à un certain nombre de maladies telles que le cancer (Shen et al. 2011), le diabète (Liou et al. 2012), avec la maladie d'Alzheimer (Ridge et al. 2012), ou encore avec la maladie de Parkinson (Ghezzi et al. 2005). Chez le porc, le génome mitochondrial influence certains caractères de production puisqu'il existe des associations avec le poids corporel (Yen et al. 2007), et avec la qualité de la viande (Fernández et al. 2008), notamment avec la densité de la graisse, la profondeur des muscles, le rapport entre gras et maigre (St John, Tsai 2018). Il influence également la fonction de reproduction, en intervenant dans la taille de la portée chez les porcs (Tsai, Rajasekar, St John 2016), et semble présenter une relation étroite avec l'ovocyte et l'embryon. Par exemple, le nombre de copies dans l'ovocyte et les variants de l'ADNmt dans l'embryon présentent une influence particulière pour le développement de ces derniers (St John et al. 2022).

A l'heure actuelle, l'ADNmt est principalement utilisé pour des études de diversité génétique chez les espèces d'élevage. Or, il a été montré un lien entre génome mitochondrial et certains phénotypes agronomiques suggérant donc un intérêt potentiel à prendre en considération ce facteur dans équations d'estimation des valeurs phénotypiques. En effet, l'ajout de cette composante permettrait d'affiner un peu plus les modèles de génétique quantitative visant à diminuer progressivement « l'hérédité manquante ».

2.2.4. Héritage culturel

L'héritage culturel fait référence à la transmission sur plusieurs générations de comportements, de culture et de traditions écologiques (Kappeler 2010; David et al. 2019). Il présente la particularité d'être transmis de façon non mendélienne, indépendamment de la transmission génétique et sans vecteur transmissible d'information a priori. Bien que tout un pan de la littérature, principalement en écologie évolutive, s'intéresse à ces concepts, ils ne seront pas

développés ici en détails (Encadré 4). En effet, leur vecteur ou support de transmission ne sont pas quantifiables par des technologies omiques qui constituent le cœur de ce chapitre. L'importance de ces mécanismes pour la sélection chez les espèces d'élevage est abordée dans la revue de (David et al. 2019).

Encadré 4 : Héritages culturel et écologique.

L'héritage culturel correspond à la transmission de compétences sociales et d'innovations au niveau du groupe (David et al. 2019). En effet, au niveau d'un groupe un apprentissage social peut se mettre en place par mimétisme. Cette forme de transmission sociale de l'information permet la diffusion de (nouvelles) connaissances, capacités et traditions culturelles du groupe. Chez les chimpanzés, un exemple d'héritage culturel est le comportement d'ingestion de végétaux permettant de juguler les infestations de parasites digestifs (Huffman, Hirata 2004). L'héritage écologique est la transmission aux descendants de ressources et conditions héritées grâce à la construction de niches (Odling-Smee, Lala, Feldman 2003; David et al. 2019). Contrairement aux espèces sauvages, les espèces d'élevage se développent dans un environnement qui leur est imposé, c'est pourquoi l'héritage écologique pourrait être considéré comme négligeable chez ces espèces.

Actuellement, la sélection des animaux d'élevage est réalisée selon leur potentiel génétique. Des facteurs non génétiques sont aussi transmis de génération en génération (épigénétique, microbiote, génome mitochondrial) et influencent les performances des animaux. Ces facteurs pourraient donc également être utilisés pour sélectionner les futurs reproducteurs (David et al. 2019) et ainsi améliorer les précisions des modèles utilisés. Pour permettre la mise en œuvre de cette stratégie de sélection, un modèle d'analyse des caractères qui prend en compte l'ensemble des facteurs transmissibles (qu'ils soient génétiques ou non) dans l'estimation du potentiel d'un animal à transmettre ses caractéristiques à sa descendance a été développé (David, Ricard 2019). Ce nouveau modèle de transmissibilité, d'ores et déjà appliqué aux données d'efficacité alimentaire de plusieurs espèces, nécessite d'être validé autrement que par des simulations via l'acquisition de données omiques de natures différentes. Parmi elles, des connaissances sur les mécanismes d'empreinte génomique parentale pourraient également être implémentées dans ce nouveau modèle de transmissibilité.

III. LES GENES SOUMIS A EMPREINTE PARENTALE

Les gènes soumis à empreinte parentale sont un cas particulier d'apposition de marques épigénétiques conduisant à une expression génique singulière. En effet, les mammifères sont des organismes diploïdes : chaque gène est présent sous la forme de deux allèles (l'un d'origine maternelle et l'autre d'origine paternelle) susceptibles de donner lieu à une expression bi-allélique ; chaque allèle parental est alors exprimé de façon équivalente. Cependant, pour les gènes soumis à empreinte parentale, seul un des deux allèles parentaux est exprimé : on parle d'expression mono-allélique dépendante de l'origine parentale. L'empreinte parentale est un exemple de régulation épigénétique qui aboutit à une expression parent-dépendante d'un allèle (Figure 19).

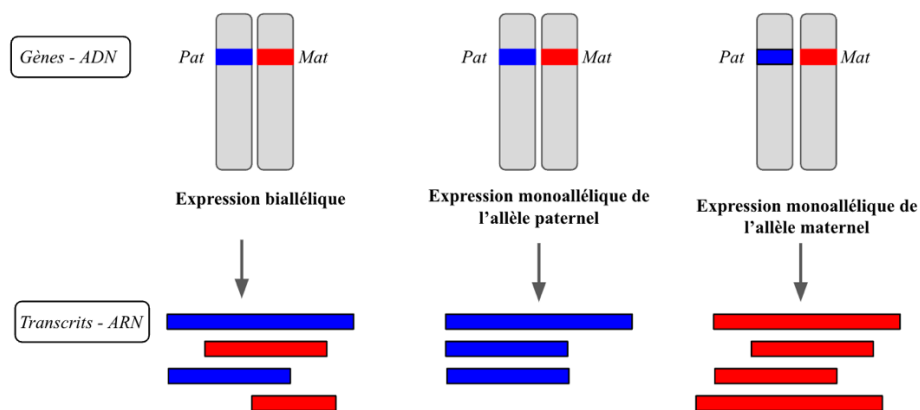


Figure 19 : Schématisation de l'expression mono-allélique parent-dépendante des gènes soumis à empreinte.

3.1. Une expression génique singulière

3.1.1. Découverte des gènes soumis à empreinte parentale

Plusieurs revues sur le sujet de l'empreinte parentale relatent l'historique de cette découverte (Ferguson-Smith, Bourc'his 2018; Tucci et al. 2019).

Tester l'existence de la parthénogenèse chez les mammifères a suscité de l'intérêt dans les années 1980. La parthénogenèse est un mode particulier de reproduction de nombreux invertébrés et de certains végétaux, qui aboutit à la formation d'un nouvel organisme à partir d'un ovule non fécondé (CNRTL). Par exemple, chez les abeilles, la parthénogenèse est possible et donne naissance aux faux bourdons, tandis que les femelles proviennent d'une fécondation classique (Webster 2020). De nombreuses expériences ont eu pour objectif d'étudier la possibilité d'une parthénogenèse chez les mammifères. Dans les différentes études, une anomalie de développement des tissus extra embryonnaires a été observée conduisant à la

conclusion d'une impossibilité de ce mode de reproduction chez les mammifères (Rougier, Werb 2001). Une cellule issue de la parthénogénèse étant une cellule haploïde, l'hypothèse selon laquelle le manque d'une copie du génome pourrait être à l'origine de cet échec de développement a été émise, ce qui a conduit aux expérimentations de Davor Solter.

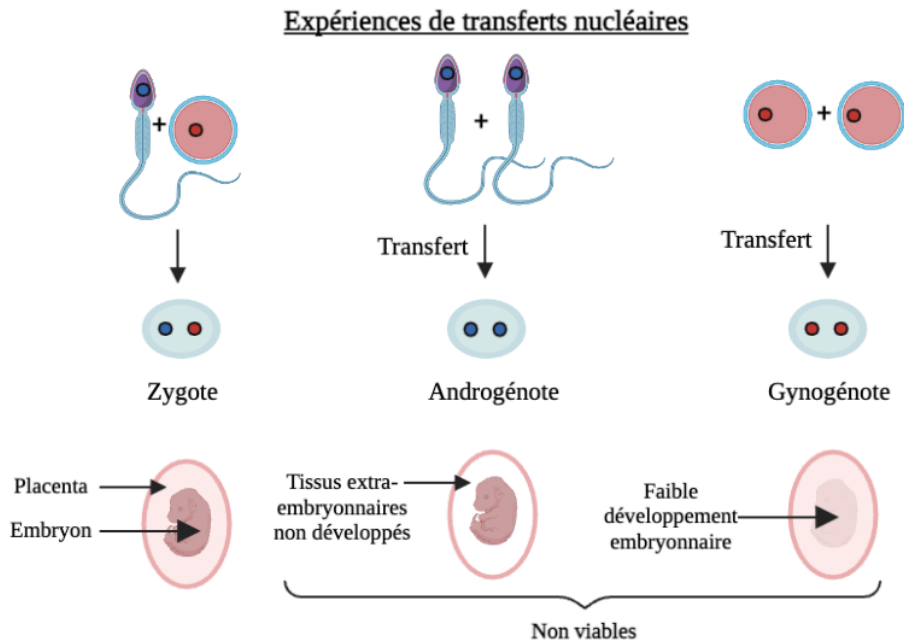


Figure 20 : Expériences de transfert de noyaux (Solter).

Inspirée de (Tucci et al. 2019).

En 1984, Davor Solter réalise une expérimentation reposant sur la technique de transfert nucléaire chez la souris, lui permettant de transférer un noyau de spermatozoïde dans un autre spermatozoïde et un noyau d'ovule dans un autre ovule. Des gynogénotes (zygote avec deux noyaux femelles) et androgénotes (zygote avec deux noyaux mâles) ont ainsi été étudiés. Le développement de ces embryons s'est avéré impossible (Figure 20). Ainsi, une complémentarité des actions des génomes mâle et femelle a été mise en évidence : tous deux ne jouent pas le même rôle dans le développement embryonnaire des mammifères. Deux hypothèses ont alors été émises : soit un (ou plusieurs) gène spécifique est hérité dans une forme fonctionnelle d'un seul parent, tandis que l'autre parent transmet une forme non fonctionnelle de celui-ci ; soit l'expression génique requiert l'interaction entre les génomes paternel et maternel. Au-delà de la mise en évidence de la létalité des gynogénotes et androgénotes, ces expériences ont également permis d'observer les causes de cette létalité. En effet, les gynogénotes sont non viables en raison d'une atteinte extra-embryonnaire tandis que les fœtus androgénotes sont non viables en

raison d'un développement insuffisant du tissu embryonnaire. C'est ainsi que le concept de non-équivalence des génomes paternel et maternel a été abordé puisque ces expérimentations ont permis de mettre en exergue la complémentarité des génomes maternel et paternel (McGrath, Solter 1984; Surani, Barton, Norris 1984).

A ce moment, il n'était pas possible de savoir si cette non-équivalence avait pour objet le génome entier ou seulement certaines régions ou gènes. Cette question a pu être résolue via les expériences de translocation de Cattanach et ses collègues en 1985. Cette expérimentation a consisté à comparer le développement de souris disomiques pour certaines régions. Il a été montré que les souris disomiques pour le chromosome 13 présentaient un phénotype comparable aux autres souris disomiques pour d'autres chromosomes, tandis que les souris disomiques pour le chromosome 11 présentaient une taille différente des autres. Les souris dont la disomie pour une région du chromosome 11 était d'origine maternelle étaient plus petites, et les souris pour lesquelles la disomie pour une région du chromosome 11 était d'origine paternelle étaient plus grandes que les souris disomiques pour d'autres chromosomes (Figure 21) (Cattanach, Kirk 1985). Ainsi, cette étude a permis de conclure que la non-équivalence des génomes pour le développement était liée à une non-équivalence valable pour certains chromosomes particuliers et non à une non-équivalence de l'entièreté du génome.

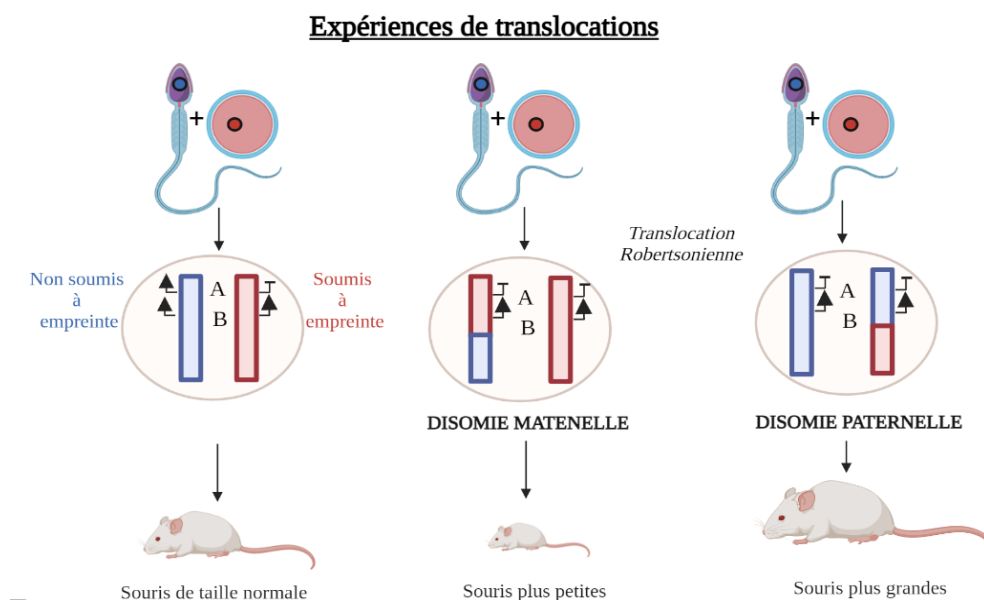


Figure 21 : Expériences de translocation (Cattanach).

Inspirée de (Tucci et al. 2019).

Enfin, les premiers gènes soumis à empreinte ont été mis en évidence dans les années 1990 : *insulin-like growth factor 2 (Igf2)* et *insulin-like growth factor 2 receptor (Igf2r)*. La technique utilisée est une technique de croisements réciproques entre des souris pour lesquelles ces gènes ont été invalidés. Concernant *Igf2r*, dont le locus se situe au niveau du chromosome 17, l'allèle maternel est exprimé tandis que l'allèle paternel est réprimé (Barlow et al. 1991). Concernant *Igf2*, il a été démontré que seul l'allèle paternel de *Igf2* s'exprimait et que l'allèle maternel était silencieux (sauf pour le plexus choroïde et les leptoméniges au sein desquels les deux allèles sont actifs) (DeChiara, Robertson, Efstratiadis 1991). La même année, l'empreinte du gène *H19*, situé en amont du gène *Igf2*, a également été étudiée : pour ce dernier, seule la copie maternelle est transcrite (Bartolomei, Zemel, Tilghman 1991).

Depuis leur découverte dans les années 1980, les gènes soumis à empreinte n'ont cessé de soulever des questionnements de la part de la communauté scientifique. Les connaissances à ce sujet ont été enrichies avec le temps, et notamment la compréhension du cycle des gènes soumis à empreinte et donc de leur mode de transmission entre générations.

3.1.2. Cycle de l'empreinte

Parmi les marques transgénérationnelles, se trouvent les marques d'empreinte qui sont à l'origine d'une expression parentale spécifique des gènes et qui sont communes à tous les individus d'une même espèce. Plusieurs types de marques d'empreinte coexistent, dont les marques de méthylation de l'ADN et les modifications des histones (Barlow, Bartolomei 2014; Inoue et al. 2017; Tucci et al. 2019).

Les marques d'empreinte parentale concernent les cellules somatiques et les cellules germinales. Les cellules somatiques conservent cette empreinte tout au long de leur vie puisqu'elles se divisent par mitose uniquement (étape de « maintenance/lecture » uniquement) tandis que les cellules de la lignée germinale subissent un cycle en trois étapes (Figure 22) (Sha 2008; Tucci et al. 2019) : *(i)* établissement (au sein de la lignée germinale), *(ii)* maintenance (à partir de la fécondation et durant la vie de l'individu) et *(iii)* effacement en vue de la mise en place d'une nouvelle empreinte dans la descendance (au sein des cellules germinales du nouvel individu).

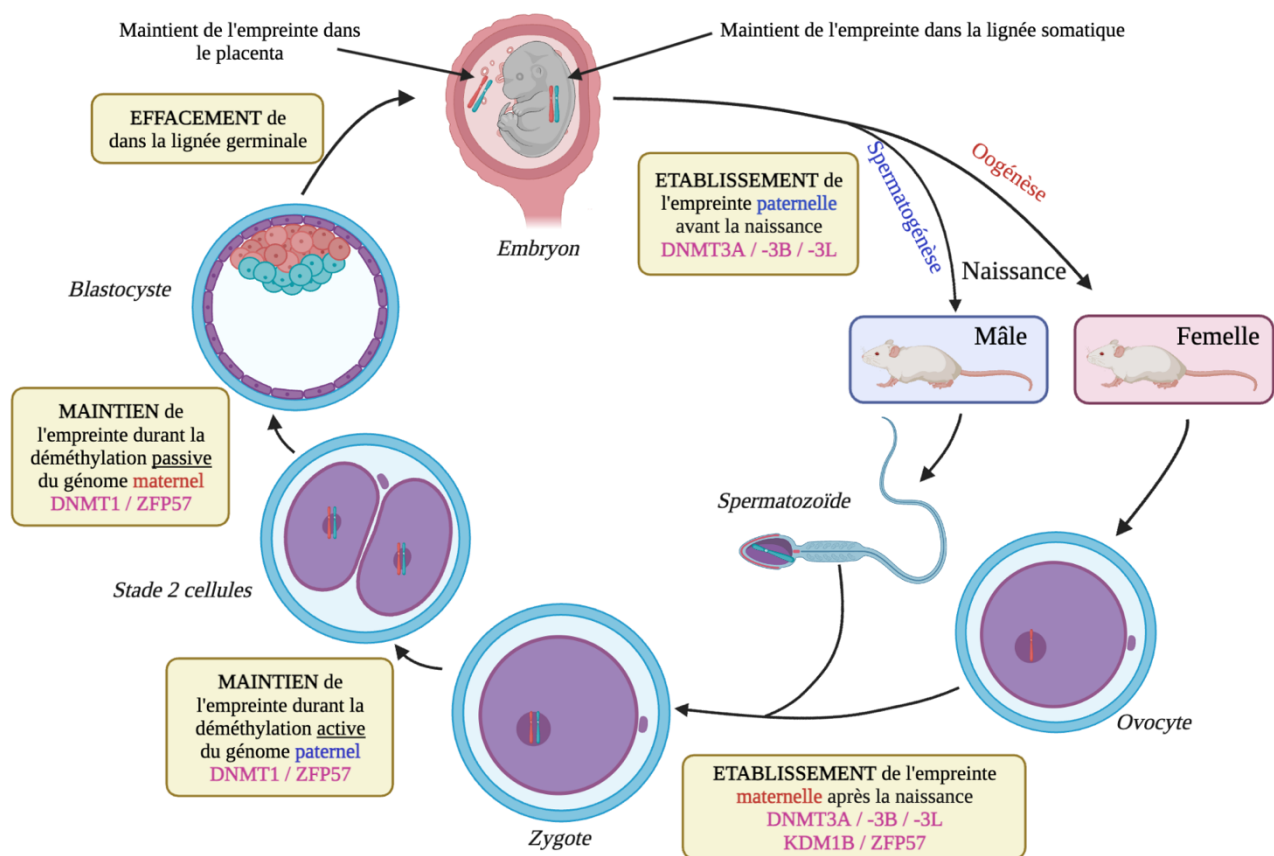


Figure 22 : Cycle de l'empreinte parentale chez la souris.

3.1.2.1. Étape 1 : l'établissement de l'empreinte parentale

L'établissement de l'empreinte se réalise durant la gamétogenèse. Des motifs de méthylation d'ADN spécifiques des gamètes ont été observés (Reik, Walter 2001). Les régions à empreinte maternelle sont méthylées durant la croissance de l'ovocyte, tandis que la méthylation pour l'empreinte paternelle est mise en place plus précocement au stade de prospermatogonie (Figures 22 et 23). La mise en place de la méthylation est permise par des méthyltransférases : les ADN méthyltransférases 3 (DNMT3A, -3B et -3L) ont pour fonction de méthyler l'ADN de novo (dans les cellules germinales, par exemple). DNMT3L est un facteur accessoire important avec une fonction de régulateur de DNMT3A (DNMT3A et DNMT3L forment un tétramère, avec deux molécules de DNMT3A au milieu et DNMT3L de chaque côté) (Zeng, Chen 2019). Par ailleurs, deux autres protéines participent à l'établissement de l'empreinte maternelle de certains gènes : KDM1B (lysine déméthylase 1B) qui est notamment nécessaire pour la mise en place de l'empreinte maternelle des gènes *Peg1/Mest*, *Grb10* ; et la protéine en doigt de zinc 57 (ZFP57), nécessaire à la mise en place de l'empreinte du gène *small nuclear ribonucleoprotein polypeptide N (Snrpn)* (Li, Sasaki 2011)).

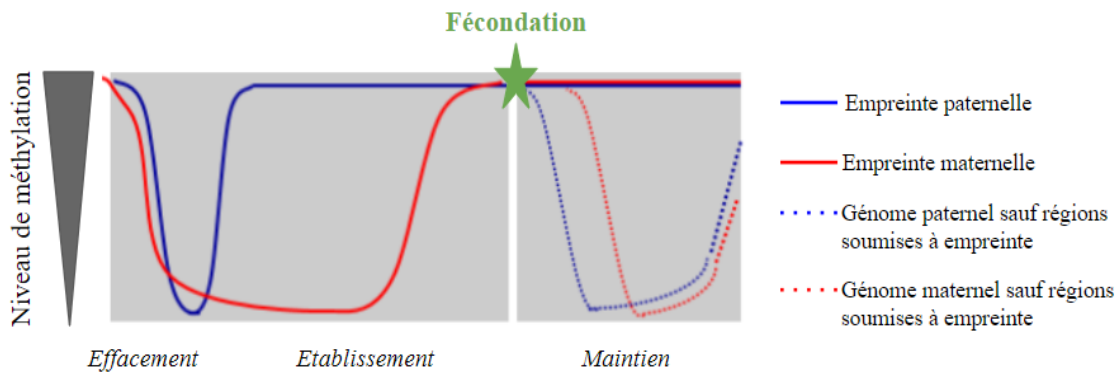


Figure 23 : Variation du niveau de méthylation au cours du cycle de l'empreinte.

3.1.2.2. Étape 2 : le maintien et la lecture

Après la fécondation, les génomes paternel et maternel subissent une déméthylation quasiment complète, à l'exception des régions génomiques soumises à empreinte pour lesquelles la méthylation est maintenue (Figures 22 et 23). La déméthylation du génome peut être active ou passive. La déméthylation active de l'ADN fait référence à un processus enzymatique impliquant des enzymes de translocation 10 ; 11 (ou « ten eleven translocation enzymes », TET) qui éliminent ou modifient le groupe méthyle d'une 5-méthylcytosine (5mC). En revanche, la déméthylation passive de l'ADN fait référence à une perte de 5mC durant les cycles successifs de réplication en l'absence de maintien fonctionnel de la machinerie de méthylation de l'ADN. Tandis que le génome paternel est déméthylé de manière active médiée en partie par l'enzyme TET3, le génome maternel subit une déméthylation passive (Gu et al. 2011 ; Kohli, Zhang 2013 ; Zeng, Chen 2019).

En parallèle, les gènes soumis à empreinte doivent maintenir leur méthylation différentielle face à cette vague de déméthylation. Ce maintien se fait majoritairement par la méthyltransférase DNMT1 (Hirasawa et al. 2008). Pour illustrer le rôle de cette méthyltransférase, une expérimentation a permis d'observer que dans un contexte de perte de fonction de DNMT1, le gène *H19* est exprimé de manière bi-allélique alors que seul l'allèle maternel est normalement transcrit (Li, Beard, Jaenisch 1993). D'autres protéines interviennent dans l'étape de maintien de l'empreinte, notamment ZFP57 (Li et al. 2008) qui va se lier à l'ADN et participer à recruter d'autres facteurs intervenant dans la maintenance de la méthylation (SanMiguel, Bartolomei 2018). Plus récemment, une deuxième protéine en doigts de zinc a été mise en évidence comme intervenant dans l'étape de maintenance : la protéine en doigt de zinc 445 (ZFP445) (Takahashi et al. 2019).

Les tissus extra-embryonnaires et embryonnaires peuvent utiliser des mécanismes différents pour maintenir l'empreinte. Des expériences ont montré que les gènes soumis à empreinte spécifiques du placenta, comme ceux du cluster *potassium voltage-gated channel subfamily Q member 1 (Kcnq1)* (dont *oxysterol binding protein like 5 (Osbp15)*, *tumor suppressing subtransferable candidate 4 (Tssc4)*, *CD81* et *achaete-scute family bHLH transcription factor 2 (Ascl2)*) maintiennent leur statut d'empreinte parentale en l'absence de DNMT1 (Lewis et al. 2004; Umlauf et al. 2004; Weaver et al. 2010). Ces gènes sont marqués de façon différentielle en fonction de l'origine parentale par des modifications d'histones dans le placenta. Ces observations suggèrent que la DNMT1 somatique ne serait pas nécessaire pour le maintien de l'empreinte dans le cas de marques d'empreinte de type modifications post-traductionnelles des histones (Kalish, Jiang, Bartolomei 2014).

Pour une vision plus exhaustive des protéines intervenant dans l'étape de maintenance, il est possible de se référer à la revue de SanMiguel et Bartolomei (SanMiguel, Bartolomei 2018).

3.1.2.3. Étape 3 : l'effacement des marques d'empreinte

Au moment de la gamétogenèse, alors que les cellules somatiques doivent maintenir et exprimer les empreintes parentales, la lignée germinale doit réinitialiser les empreintes de la génération précédente et déposer ses propres empreintes sexe-spécifiques. Comme pour la déméthylation du génome dans son ensemble, la déméthylation des gènes soumis à empreinte peut être une déméthylation active ou passive, cependant le caractère rapide de cette déméthylation suggère qu'il s'agirait plutôt d'une déméthylation active (Hajkova et al. 2002).

Les protéines à l'origine de cette déméthylation sont les protéines de la famille de translocation (TET1, TET2 et TET3, avec une action majoritaire de TET1 dans les cellules germinales (Caldwell, Bartolomei 2022) qui catalysent la conversion de 5mC en 5-hydroxyméthylcytosine (5hmC). Cette conversion est probablement l'étape centrale de la déméthylation et peut conduire par la suite soit à un effacement passif soit à un effacement actif de la méthylation de l'ADN (Hackett, Surani 2013; Caldwell, Bartolomei 2022) (Figure 24). La déméthylation passive se produit par dilution lors de la réplication en raison de l'absence de rétablissement de la méthylation de l'ADN sur le nouveau brin fille après la synthèse de l'ADN. La déméthylation active se produit par l'élimination enzymatique directe d'une modification de l'ADN indépendante de la réplication de l'ADN (Hackett, Surani 2013).

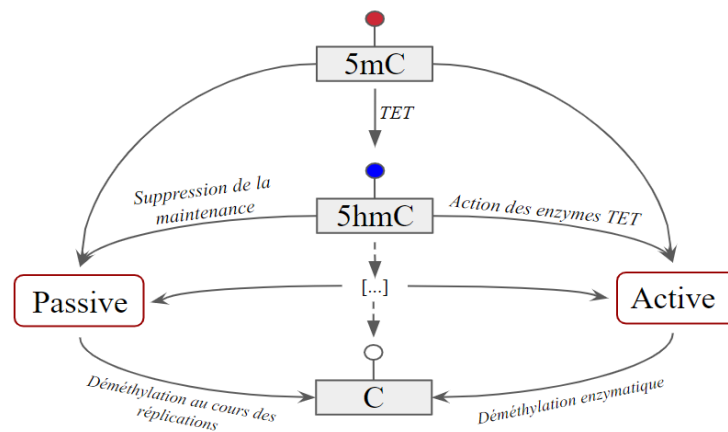


Figure 24 : Les mécanismes de déméthylation : la déméthylation passive versus la déméthylation active via l'action des enzymes TET.

Inspiré de (Hackett, Zylitz, Surani 2012; Hackett, Surani 2013).

Au cours de cette déméthylation au sein de la lignée germinale, les marques de méthylation des gènes soumis à empreinte sont elles aussi effacées, ce qui implique la nécessité d'une forme de « mémoire » de ces marques d'empreinte puisqu'elles subiront à nouveau l'étape d'établissement lors de la formation des gamètes chez l'embryon. Cette mémoire est un point clé de l'empreinte parentale. Le mécanisme moléculaire à l'origine de cette mémoire n'est pas encore complètement élucidé : il est probable qu'elle résulte d'une interaction complexe entre les régulateurs de méthylation de l'ADN, les régulateurs de la chromatine, la transcription et les ARN longs non codants (Gahurova et al. 2017; SanMiguel, Bartolomei 2018).

L'effet transgénérationnel des gènes soumis à empreinte est permis par un cycle de l'empreinte finement régulé par l'intervention de nombreux facteurs (protéines, enzymes, marques épigénétiques...). Cependant, une fois que les marques d'empreinte sont correctement apposées dans la nouvelle lignée germinale, celle-ci devront être correctement régulées afin de conduire à l'expression mono-allélique parent-dépendante caractéristique des gènes soumis à empreinte.

3.1.3. Mécanismes moléculaires impliqués dans la régulation des gènes soumis à empreinte parentale

Les marques d'empreinte sont de nature biochimique et peuvent être : des méthylation de dinucléotides CpG (base de la mise en place de l'empreinte dite canonique), ou des modifications d'histones de type méthylation, acétylation ou encore phosphorylation (base de la mise en place de l'empreinte dite non canonique). Nous nous focaliserons dans ce document

sur les marques de type méthylation, et nous aborderons plus succinctement l’empreinte non canonique.

Concernant la régulation des gènes portés par de l’ADN eucaryote, les promoteurs des gènes abritent généralement des régions riches en CpG. Ces îlots CpG sont aussi des cibles d’apposition de marques épigénétiques comme la méthylation. Ainsi, l’état de méthylation des 5mC module directement la transcription des gènes (Chao, D’Amore 2008). Ces CpG sont réunis en clusters au sein du génome, ou présents seuls pour certains d’entre eux (Bartolomei, Ferguson-Smith 2011).

Il existe des cas où les CpG vont être méthylés de façon particulière : ces régions sont nommées « régions différentiellement méthylées » ou DMR en anglais pour « Differentially Methylated Region ». Dans le cas des gènes soumis à empreinte, ces DMR présentent la particularité d’être différentiellement méthylées en fonction de l’origine parentale. Deux types de DMR se distinguent et correspondent à deux étapes de la mise en place de l’empreinte : les DMR primaires et secondaires. Les DMR primaires acquièrent une méthylation allèle-spécifique durant la gamétogenèse et portent l’information de l’origine parentale. Les DMR secondaires acquièrent une méthylation spécifique de l’origine parentale après la fécondation et peuvent différer entre tissus (Tucci et al. 2019). Enfin, plusieurs gènes soumis à empreinte peuvent être régulés par une même DMR agissant à distance et appelée alors « centre de contrôle d’empreinte » ou ICR en anglais pour « Imprinting Control Region » ; on parle alors de cluster de gènes soumis à empreinte (Barlow, Bartolomei 2014; Ferguson-Smith, Bourc’his 2018).

Des études semblent suggérer qu’il existerait deux grands types de mécanismes moléculaires pour contrôler l’empreinte parentale de plusieurs gènes d’un même cluster : les insulateurs et les longs ARN non codants (ARNlnc). Ces deux mécanismes ont pour point commun d’agir de façon locale en *cis*.

3.1.3.1. Mécanisme insulateur

L’exemple le plus connu est l’exemple du cluster de gènes soumis à empreinte *Igf2/H19*. *Igf2* est exprimé à partir de l’allèle paternel. Il existe un ICR, appelé ICR1, entre les gènes *H19* et *Igf2* qui est méthylé sur l’allèle paternel et non méthylé sur l’allèle maternel (Figure 25). Un enhancer ayant pour rôle de favoriser l’expression de ces deux gènes est situé en aval. La méthylation de ICR1 sur l’allèle paternel empêche la fixation du facteur CTCF. Une boucle chromatinienne peut alors se former aboutissant au rapprochement de l’enhancer et du

promoteur d'*Igf2* permettant ainsi l'expression du gène IGF2 à partir de l'allèle paternel. En revanche, la fixation de CTCF, un insulateur, sur l'allèle maternel isole le gène *Igf2* de ces enhancers ; mais n'empêche pas l'expression du gène *H19* à partir de l'allèle maternel. En effet, une boucle chromatinienne va se former et éloigner l'enhancer d'*Igf2* mais le rapprocher d'*H19* (Engel et al. 2008; Dekker, Misteli 2015). En conclusion, chez la souris, *Igf2* présente une expression de l'allèle paternel tandis que *H19* est transcrit à partir de l'allèle maternel (Chao, D'Amore 2008; Bartolomei, Ferguson-Smith 2011; Barlow, Bartolomei 2014; Ferguson-Smith, Bourc'his 2018).

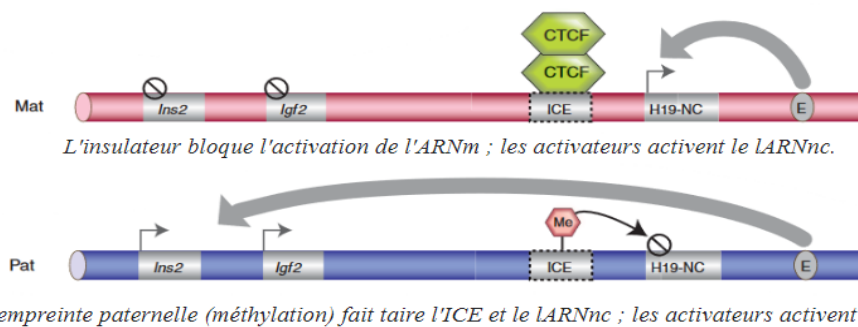


Figure 25 : Régulation épigénétique basée sur le modèle insulateur : exemple des gènes *Igf2* et *H19*.

ICE = « Imprinting Control Element » (équivalent à ICR) ; E = « enhancer » correspondant à un activateur. (Bartolomei, Ferguson-Smith 2011).

3.1.3.2. Mécanisme avec action de longs ARN non codants (ARNlnc)

Un ARNlnc est un transcrit de plus de 200 nucléotides qui ne sera pas traduit en protéine. Pour illustrer le rôle des ARNlnc, il est possible de prendre l'exemple du cluster de gènes soumis à empreinte *Igf2r/Airn* qui présente un ARNlnc impliqué dans l'empreinte parentale de ce locus : l'ARN *Airn* (Figure 26). Au sein de ce cluster, *Airn* est transcrit à partir de l'allèle paternel et agit *en cis* pour réprimer l'expression de *Igf2r*, *Slc22a3* et *Slc22a2* (« solute carrier family 2 member 3/member 2 »). L'ICR de ce cluster est un grand îlot CpG localisé dans un intron du gène *Igf2r*, il est méthylé sur l'allèle maternel et non méthylé sur l'allèle paternel. Or la région promotrice de *Airn* est située au sein de l'ICR donc il n'est transcrit que lorsque l'ICR n'est pas méthylée (donc uniquement à partir de l'allèle paternel).

Dans le placenta, *Airn* exercerait son action répressive via différents mécanismes selon les gènes cibles (Barlow, Bartolomei 2014; El Khattabi 2015) : (i) diminution du recrutement de d'ARN polymérase II fonctionnelle (Latos et al. 2012), (ii) formation d'une enveloppe autour de la chromatine de l'allèle réprimé empêchant sa transcription (Nagano et al. 2008) et (iii) recrutement d'enzymes de modification des histones (conduisant à la méthylation de la lysine

9 de l'histone 3 (H3K9) qui est une modification caractéristique des promoteurs réprimés) (Nagano et al. 2008).

Ce modèle impliquant des ARNlnc s'applique également à d'autres loci soumis à l'empreinte tels que les loci *KCNQ1 antisense (Kcnq1ot1)*, *GNAS*, et *ubiquitin protein ligase E3A (Ube3a)* (Koerner et al. 2009).

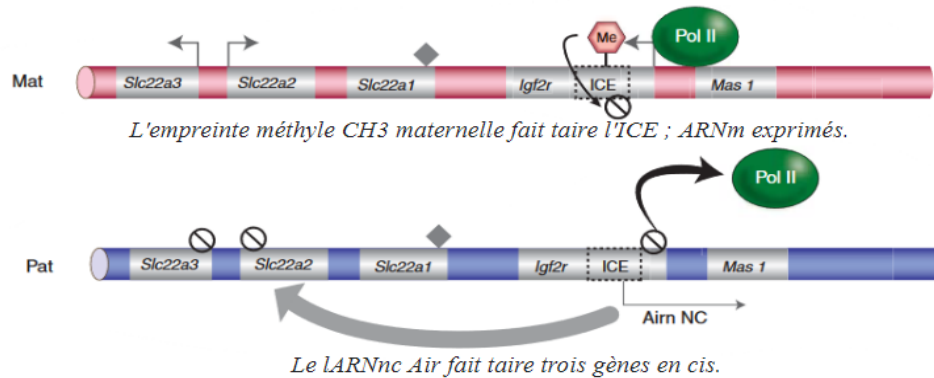


Figure 26 : Régulation épigénétique basée sur le modèle insulateur : exemple du cluster *Igf2r*.

ICE = « Imprinting Control Element » (équivalent à ICR) ; E = « enhancer » correspondant à un activateur. (Bartolomei, Ferguson-Smith 2011).

3.1.3.3. Rôles des petits ARN non codants

Si les ARNlnc peuvent intervenir dans le mécanisme d'action de l'expression parent-spécifique, les petits ARN non codants semblent également importants. Les petits ARN non codants sont classés en 5 familles : les petits ARN nucléolaires (snoARN), les petits ARN nucléaires (snARN), les ARN interférents (siARN), les microARN (miARN) et les ARN interagissant avec les protéines PIWI (piARN). Les protéines PIWI sont responsables de la stabilité du taux de divisions cellulaires dans la lignée germinale. Ces familles diffèrent par leur rôle au sein de la cellule. Les snoARN et snARN sont impliqués dans la maturation des ARN transcrits : les snARN interviennent dans l'épissage des ARNm tandis que les snoARN permettent la maturation des ARNr. Les siARN participent à la dégradation des ARNm tandis que les miARN sont des régulateurs de la traduction. Les piARN sont quant à eux spécialisés dans la protection des génomes contre les éléments transposables (Jammes, Renard 2010).

Les petits ARN non codants impliqués dans l'empreinte parentale sont les snoARN (60 à 300 nucléotides), les piARN et les miARN (22-24 nucléotides). Les régions soumises à empreinte présentent une densité de petits ARN régulateurs (certaines comprenant 50 de ces petits ARN

par mégabase) supérieure aux autres régions du génome. Les snoARN sont impliqués dans la maturation post-transcription des ARN 5, 8S, 16S et 28s (ARNr). Deux groupes de snoARN existent : les C/D snoARN et les H/ACA snoARN. Les C/D snoARN sont impliqués dans la méthylation 2'-O et sont les seuls pour lesquels un contrôle par l’empreinte parentale a été mis en évidence. Les miARN matures guident la protéine Argonaute (AGO2) du complexe RISC (RNA-induced silencing complex) pour diriger la répression post-transcriptionnelle des ARNm cibles (Girardot, Cavaillé, Feil 2012). Certains d’entre eux sont soumis à empreinte, les miARN soumis à empreinte sont majoritairement contenus dans trois régions génomiques distinctes : le cluster C2MC spécifique aux rongeurs (exprimé paternellement situé dans l'intron 10 du gène *Scm like with four mbt domains 2 (Sfmbt2)*), le cluster C19MC spécifique aux primates (paternellement exprimé positionné sur le chromosome 9 chez l’Homme) et le cluster C14MC spécifique aux euthériens (maternellement exprimé situé dans les domaines de *Dlk1-Dio3* chez l’Homme). Ces miARN présentent un rôle dans la placentation. Certains petits ARN de la classe des piARN sont impliqués dans l’acquisition de la méthylation à empreinte parentale (locus *Rasgrf17* chez la souris) (Watanabe et al. 2011). Cependant cette implication dans le cycle de l'empreinte n’est pas généralisable à tous les loci soumis à empreinte ni à toutes les espèces (Girardot, Cavaillé, Feil 2012).

3.1.3.4. L’empreinte non-canonique : un autre type de marque d’empreinte

L’empreinte parentale non canonique est un autre mécanisme de mise en place de l’empreinte parentale ne reposant pas sur la méthylation de l’ADN mais sur des modifications des histones (Kobayashi 2021). Ce mode d’empreinte différent de l'empreinte canonique précédemment décrite a été mise en évidence chez la souris (Okoe et al. 2014). Dans le cas de l’empreinte non canonique, décrite uniquement dans l’ovocyte jusqu’à présent, l’apposition de marques épigénétiques au niveau des histones H3K27me3 (triméthylation de l’histone H3 au niveau de la lysine 27) sur des séquences rétrovirales servirait d’empreinte pour l’apposition post-fécondation de DMR secondaires (Inoue et al. 2017).

Il a longtemps été admis que les mécanismes d’empreinte parentale étaient conservés entre espèces. Cependant les connaissances progressant, des spécificités entre espèces, entre tissus et entre stades de développement ont été observées.

3.1.4. Une expression tissu-dépendante des gènes soumis à empreinte

L'empreinte parentale peut dans certains cas être conservée entre les tissus, mais pour de nombreux loci elle est spécifique d'un tissu. En effet, certains gènes peuvent être soumis à empreinte uniquement dans un tissu, ou bien soumis à empreinte dans divers tissus mais avec une empreinte (maternelle ou paternelle) différente selon le tissu (Prickett, Oakey 2012). Deux organes sont le lieu d'expression de nombreux gènes soumis à empreinte parentale : le placenta et le cerveau. Dans certains cas, l'expression parentale dépendante peut être spécifique de l'un de ses deux organes. Le cas du placenta est développé ici.

3.1.4.1. Le placenta, un organe clef pour les mécanismes d'empreinte parentale

Les mécanismes d'empreinte parentale et le placenta semblent liés par leur moment d'apparition au cours de l'évolution. En effet, l'empreinte parentale a été mise en évidence chez les organismes euthériens et chez les marsupiaux uniquement, c'est-à-dire chez les organismes utilisant la placentation comme mode de gestation (Figure 27) (Hore, Rapkins, Graves 2007).

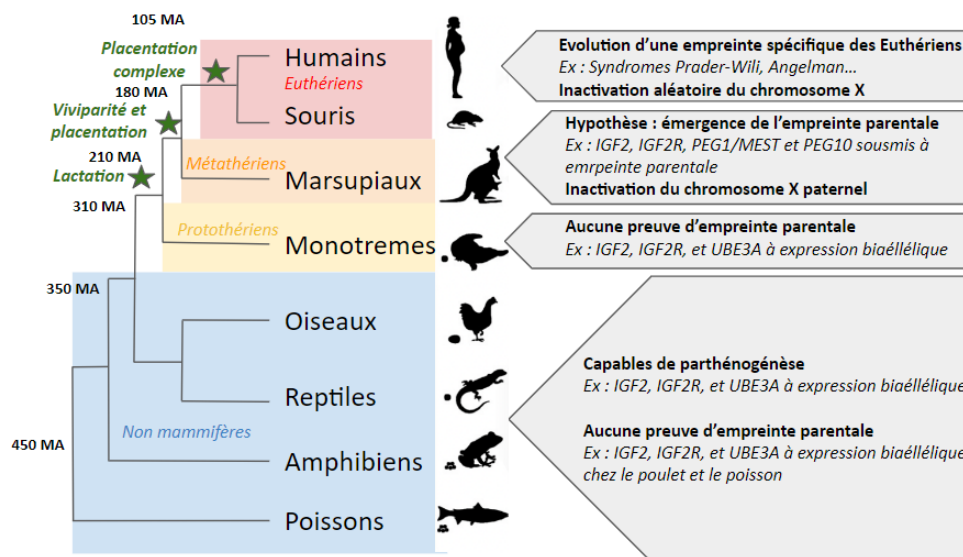


Figure 27 : Comparaison des stratégies de reproduction et de l'existence de l'empreinte génomique chez différents groupes de vertébrés.

Traduit et adapté de (Hore, Rapkins, Graves 2007).

L'empreinte est par ailleurs étudiée chez d'autres vertébrés comme les oiseaux, mais aucune preuve de l'existence d'un mécanisme d'empreinte parentale n'a jusqu'à présent été mis en évidence, cependant cette hypothèse de l'existence d'une empreinte parentale chez d'autres espèces vertébrées non placentaires ne peut être exclue (Frésard et al. 2014). Les gènes soumis

à empreinte semblent donc présenter un lien tout particulier avec la placentation chez les vertébrés. Par ailleurs, l’empreinte parentale existe également chez les plantes angiospermes et certains insectes (Montgomery, Berger 2021).

Plusieurs types de placentation se distinguent : hémochoriale, épithélio-choriale (avec la placentation synépithélio-choriale qui est un sous-type avec présence de syncytia qui sont des masses cellulaires multinucléées se développant parmi les cellules épithéliales utérines), et endo-choriale (Figure 28). Une différence entre notre espèce d’étude (le porc) et les espèces modèles (l’Homme et la souris) est à noter : l’Homme et la souris présentent une placentation hémochoriale tandis que l’espèce porcine présente une placentation épithélio-choriale qui représente une barrière foeto-maternelle plus importante et limite les échanges (Chavatte-Palmer, Tarrade 2016). Le placenta est un organe barrière présentant trois rôles : (i) assurer les échanges de nutriments, gaz sanguins et déchets entre le fœtus et la mère, (ii) assurer la protection du fœtus vis-à-vis du système immunitaire maternel et (iii) assurer diverses productions endocrines (Chemineau 2014).

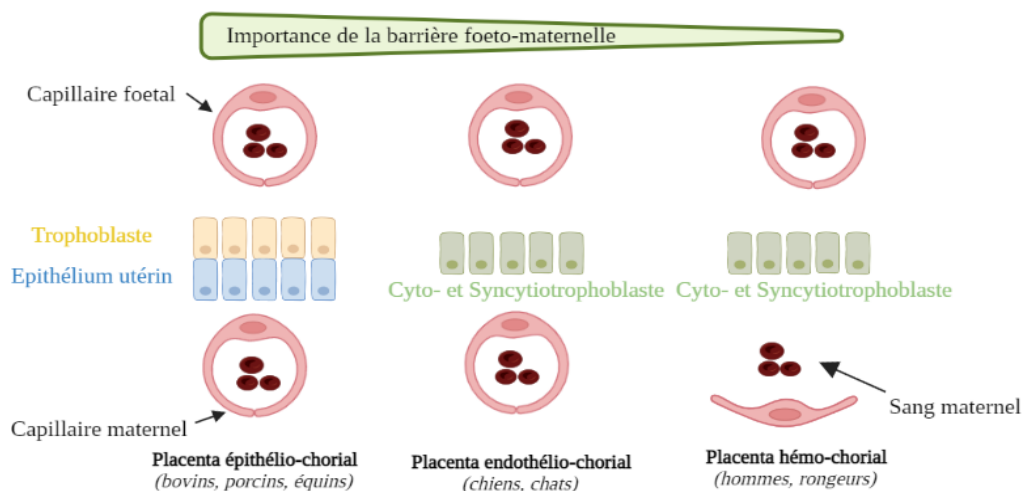


Figure 28 : Comparaison histologique simplifiée des différentes placentations en fonction des espèces.

Inspiré de (Chemineau 2014).

De nombreux gènes soumis à empreinte sont spécifiques du placenta chez l’Homme et la souris, certains sont soumis à empreinte dans le placenta pour les deux espèces tandis que d’autres sont soumis à empreinte pour une espèce et à expression bi-allélique pour l’autre (Monk 2015). Par exemple, le gène *Sfmbt2* est exprimé à partir de la copie paternelle dans le placenta murin tandis que l’expression est bi-allélique dans placenta humain (Miri et al. 2013).

3.1.4.2. Exemple du gène growth factor receptor bound protein 10 (Grb10)

L'expression tissu-dépendante peut s'illustrer par l'exemple du gène *growth factor receptor bound protein 10* (*Grb10*). En effet, chez la souris, ce gène présente une expression paternelle dans le cerveau tandis que l'expression est maternelle dans le placenta et de nombreux tissus embryonnaires. Chez l'Homme, l'expression du gène (*GRB10*) se fait également à partir de l'allèle paternel dans le cerveau, tandis que l'expression allèle-spécifique maternelle est conservée uniquement dans une zone spécifique du placenta : le trophoblaste villositaire placentaire, puisque dans les autres tissus placentaires testés présentaient une expression bi-allélique (Monk et al. 2009).

Chez le porc, *GRB10* présente une expression bi-allélique dans les tissus fœtaux (cerveau, poumons, muscle) (Congras et al. 2014). De la même façon, *IGF2* présente une expression tissu-dépendante et temps-dépendante (variable en fonction du stade de développement) permise par l'existence de plusieurs promoteurs chez l'Homme et chez la souris (Stringer et al. 2012). Cette expression tissu-spécifique d'*IGF2* permet de comprendre la pléiotropie de ce gène. En effet, *IGF2* est un facteur de croissance agissant sur des tissus très variés (Figure 29).

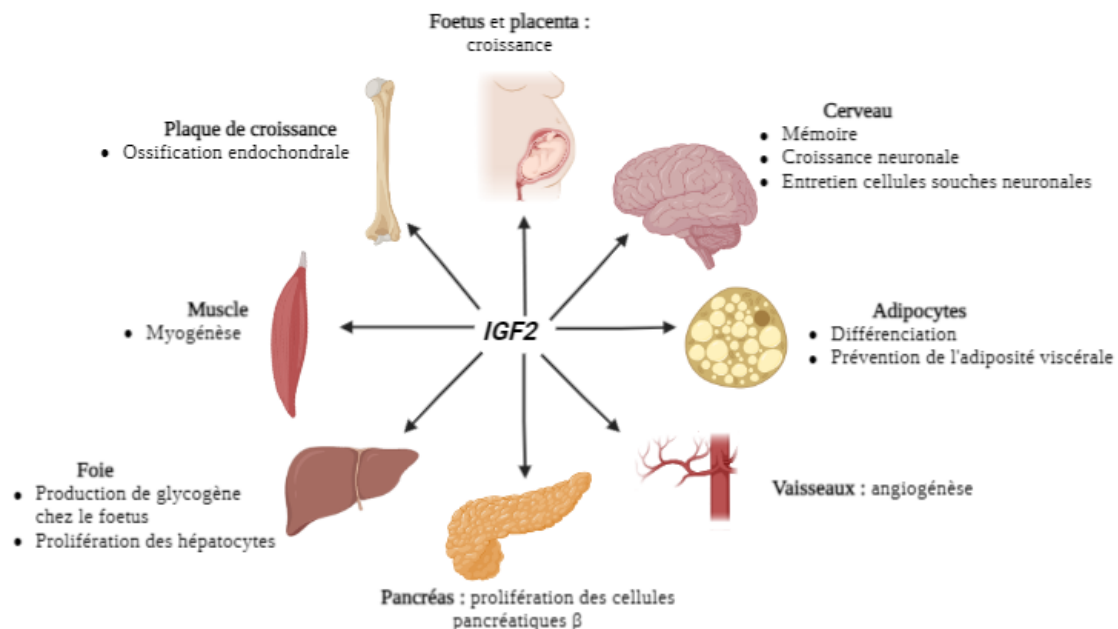


Figure 29 : Schéma résumant les rôles physiologiques de la protéine IGF2 décrits chez l'Homme et modèles de souris.

Inspiré de (Sélénou et al. 2022).

L'expression tissu-spécifique de certains gènes soumis à empreinte parentale, notamment dans le placenta ou encore dans le cerveau, peut être mise en lien avec les rôles joués par ces derniers dans le développement et la croissance.

3.2. Une information transgénérationnelle jouant un rôle dans les fonctions de développement et croissance

3.2.1. Rôle des gènes soumis à empreinte parentale

3.2.1.1. Rôles dans le développement embryonnaire

Certains gènes soumis à empreinte qui s'expriment dans le placenta participent au contrôle du développement placentaire, il s'agit par exemple du gène *Peg10* dont l'expression n'est pas spécifique du placenta, mais qui est soumis à empreinte avec expression paternelle et qui impacte le développement du trophoblaste chez l'Homme (Ono et al. 2006; Chen et al. 2015). Une autre étude s'intéressant à l'espèce porcine a mis en évidence que cette empreinte avec expression de l'allèle paternel était conservée chez le porc (Ahn, Hwang, et al. 2020a).

Outre le rôle des gènes soumis à empreinte dans le développement du placenta, certains influencent l'allocation des ressources maternelles lors de la gestation (Cassidy, Charalambous 2018). Quelques exemples seront décrits ci-dessous : l'exemple des gènes *Grb10* et *Igf2* qui interviennent sur la surface d'échange du placenta et l'exemple du gène *aquaporine (Aqp)* qui influence la densité vasculaire placentaire.

Chez la souris, le labyrinthe constitue une portion du placenta dans laquelle les échanges foeto-maternels sont intenses. Or, dans cette espèce, certains gènes soumis à empreinte exprimés dans le placenta sont capables de moduler la taille du labyrinthe qui constitue une surface d'échanges, notamment *Igf2* (Sibley et al. 2004) et *Grb10* (Charalambous et al. 2010). *Grb10* code pour une protéine adaptatrice de signalisation, qui est exprimée principalement à partir de l'allèle maternel et qui agit pour limiter la croissance fœtale et placentaire. En effet, la délétion de l'allèle maternel chez la souris entraîne une augmentation du volume du labyrinthe du placenta, ce qui n'est pas le cas lorsque l'allèle paternel seul est délété. Ainsi, *Grb10* est le premier exemple d'un gène qui agit pour limiter la taille et l'efficacité du placenta (Charalambous et al. 2010). Concernant *Igf2*, le transfert placentaire *in vivo* de trois solutés (14C-mannitol, 51CrEDTA et 14C-inuline) a été comparé entre des souris chez lesquelles l'*Igf2* spécifique du placenta avait été supprimée (P0), et des souris témoins. Les résultats mettent en évidence une capacité de diffusion théorique moyenne diminuée de 40% chez les individus knock-out P0 par rapport aux témoins sauvages et une augmentation de l'épaisseur du labyrinthe placentaire et une diminution de sa surface chez les souris P0 par rapport aux individus témoins. Ces données montrent que l'*Igf2* placentaire participe à la régulation des échanges placentaires chez la souris (Sibley et al. 2004).

L'allocation des ressources peut également être modulée via la densité vasculaire au sein du placenta. Or, certains gènes soumis à empreinte participent à ce contrôle de densité vasculaire, comme le gène codant pour l'aquaporine 1 (*Aqp1*), qui est maternellement exprimé dans le placenta. Ce gène pourrait jouer un rôle inhibiteur spécifique dans le développement vasculaire placentaire de la souris (Guo et al. 2016). Les gènes soumis à empreinte peuvent également avoir un impact sur les fonctions endocrines du placenta (John 2017). Par exemple, des souris pour lesquelles ICR1 du locus *Igf2-H19* avait été délété spécifiquement dans la jonction endocrine du placenta présentaient une augmentation de l'expression d'*Igf2* et une diminution d'expression de *H19* spécifiquement dans la cette jonction endocrine, ainsi qu'une prolifération des cellules endocriniennes suggérant le contrôle de la capacité endocrinienne placentaire par le centre d'empreinte (Aykroyd, Tunster, Sferruzzi-Perri 2022).

Si les gènes soumis à empreinte présentent un lien fort avec le placenta, ils présentent également un lien fort avec deux structures anatomiques ayant co-évolué avec le placenta : l'hypophyse et l'hypothalamus. En effet, de nombreux gènes soumis à empreinte présentent une expression parent-dépendante à la fois dans le placenta et dans l'hypophyse et l'hypothalamus. Ce lien entre l'hypothalamus, l'hypophyse et le placenta concernant l'empreinte parentale de certains gènes peut être mis en parallèle avec le lien physiologique entre ces structures endocrines impliquées dans la fonction de reproduction. De façon plus générale, de nombreux gènes soumis à empreinte sont impliqués dans le développement et la physiologie du système nerveux central dans son ensemble (Keverne 2013).

3.2.1.2. Rôle sur le développement et la physiologie du système nerveux central

Plusieurs gènes soumis à empreinte sont impliqués dans le neurodéveloppement et agissent sur le développement du système nerveux du fœtus lors de la gestation, mais également sur le développement post-natal ainsi qu'à l'âge adulte (Wilkinson, Davies, Isles 2007; Perez, Rubinstein, Dulac 2016). Au sein du système nerveux, l'expression d'origine parentale de certains gènes est spécifique d'une ou quelques région(s) anatomique(s). Par exemple, chez la souris l'expression maternelle du gène *Ube3a* se limite à des neurones du bulbe olfactif, de l'hippocampe et du cervelet, tandis que dans tous les autres tissus l'expression d'*Ube3a* est bi-allélique. La perte d'expression de l'allèle maternel de ce gène conduit notamment à une des maladies d'empreinte : le syndrome Angelman (Wilkinson, Davies, Isles 2007). En effet, le lien entre système nerveux et empreinte parentale peut également être illustré par le phénotype des maladies d'empreinte qui traduisent des anomalies du neurodéveloppement (cf 3.2.4).

L’empreinte intervient dans la régulation neuronale de comportements tels que l’alimentation et le sommeil, et influence les fonctions de l’hypophyse (Higgs et al. 2022), comme l’illustre l’exemple du gène « *zinc finger DBF-type containing 2* » (*Zbdf2*) paternellement exprimé dans le cerveau et impliqué dans les voies de stimulation de l’appétit (Glaser et al. 2022).

3.2.1.3. Rôle sur le développement de la glande mammaire

Certains gènes soumis à empreinte jouent un rôle dans le développement de la glande mammaire, ce qui est en continuité avec le rôle des gènes soumis à empreinte dans l’allocation des ressources maternelles.

Les gènes *H19* et *Igf2* sont exprimés dans la glande mammaire et sont impliqués dans la maturation de cette glande. Chez l’Homme et la souris, *H19* est exprimé à partir de l’allèle maternel tandis que *Igf2* est exprimé à partir de l’allèle paternel. *H19* est transcrit dans la glande mammaire durant la puberté et la gestation et pourrait être impliqué dans la morphogénèse et/ou la migration des cellules épithéliales (Adriaenssens et al. 2002). *Igf2* est un médiateur de la formation des cellules alvéolaires induite par la prolactine (Briskin et al. 2002). Plus récemment le rôle dans la glande mammaire du gène *mesoderm specific transcript* (*Peg1/Mest*) a lui aussi été étudié. Ce gène présente une expression paternelle dans le tissu mammaire. L’expression de ce gène décroît durant la lactation. L’application d’un stimulus induisant la différenciation lactogène des cellules témoins, n’a pas induit cette différenciation chez les cellules épithéliales mammaires de souris présentant une surexpression de *Peg1/Mest*. Ce résultat suggère que le gène *Peg1/Mest* est capable de supprimer la différenciation des cellules épithéliales mammaires (Yonekura et al. 2019).

Durant la conférence « [Genomic Imprinting from Biology to Disease](#) » organisée à Hinxton (Angleterre) en 2023, l’équipe d’Anne Ferguson-Smith a abordé le rôle de l’empreinte dans les caractéristiques nutritives du lait. Il a été montré (i) qu’au moins 7 produits protéiques des gènes soumis à empreinte étaient présents dans les échantillons de lait humain et (ii) que la protéine ZFP57 était impliquée dans la morphogénèse de la glande mammaire et la mise en place de la lactation chez la souris (données et résultats non publiés). Ainsi, des recherches vont pouvoir être menées concernant ces protéines du lait issues d’une expression à empreinte parentale, dans l’espoir de mieux appréhender la signification physiologique de l’empreinte, le rôle de l’empreinte parentale dans la santé et, de mieux comprendre la relation entre la mère et sa progéniture ainsi que l’évolution de l’empreinte génomique.

3.2.2. Connaissances comparées entre l'Homme, la souris et le porc

Les espèces chez lesquelles les connaissances concernant l’empreinte parentale sont les plus riches sont la souris et l’Homme. En effet, en raison de l’utilisation fréquente des modèles murins en recherche, les données de séquençage sont riches dans cette espèce. Concernant l’espèce humaine, les recherches concernant l’empreinte parentale sont motivées par l’impact de ce mode particulier d’expression sur la santé humaine.

D’après la banque de données Geneimprint (Geneimprint 2023) (Figure 30), 129 gènes sont connus pour être soumis à empreinte dans au moins un tissu chez l’Homme. L’allèle exprimé est mentionné pour 122 de ces 129 gènes. Parmi ces 122 gènes, 4 présentent une empreinte dépendante de l’isoforme, 39 sont à expression mono-allélique maternelle et 79 sont à expression mono-allélique paternelle. La souris, espèce modèle pour la recherche, est quant à elle l’espèce chez laquelle le plus grand nombre de gènes soumis à empreinte a été mis en évidence, puisque 150 gènes sont décrits d’après la banque Geneimprint. Parmi ces 150 gènes, les données d’expression sont mentionnées pour 143 gènes : 5 sont rapportés comme isoforme-dépendants, 76 sont à exprimés à partir de la copie maternelle et 62 sont exprimés à partir de la copie paternelle.

Cependant, une réelle dichotomie doit être faite concernant les connaissances chez ces deux espèces modèles et les connaissances chez les espèces d’élevage. En effet, chez le porc par exemple, sur cette même base de données, seuls 43 gènes sont rapportés comme soumis à empreinte (2 indiqués comme tissu-dépendants, 1 indiqué comme isoforme-dépendant, 15 à exprimés maternellement et 25 exprimés paternellement). Or, comme mentionné précédemment, il semble que l’empreinte parentale soit un phénomène plutôt conservé chez les mammifères (avec des nuances de type d’empreinte, ou encore de tissus). Il est donc cohérent de supposer que de nombreux gènes soumis à empreinte restent encore à découvrir chez le porc. De plus, ces gènes sont fortement impliqués dans le développement pré- et post-natal, c’est pourquoi ces gènes pourraient représenter un intérêt pour la filière porcine puisque les porcelets représentent l’entièreté des produits de la filière. Par ailleurs, le deuxième intérêt que pourrait représenter l’approfondissement des recherches portant sur les gènes soumis à empreinte chez le porc est l’utilisation du porc comme un modèle animal pour des maladies ou syndromes humains. Enfin, les génomes de référence ayant été établis chez la majorité des espèces d’élevage, il est aujourd’hui possible de développer les recherches concernant l’empreinte parentale chez ces espèces en caractérisant l’empreinte de gènes connus comme soumis à empreinte chez l’Homme et/ou la souris.

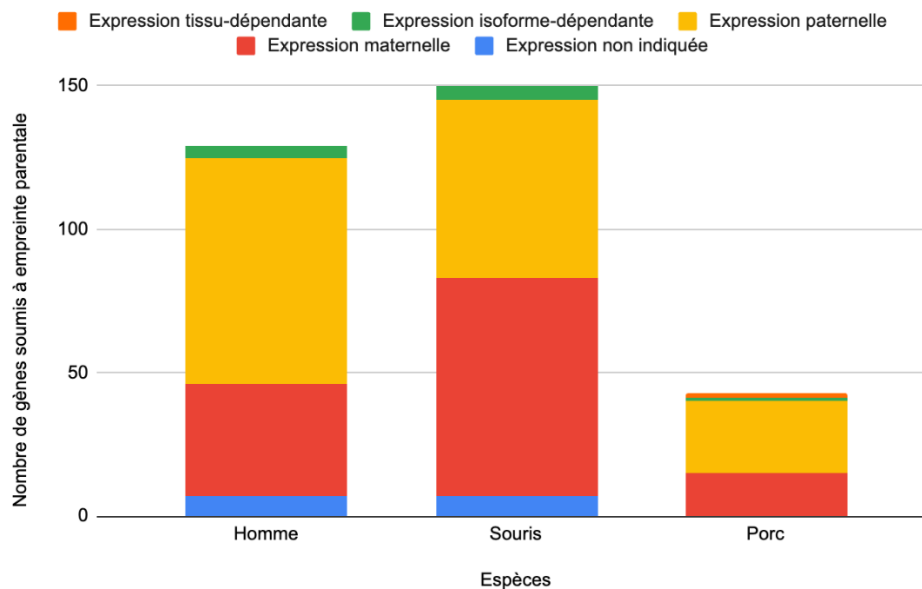


Figure 30 : Répartition des gènes soumis à empreinte connus chez l'Homme (a), la souris (b) et le porc (c).
Données (Geneimprint 2023).

Ce mécanisme d'expression génique particulière qu'est l'empreinte parentale n'est malheureusement pas infaillible ; en effet, des anomalies dans l'apposition des marques d'empreinte peuvent conduire à des modifications phénotypiques se caractérisant par des présentations cliniques variées.

3.2.3. Lien entre empreinte parentale et santé chez l'Homme

L'empreinte conduit à une expression particulière des gènes, puisqu'une seule copie est exprimée. Un désordre de ces marques d'empreinte peut donc modifier l'expression génique et donc avoir des conséquences phénotypiques. Plusieurs maladies liées à une anomalie de l'empreinte parentale sont aujourd'hui connues. Comme évoqué précédemment, les gènes soumis à empreinte sont très impliqués dans le développement embryonnaire, la croissance pré- et post-natale et la maturation du système nerveux. Les maladies affectant ces gènes particuliers se traduisent généralement par des anomalies du développement (du placenta et/ou du fœtus) entraînant une altération de la croissance et de l'acquisition des nutriments lors de la gestation et/ou des troubles du comportement. Quelques exemples sont détaillés ci-dessous.

3.2.3.1. Syndrome d'Angelman (SA)

Le syndrome d'Angelman (SA) est une maladie neuro-génétique dont la prévalence varie entre 1 enfant sur 12 000 et 1 enfant sur 20 000 (Association française du syndrome d'Angelman 2021). Ce syndrome se caractérise par des dysfonctions motrices (ataxie, tremblements, saccades), une déficience intellectuelle sévère, des troubles de la parole, des convulsions, une hyperactivité et trouble du spectre autistique (TSA) (Margolis et al. 2015). Quatre mécanismes moléculaires peuvent être à l'origine du syndrome d'Angelman : une disomie uniparentale paternelle (3–5 %), une mégadélétion maternelle de novo de la région chromosomique 15q11-q13 (70–80 %), des mutations intragéniques dans le gène *UBE3A* (10–20 %), et des défauts d'empreinte parentale qui affectent l'expression maternelle d'*UBE3A* (3 à 5 %) (Maranga et al. 2020). Ce syndrome est lié à la perte de fonction du gène *UBE3A* hérité de la mère sur la région chromosomique 15q11-q13 (Figure 31). Or cette région chromosomique est soumise à empreinte parentale, avec expression de l'allèle maternel dans les neurones. Un ARN antisens (*Snhg14*) transcrit à partir de l'allèle paternel va interagir au niveau du centre de contrôle d'empreinte de ce cluster non méthylé, et fait taire l'expression de la copie paternelle du gène *UBE3A* (Maranga et al. 2020). Chez les patients dont la cause du SA est un défaut d'établissement de l'empreinte, le promoteur maternel n'est pas méthylé correctement ce qui provoque l'expression du transcrit antisens de *UBE3A* sur l'allèle maternel. L'expression du transcrit antisens induit une inhibition de l'expression du gène *UBE3A* maternel et donc de la protéine UBE3A. Comme la copie paternelle est éteinte, le gène n'est plus exprimé et la protéine UBE3A est absente (Association française du syndrome d'Angelman 2021).

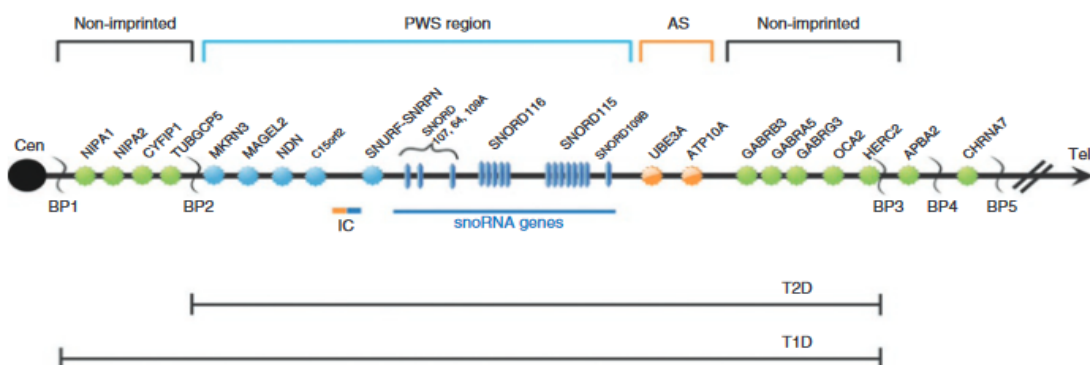


Figure 31 : Schéma de la région 15q11-2-q13 impliquée dans les syndromes d'Angelman et Prader-Willi.

La région du syndrome de Prader-Willi (PWS) (représentée en bleu) possède cinq gènes à expression paternelle (région PWS) qui codent pour des polypeptides (*MKRN3*, *MAGEL2*, *NECDIN* et *SNURF-SNRPN*) et une famille de six gènes exprimés paternellement qui codent pour des snoARN. Seuls *UBE3A* et *ATP10A* (représentés en orange), liés au syndrome d'Angelman (SA), ont une expression maternelle préférentielle, limitée à certaines régions spécifiques de tissus (en particulier le cerveau). Le centre de contrôle d'empreinte (IC) se trouve à proximité du gène bi-cistronique *SNURF-SNRPN* et dans la région à empreinte de 2,5 Mb. Le groupe de gènes du récepteur GaBa (*GABRB3*, *GABRA5* et *GABRG3*), *OCA2* (albinisme de type ii) et *HERC2* ne sont pas soumis à empreinte et ont une expression biparentale (en vert). (Cassidy et al. 2012).

Cette région chromosomique, lorsque l’empreinte parentale y est modifiée, peut être à l’origine d’un autre syndrome associé à des troubles du développement : le syndrome de Prader-Willi.

3.2.3.2. Syndrome Prader-Willi (PWS)

Ce syndrome possède des points communs avec le syndrome d’Angelman puisqu’il s’agit également d’un syndrome d’anomalie du neurodéveloppement et résulte d’erreurs d’expression génique impliquant également les gènes de la région chromosomique 15q11-q13. Dans ce cas, il ne s’agit pas d’un défaut d’expression maternelle mais d’un défaut d’expression paternelle des gènes de ce cluster. Il s’agit de la cause génétique (ou épigénétique) la plus connue de l’obésité potentiellement mortelle chez l’Homme, et la prévalence est d’environ 1 individu sur 15 000 (Butler, Miller, Forster 2019).

Les signes cliniques caractéristiques de ce syndrome sont des anomalies crânio-faciales (hypoplasie de l’émail, sécheresse buccale), une déficience intellectuelle, un déficit en hormone de croissance, un hypogonadisme et/ou hypogénitalisme, des anomalies du comportement, une hyperphagie menant à l’obésité dans la petite enfance. En fonction du mécanisme moléculaire à l’origine du syndrome, des troubles autistiques peuvent survenir (par exemple chez les individus atteints de ce syndrome en raison d’une disomie maternelle 15) (Butler, Miller, Forster 2019; Schwartz et al. 2021). En effet, comme pour le syndrome d’Angelman, plusieurs mécanismes moléculaires peuvent être à l’origine du syndrome PWS : une délétion paternelle ou une disomie maternelle du chromosome 15 (Butler, Miller, Forster 2019).

Au centre de la région impliquée dans le phénotype PWS, se trouve le gène *SNURF-SNRPN*. Il s’agit d’un gène bi-cistronique codant pour deux protéines différentes : la protéine *SmN*, qui est une protéine spliceosomale impliquée dans l’épissage de l’ARNm et un polypeptide de fonction inconnue. A l’extrémité 5' du *SNURF-SNRPN* se trouve un îlot CpG englobant le promoteur, l’exon 1, et l’intron 1 de ce gène. Il s’agit d’une DMR qui est non méthylée sur l’allèle exprimé paternellement hérité et méthylée sur l’allèle réprimé hérité de la mère. Le gène *SNURF-SNRPN* sert également d’hôte pour les six snoARN qui sont régulés par l’expression de *SNURF-SNRPN* (Cassidy et al. 2012).

Le phénotype PWS est principalement lié au défaut d’expression du locus *SNORD116* exprimé paternellement, cependant les autres gènes de la région PWS (Figure 31) (gènes *necdin (NDN)*, *MAGE family member L2 (MALGEL2)* notamment) peuvent aussi contribuer à ce phénotype (Mendiola, LaSalle 2021). Le locus *SNORD116* code pour un long ARN non codant qui

commence au niveau de la région de contrôle d’empreinte de la région PWS (IC), suivi de deux snoARN (*SNORD116* et *SNORD115*). Chez l’Homme, seul *SNORD115* (mais pas *SNORD116*) est exclusivement exprimé dans les neurones, tandis que *Snord116*, *Snord115* sont tous deux des transcrits spécifiques aux neurones chez la souris. *SNORD115* et *SNORD116* englobent des séquences codant pour des C/D snoARN. *SNORD116* et *SNORD115* sont classés comme « snoARN orphelins » parce que leurs cibles et leurs fonctions sont inconnus (Bratkovič, Božič, Rogelj 2020; Mendiola, LaSalle 2021). Les mécanismes précis de l’implication de *SNORD116* dans le neurodéveloppement ne sont pas totalement élucidés, cependant il est probable que ces zones d’ombres puissent être étudiées grâce au progrès technologique de séquençage (Mendiola, LaSalle 2021).

L’empreinte parentale est un mécanisme d’expression génique particulier qui concerne uniquement certains gènes particuliers (plus d’une centaine connue chez l’Homme et la souris). Elle est à l’origine d’une information transgénérationnelle jouant un rôle dans les fonctions de développement et de croissance. Ce mode d’expression particulier semble avoir un lien tout particulier avec le placenta puisqu’il n’a été découvert que chez les organismes euthériens et marsupiaux parmi les vertébrés. Ce lien avec la placentation, et le rôle des gènes soumis à empreinte parentale dans les fonctions de reproduction, sont à l’origine de nombreuses théories sur le succès évolutif de ce mode d’expression génique atypique qu’est l’empreinte parentale.

3.3. Les théories sur le succès évolutif de l’empreinte parentale

Une dizaine de théories ont tenté de formaliser les avantages associés à l’évolution de l’empreinte génomique chez les mammifères, les plantes à fleurs et certains insectes. Aucune ne permet d’explicitier l’intérêt de l’ensemble des gènes soumis à empreinte puisque, comme abordé précédemment, ces derniers interviennent dans un grand nombre de fonctions physiologiques. Dans leur revue, Patten et ses collègues mettent en avant trois théories principales que sont : la théorie du conflit parental, la théorie de l’antagonisme sexuel, et la théorie de la coadaptation maternelle. Ces trois théories ont pour point commun une asymétrie entre les allèles hérités du père et de la mère à un locus conduisant à une sélection favorisant l’expression différentielle des allèles (Patten et al. 2014). Deux autres théories historiques seront exposées dans les paragraphes suivants. Cependant pour une vision plus exhaustive de l’ensemble des théories concernant l’évolution de l’empreinte parentale (hors théorie du conflit parental), il est possible de se référer à la revue : (Spencer, Clark 2014).

3.3.1. Théorie de la prévention de la parthénogénèse

Comme évoqué précédemment, l’empreinte parentale a été découverte en essayant de répondre à la question suivante : les mammifères sont-ils capables de se reproduire par parthénogénèse ? Les expériences de Mc Grath et Solter en 1984 ont permis de constater l’impossibilité de parthénogénèse chez les mammifères (McGrath, Solter 1984), mais ont également conduit à la découverte des gènes soumis à empreinte en 1991. Plusieurs expériences similaires de parthénogénèse ont échoué jusqu’en 2007 où des embryons bi-maternels ont été produits à partir d’ovocytes présentant une double délétion dans les DMR de *H19* et *Dlk1-Dio3*. Ces résultats prouvent que les gènes soumis à empreinte régulés par ces deux DMR paternellement méthylés sont la seule barrière paternelle empêchant le développement normal des fœtus de souris bi-maternelles (Kawahara et al. 2007).

Ainsi, l’empreinte pourrait avoir pour rôle de prévenir les phénomènes reproductifs de parthénogénèse qui peuvent être considérés comme une impasse évolutive au sein d’une population. En effet, la parthénogénèse ne permet pas de profiter des phénomènes de recombinaison apparaissant lors de la reproduction sexuée, qui participent à maintenir de la diversité génétique (Spencer, Clark 2014). Cependant, le maintien de réseaux de plus de 100 gènes soumis à empreinte paternelle ou maternelle, comme c’est le cas chez les mammifères, en ferait un système trop coûteux évolutivement pour éviter le phénomène de parthénogénèse. De plus, les conséquences fonctionnelles de l’empreinte ne se limitent pas au placenta et à la croissance fœtale (Baran et al. 2015), comme en témoigne l’exemple du rôle de l’empreinte parentale sur les fonctions de la glande mammaire.

3.3.2. Théorie de la « bombe à retardement ovarienne »

Cette proposition repose sur l’hypothèse selon laquelle l’empreinte serait avantageuse afin d’empêcher toute expansion trophoblastique maligne résultant d’une parthénogénèse ovocytaire. L’ovocyte est ainsi vu comme une « bombe à retardement » qui serait désamorcée via l’empreinte parentale, permettant d’éviter les conséquences tumorigènes de la parthénogénèse chez certaines espèces vivipares (Varmuza, Mann 1994). Si les premiers exemples de loci soumis à empreinte peuvent s’accorder avec cette vision (exemple d’*Igf2*), cette hypothèse n’explique pas l’empreinte de gènes non impliqués dans le développement du trophoblaste, ni pourquoi l’empreinte persiste longtemps après la fécondation, chez le fœtus et l’adulte. Par ailleurs, une expansion trophoblastique n’aurait pas la même importance chez

toutes les espèces puisqu'elle va dépendre de l'histologie du placenta. Ainsi, les espèces à placenta peu invasif sont peu confrontées à cette problématique, or ces espèces présentent tout de même des gènes soumis à empreinte parentale. L'hypothèse de la « bombe à retardement ovarienne » ne semble donc pas avoir un pouvoir explicatif généralisable à l'ensemble des phénomènes d'empreinte (Spencer, Clark 2014).

3.3.3. Théorie de l'antagonisme sexuel

Day, Bonduriansky et leurs collègues soutiennent que l'empreinte peut résulter d'une sélection sexuellement antagoniste, par laquelle différents allèles d'un même locus sont plutôt favorisés chez les mâles ou chez les femelles (Day, Bonduriansky 2004). Dans cette théorie, les patrigènes (allèles transmis par le père) seront enrichis pour les allèles mâles bénéfiques et les matrigènes (allèles transmis par la mère) seront enrichis pour les allèles femelles bénéfiques (Patten et al. 2014; Spencer, Clark 2014). Cependant cette hypothèse nécessite l'existence d'une empreinte sexe-spécifique : c'est-à-dire une empreinte dans laquelle l'expression ou la non-expression d'un gène dépend non seulement du sexe du parent qui transmet le gène, mais aussi du sexe de la progéniture réceptrice. Certains loci semblent en effet être soumis à empreinte avec une expression sexuellement dimorphique, notamment dans le cerveau (Hager et al. 2008; Gregg et al. 2010). Cependant, la majorité des gènes soumis à empreinte étudiés et évoqués précédemment ne présentent pas une expression dépendante du sexe de l'individu : que le descendant soit un mâle ou une femelle, si un gène est soumis à empreinte la copie exprimée de ce gène sera la même pour tous les descendants (paternelle pour tous ou maternelle pour tous). Par ailleurs, si cette théorie a un pouvoir explicatif potentiellement étendu, elle suggère que l'empreinte parentale ne serait pas présente chez l'ensemble des organismes gonochoriques. Or, l'empreinte parentale n'a jusqu'à présent été découverte que chez les mammifères thériens, les angiospermes et certains insectes.

3.3.4. Théorie du conflit parental (ou « la bataille des sexes »)

Selon la théorie du conflit parental, l'apposition différentielle de l'empreinte serait liée à des intérêts différents entre le père et la mère : l'intérêt de la mère (qui est la principale source nourricière de la progéniture) serait d'économiser des ressources au cours de la gestation pour conserver une part des ressources pour les portées suivantes, tandis que l'intérêt du père serait

de maximiser les ressources accessibles pour les fœtus afin de favoriser leur croissance. Ainsi, l’empreinte parentale fonctionnerait de sorte que les gènes exprimés maternellement et paternellement aient tendance à réduire et à augmenter la demande des ressources maternelles, respectivement (Moore et al. 1991; Haig 2014). En effet, comme évoqué précédemment de nombreux gènes soumis à empreinte sont impliqués dans l’allocation des ressources maternelles, et dans le développement fœtal. Cette théorie postule que les gènes impliqués dans l’acquisition des ressources (comme le facteur de croissance *Igf2*) seraient exprimés paternellement, alors que les gènes impliqués dans la conservation des ressources (tels que le gène codant pour la protéine GRB10 qui agit pour limiter la croissance placentaire et fœtale) seraient exprimés maternellement (Sibley et al. 2004; Charalambous et al. 2010).

Patten et ses collègues font le lien entre la théorie du conflit parental et celle de la valeur sélective inclusive. La théorie de la valeur sélective inclusive fait référence à la capacité d’un individu à transmettre ses gènes, par l’intermédiaire de sa progéniture et de celle de parents proches ayant des gènes communs (Haig 1997). En d’autres termes, si un animal aide un individu apparenté à améliorer son succès reproducteur, cela augmente sa valeur sélective inclusive (Patten et al. 2014). Or, la femelle « aurait intérêt » à favoriser le nombre de portées qu’elle pourra mener à bien afin de favoriser sa valeur sélective inclusive, tandis que pour le mâle le coût investi dans le développement d’une portée ne risquera pas de dégrader sa valeur sélective, il aurait donc plutôt pour intérêt à assurer le bon développement de la portée en cours.

3.3.5. Théorie de la coadaptation fœto-maternelle

La théorie de la coadaptation fœto-maternelle repose sur un organe au sein duquel l’empreinte parentale joue un rôle majeur : le placenta. En effet, tout comme la théorie du conflit parental, la théorie de la coadaptation fœto-maternelle met la fonction de reproduction au cœur du rôle des gènes soumis à empreinte. Cependant, dans ce cas, l’hypothèse faite n’est pas celle d’un conflit mais plutôt d’une relation positive pour les deux individus. Cette théorie a été proposée par Wolf et Hager en 2006 suite à la constatation du fait que les interactions entre la mère et le fœtus sont souvent sélectionnées de sorte à favoriser le développement de la progéniture (Wolf, Hager 2006). Certains allèles peuvent être incompatibles, ainsi, les gènes soumis à empreinte auraient pour rôle de faire taire l’un des deux allèles afin d’éviter de possibles interactions néfastes, dans le but de favoriser le développement de la progéniture (Wilkins 2014). Par exemple, le gène *paternally expressed gene 3 (Peg3)* améliore les capacités de reproduction par

un double rôle. Son expression chez l'adulte contrôle la production de lait chez la mère, tandis qu'il influence la croissance placentaire et post-natale et le réflexe de succion chez le petit. Une étude examine comment une mutation ciblée (de l'allèle paternel ou maternel) de ce gène régule de multiples aspects du développement fœtal et postnatal, y compris la taille placentaire, la croissance fœtale, l'allaitement et croissance postnatale, l'âge de sevrage et début de la puberté. Les mères pour lesquelles l'allèle paternel a été délété présentent un succès de reproduction moindre avec des soins maternels réduits, une réduction de l'apport alimentaire maternel pendant la gestation et une altération de l'éjection du lait, qui à son tour réduit la croissance du nourrisson et retarde le sevrage et le début de la puberté. Lorsque la mutation était chez le fœtus, les mères de type sauvage mangeaient moins et n'augmentaient pas leur prise alimentaire durant la dernière semaine de gestation, suggérant une altération des signaux endocriniens placentaires qui sont en partie responsables de la régulation de l'alimentation maternelle. Ceci est un exemple de synchronisation de traits co-adaptatifs chez la mère et la progéniture par le même gène à empreinte exprimé paternellement (Curley et al. 2004).

A ce jour aucune théorie ne permet d'expliquer en totalité le succès évolutif de l'empreinte parentale. L'exploration d'une telle question nécessite en particulier des approches génomiques comparatives des loci ciblés par l'empreinte et de leurs mécanismes de contrôle, ce qui est notamment rendu possible grâce au développement de technologies de séquençage toujours plus sophistiquées, rapides et moins coûteuses.

PARTIE 2 : Étude expérimentale des signatures moléculaires de l’empreinte parentale sur la région *H19-IGF2/KCNQ1-CDKN1C*

I. CONTEXTE

1.1. Intérêt du croisement réciproque pour l’étude de l’empreinte

L’empreinte parentale est un exemple de régulation épigénétique qui aboutit à une expression parent-dépendante d’un allèle. Cette expression mono-allélique dépendante de l’origine parentale résulte de l’apposition de marque épigénétiques (méthylation des dinucléotides CpG et marques d’histones) de façon parent-spécifique. Ainsi, pour caractériser l’expression et/ou la méthylation dépendante de l’origine parentale, il est nécessaire de distinguer l’origine parentale des deux allèles. Pour cela, on exploite le polymorphisme de l’ADN pour identifier des variants dits informatifs dans des croisements réciproques.

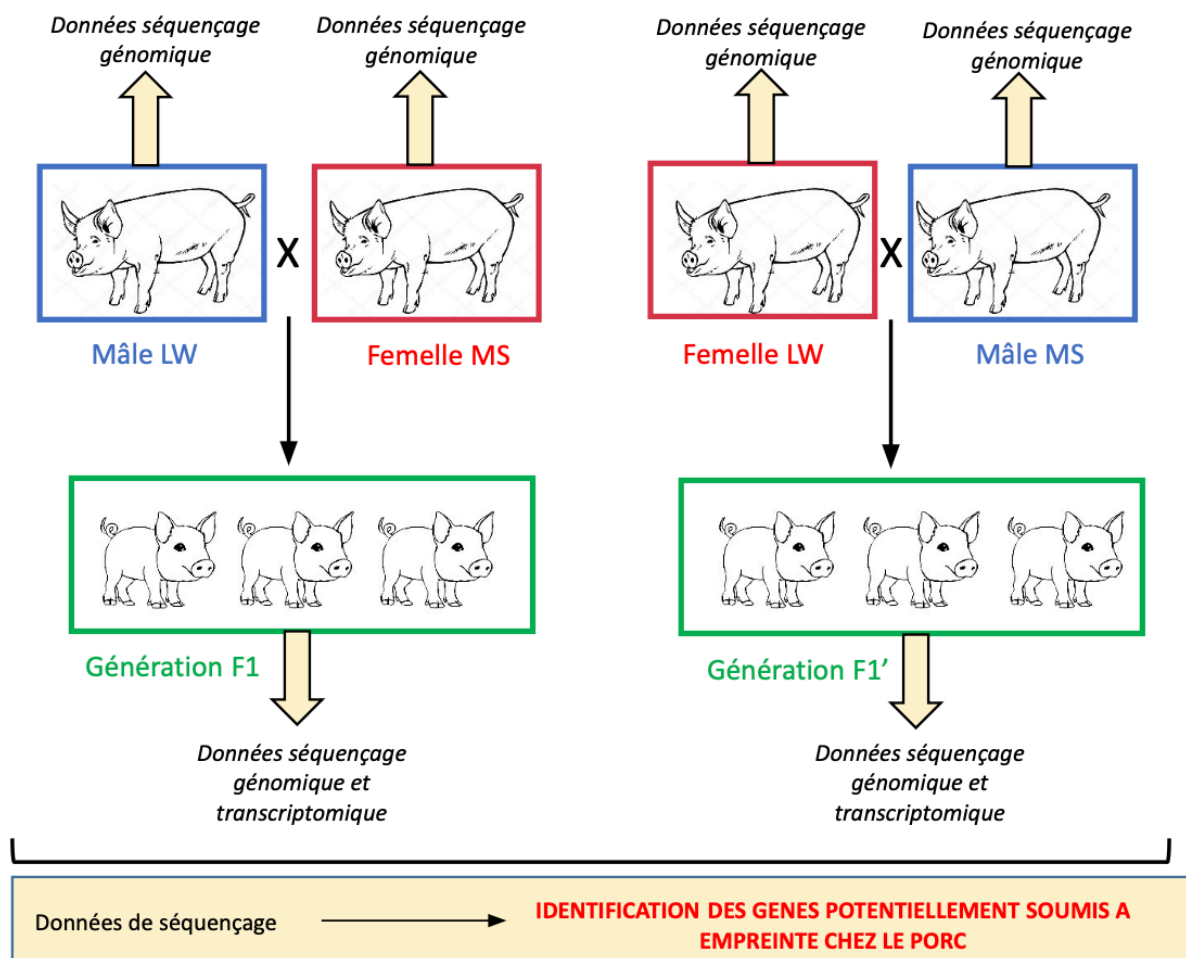


Figure 32 : Croisement réciproque dans le cadre de l’étude.

Ces croisements permettent de distinguer une expression et/ou méthylation mono-allélique classique d'une expression mono-allélique dépendante de l'origine parentale, comme cela a pu être montré chez l'abeille (Xin Wu et al. 2020). Pour maximiser le nombre de variants informatifs chez les descendants et ainsi identifier sans ambiguïté l'origine parentale, un croisement réciproque entre des races phylogénétiquement éloignées est le plus pertinent. Dans le cadre de la caractérisation des mécanismes d'empreinte parentale chez le porc, les races Large White (LW) et Meishan (MS) ont été utilisées (Ollivier et al. 2004). Un croisement réciproque a donc été réalisé entre les individus de ces deux races, soit le croisement d'une femelle MS avec un mâle LW et d'une femelle LW avec un mâle MS (Figure 32).

1.2. Données nécessaires à l'évaluation des mécanismes d'empreinte

L'étude de l'empreinte parentale peut être réalisée par des approches omiques variées (Hubert, Demars 2022). La prise en compte simultanée des données de génotypage, épigénotypage et expression doivent permettre de conclure sur l'origine parentale des allèles transcrits et/ou méthylés.

1.2.1. Les données génomiques

L'identification de variants informatifs permettant de déterminer l'origine parentale des allèles peut être faite de façon globale à partir du séquençage de tout le génome (Paragraphe 2.1.4 de la partie bibliographie) ou de façon ciblée dans une région d'intérêt. Alors que les outils bioinformatiques de traitement des données génomiques permettent l'accès à une liste de variants informatifs, les différentes approches ciblées requièrent en amont la connaissance du variant d'intérêt. Pour ces dernières, le choix de la technique de biologie moléculaire va dépendre du coût, du nombre d'individus à génotyper et de la facilité de l'expérimentation. Ici, un focus sur une approche d'amplification allèle-spécifique, nommée PCR Allelic Competitive Extension (PACE®), est présentée dans l'Encadré 5.

Encadré 5 : Principe de la PACE®.

Cette technologie repose sur la distinction des allèles selon leur fluorescence. Chacune des deux amorces ciblant un allèle est équipée d'une séquence codant pour un fluorochrome différent. Par compétition entre les deux allèles, les amorces vont s'hybrider (ou non) selon si l'allèle qu'elles ciblent est présent ou non (Semagn et al. 2014). Un signal fluorescent est ensuite émis, dont la couleur permettra, après lecture en point final, de déterminer le ou les allèles présents pour un individu au niveau de ce SNP.

1.2.2. Les données épigénomiques

L'identification des marques épigénétiques et plus particulièrement de la méthylation des dinucléotides CpG peut être faite à l'échelle du génome entier ou bien de régions ciblées comme dans le cas des DMRs. La majorité des techniques actuelles reposent soit sur un traitement chimique ou enzymatique permettant de discriminer les cytosines méthylées des dinucléotides CpG, des autres cytosines (Li 2021; Mehrmohamadi et al. 2021; Halabian et al. 2021). Les approches les plus récentes de séquençage de la molécule native d'ADN permettent de lire directement la méthylation au cours du séquençage.

L'approche la plus utilisée jusqu'à présent pour évaluer la méthylation à l'échelle du génome est le « Whole Genome Bisulfite Sequencing » (WGBS), qui consiste à séquencer tout le génome après traitement au bisulfite de sodium. Ce traitement chimique a pour effet de transformer toutes les cytosines non méthylées en Uracile (U), qui seront donc lues comme des Thymines (T) lors du séquençage (Figure 33). Néanmoins, le WGBS présente des limites qui sont la dégradation de la molécule d'ADN par le bisulfite de sodium, les ressources informatiques qu'il nécessite mais aussi son coût. Afin d'optimiser l'utilisation de ce système, il est possible de réaliser un « Targeted Bisulfite Sequencing », consistant à ne séquencer que des régions d'intérêt préalablement identifiées. Deux types de méthodes TBS existent : la méthode d'enrichissement ciblé (ou capture) qui implique l'utilisation de sondes qui vont permettre d'identifier les zones d'intérêt à séquencer, et la méthode d'enrichissement non spécifique qui consiste à séquencer des zones riches en îlots CpG.

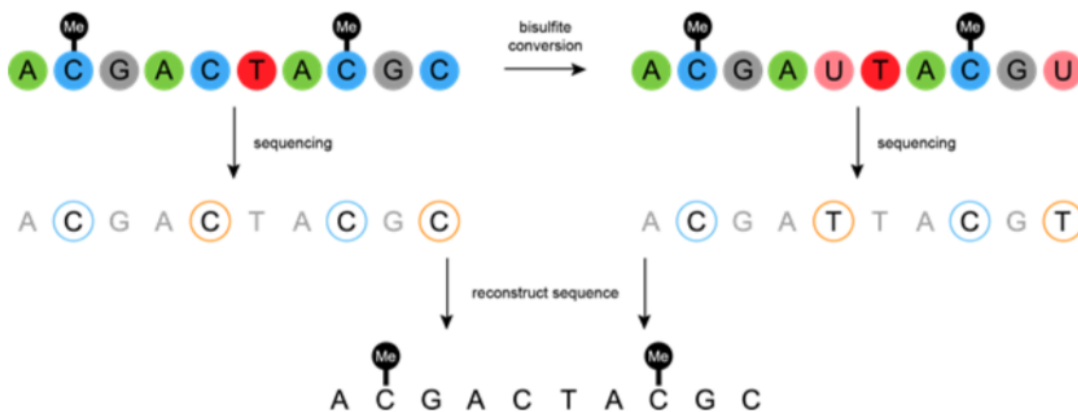


Figure 33 : Principe de la conversion au bisulfite de l'ADN.

L'objectif de toutes ces technologies est non seulement d'identifier quelles sont les cytosines méthylées et d'évaluer ainsi leur degré de méthylation qui peut varier entre 0 et 100%. Dans le cas des DMR identifiées dans les régions soumises à empreinte où en théorie un allèle parental est complètement méthylé alors que l'autre n'est pas du tout méthylé, on s'attend à une hémiméthylation (50% de méthylation). L'intérêt des technologies épigénomiques issues de NGS est l'accès simultané aux informations génétique (identification de variants) et épigénétique (identification de la méthylation des cytosines).

1.2.3. Les données transcriptomiques

L'identification de l'expression mono-allélique dépendante de l'origine parentale pour caractériser sans ambiguïté un gène soumis à empreinte requiert l'analyse des transcrits extraits de tissus d'intérêt. La très grande majorité des transcrits présents dans un tissu à un instant particulier présente une expression bi-allélique signifiant que les deux allèles parentaux sont transcrits de façon équivalente. En revanche, dans le cas des gènes soumis à empreinte parentale, un seul des deux allèles parentaux sera transcrit et une expression mono-allélique sera alors observée. Alors que le séquençage du transcriptome d'un tissu d'intérêt permet de caractériser en une seule expérience le niveau d'expression d'un transcrit et son type d'expression (bi- ou mono-allélique), l'évaluation de gènes particuliers requiert en amont la connaissance de l'origine parentale des allèles pour la mise en place des expériences et l'analyse des résultats.

Plusieurs études, basées sur le séquençage du transcriptome chez des individus issus de croisements réciproques, ont eu pour objectif d'évaluer la présence d'empreinte parentale dans une espèce (Frésard et al. 2014) ou de déterminer la carte des gènes soumis à empreinte (Yin-Qiao Wu et al. 2020). Ces approches sont aujourd'hui facilement réalisables et peu coûteuses ; elles permettraient de mieux comprendre l'origine évolutive des mécanismes d'empreinte parentale.

1.3. Choix de la région d'étude

Le choix de la région s'est porté sur un cluster dont l'empreinte parentale est bien connue chez les espèces modèles que sont l'Homme et la souris. Ce dernier se trouve sur le chromosome 2, entre les positions 1 381 838 pb et 2 059 694 pb (correspondant à la région 11p15). Au niveau de cette partie du chromosome, sont localisés notamment les gènes *IGF2*, *H19*, *CDKN1C*, et *KCNQ1*. Cette région a également été choisie en raison de l'implication du gène *IGF2* dans le caractère de croissance, caractère intéressant la filière porcine (Nezer et al. 1999; Van Laere et al. 2003).

Le cluster comportant les gènes *H19/IGF2 - KCNQ1/CDKN1C*, soumis à empreinte parentale, est impliqué chez l'Homme dans les syndromes Silver-Russell and Beckwith-Wiedemann (Sélénou et al. 2022). La protéine IGF-II est impliquée dans la croissance fœtale mais également dans le développement de néoplasies. La perte d'empreinte parentale au niveau du cluster *IGF2* peut être utilisée comme un marqueur épigénétique du risque de cancer chez l'homme (Cui 2007; Kasprzak, Adamek 2019). Chez le porc, la région du gène *IGF2* influence des caractères de composition corporelle, des caractères de qualité de la viande ainsi que le nombre de tétines (Sanchez et al. 2003). De plus, une mutation dans cette région a été mise en évidence comme influençant la musculation des porcs (Van Laere et al. 2003). L'empreinte parentale au niveau de la région *IGF2* chez le porc a été mise en évidence en 1999 à l'aide de croisements entre les races Large White et Piétrain (Nezer et al. 1999), puis a été confirmé dans un croisement entre sanglier et Large White (Jeon et al. 1999), et enfin avec un croisement entre races Landrace et Meishan (de Koning et al. 2000; Amarger et al. 2002).

L'objectif de mon stage a consisté en une caractérisation moléculaire de la région du chromosome 2 porcine comportant les gènes *H19/IGF2 - CDKN1C/KCNQ1* connus pour être soumis à empreinte chez l'Homme et la souris et dont les données chez le porc semblent largement incomplètes. Pour ce faire, à partir de données obtenues sur les individus du croisement réciproque LW x MS, j'ai identifié, sélectionné et confirmé (i) des variants informatifs permettant de discriminer l'origine parentale des deux allèles, (ii) des DMR présentant une méthylation dépendante de l'origine parentale et (iii) des transcrits présentant une expression mono-allélique parentale-spécifique.

II. MATERIELS ET METHODES

L'étude expérimentale s'est organisée en trois axes (Figure 34) : le génotypage, l'épigénotypage et l'étude de l'expression avec un focus plus important sur l'étude de la méthylation qui présente un aspect plus novateur.

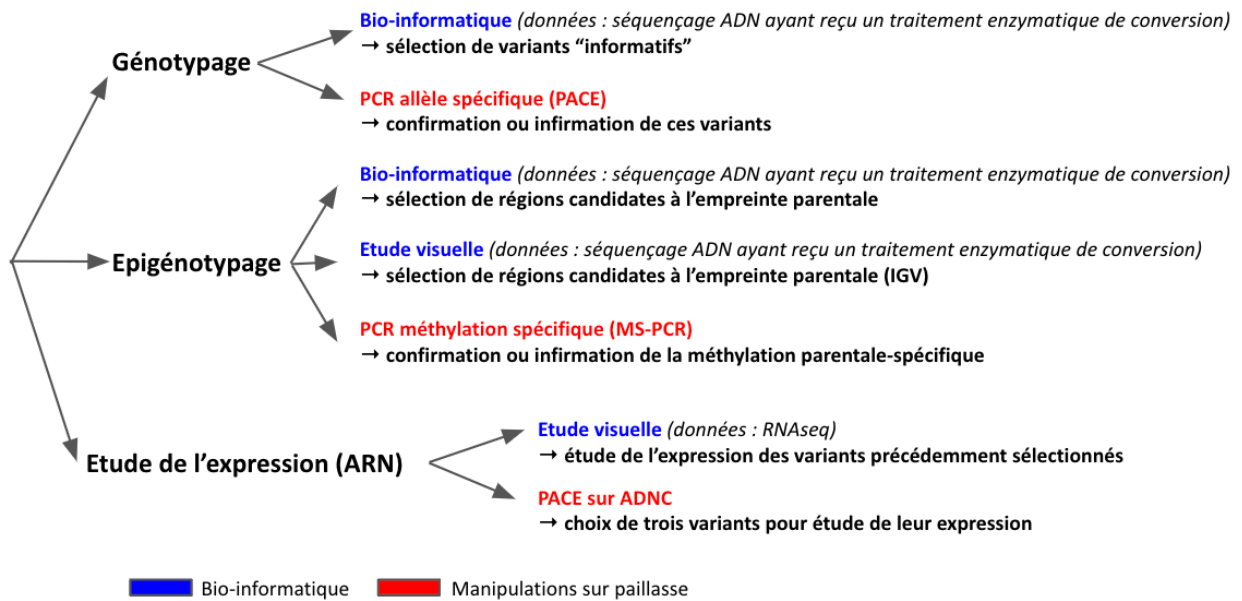


Figure 34 : Organisation de la partie expérimentale en trois axes principaux.

2.1. Les éléments disponibles avant mon arrivée

2.1.1. Les animaux et échantillons

Les échantillons disponibles proviennent de verrats, truies et porcelets issus de trois familles de croisements réciproques LW x MS (Figure 35). Ces échantillons (ADNg, ADNg traités au bisulfite de sodium et ARN) avaient été extraits de divers tissus avant mon arrivée (sang, cerveau, muscle) et sont présentés dans le Tableau 3.

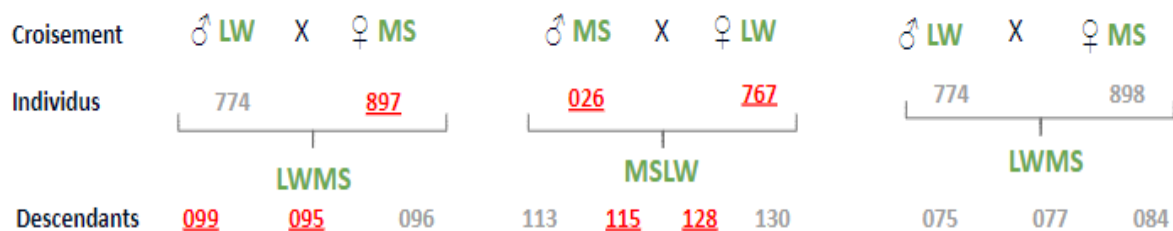


Figure 35 : Les trois familles utilisées pour cette étude. En rouge soulignés : les individus dont les données de séquençage d'ADNg traité sont disponibles.

2.1.2. Les données

Les données épigénomiques et transcriptomiques avaient déjà été générées avant mon arrivée :

- Données épigénomiques - dans le cadre de ce projet, une méthode de capture ciblée aux gènes soumis à empreinte a été développée à partir d'ADN de sang ; celle-ci repose sur une conversion enzymatique pour évaluer la méthylation (Hubert et al. 2022a; 2022b; Hubert, Perret, et al. 2023). A mon arrivée, les traitements bioinformatiques préliminaires avaient été réalisés et je disposais des fichiers d'analyse pour la visualisation des génotypes des variants et des épigénotypes de la méthylation. Seuls quelques individus du croisement réciproque ont été obtenus avec cette technologie (animaux mentionnés en rouge sur la Figure 35) ;
- Données transcriptomiques - des données de séquençage tout génome à partir des ARN extraits de cerveau et de muscle étaient également disponibles. La totalité des descendants des croisements réciproques ont été analysés avec cette approche (Tableau 1). Les premiers traitements bioinformatiques avaient été réalisés et je disposais des fichiers de visualisation permettant de quantifier l'expression des transcrits et d'évaluer une éventuelle expression mono-allélique dépendante de l'origine parentale.

Race	Individu	Père	Mère	Sexe	ADNg*	ADNg traité**	ARN
MS	026	/	/	M	X		
MS	897	/	/	F	X		
MS	898	/	/	F	X		
LW	767	/	/	F	X		
LW	608	/	/	M	X		
LWMS	075	774	898	F	X	C, M, S	C, M
LWMS	077	774	898	F	X	C, M, S	C, M
LWMS	084	774	898	F	X	C, M, S	C, M
LWMS	095	774	897	F	X	C, M, S	C, M

LWMS	096	774	897	F	X	C, M, S	C, M
LWMS	099	774	897	M	X	C, M, S	C, M
MSLW	113	026	767	F	X	C, M, S	C, M
MSLW	115	026	767	F	X	C, M, S	C, M
MSLW	128	026	767	M	X	C, M, S	C, M
MSLW	130	026	767	M	X	C, M, S	C, M

Tableau 3 : Échantillons et données disponibles.

**ADNg : ADN génomique extrait à partir de sang ; **ADNg traité : ADNg ayant reçu le traitement enzymatique (transformation des Cytosines non méthylées en Uracile) C : cerveau, M : muscle et S : sang*

2.2. Analyses bioinformatiques

Des développements bio-informatiques permettant l'analyse de données épigénomiques étaient en cours pendant mes diverses périodes de stage. J'ai alors effectué quelques analyses par moi-même afin de mieux comprendre comment ces outils fonctionnaient sur notre jeu de données. Ces analyses sont opérées au travers de l'exécution de pipelines dédiés, c'est-à-dire de séries de tâches séquentielles visant à traiter les données omiques brutes de façon automatisée et pertinente pour répondre à la question biologique posée.

2.2.1. Séquençage de l'épigénome

Les analyses ont été réalisées sur le cluster de calcul de la plateforme GenoToul Bioinfo (<https://bioinfo.genotoul.fr/>). Les données utilisées en entrée sont les données épigénomiques présentées dans le paragraphe 2.2.2. Le pipeline se divise en deux grands axes : la détermination des haplotypes et l'identification des régions enrichies en motifs d'hémi-méthylation, que l'on peut considérer comme des DMR candidates en première approximation (Hubert, Perret, et al. 2023). Les informations tirées de ces deux axes sont ensuite réunies afin de détecter les régions dont la méthylation est allèle-spécifique (incluant donc les régions à méthylation parent-spécifique) (Figure 36).

2.2.1.1. Détermination des haplotypes

Cette branche du pipeline est destinée à identifier les haplotypes portés par les individus du croisement réciproque pour lesquels des données épigénomiques sont disponibles. L'obtention des haplotypes permettra de les utiliser chez les descendants comme des marqueurs génomiques de l'origine parentale des séquences méthylées.

De telles analyses bioinformatiques commencent classiquement par une étape de mapping, qui permet d'établir une cartographie génomique des séquences brutes, suivie d'une étape de calling, qui permet d'identifier les variants les plus probables chez chaque individu. Ces étapes ont été réalisées à l'aide de l'outil gemBS, bien adapté à la découverte du polymorphisme dans les données épigénomiques (Merkel et al. 2019). La phase des variants bialléliques est ensuite déterminée à l'aide de l'outil WhatsHap, qui utilise une approche intégrative pour identifier les variants transmis de façon groupée à la descendance (Garg, Martin, Marschall 2016).

2.2.1.2. Identification des DMRs candidates

La description détaillée de cette branche du pipeline et la justification des seuils choisis est disponible dans la publication : (Hubert, Iannuccelli, et al. 2023). Brièvement, le mapping et le calling de méthylation a été réalisé à l'aide de l'outil Bismark (Krueger, Andrews 2011) avec des paramètres standards. Pour la détection des régions hémi-méthylées (DMR candidates), une fonction R personnalisée a été utilisée et deux critères de sélection ont été comparés :

- Un critère "strict", permettant de détecter des petits motifs ponctuels d'hémi méthylation : considérer qu'une région est hémi-méthylée lorsque 5 CpG hémi-méthylées sont présents dans un intervalle de 100 pb. Dans ce cas un CpG est considéré comme hémi méthylé lorsque le taux de méthylation est entre 40% et 60%.
- Un critère "lâche", permettant de détecter de plus grandes régions hémi méthylées : considérer qu'une région est hémi-méthylée lorsque 20 CpG hémi-méthylées sont présents dans un intervalle de 350 pb. Dans ce cas un CpG est considéré comme hémi méthylé lorsque le taux de méthylation est entre 35% et 65%.

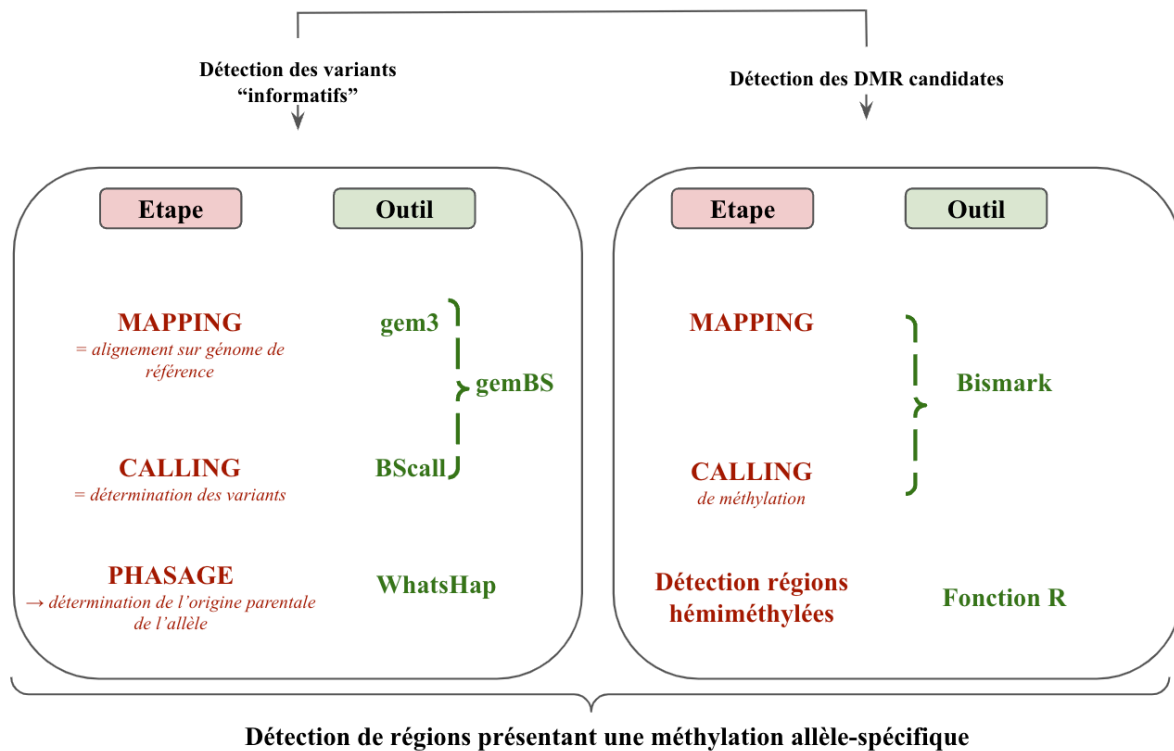


Figure 36 : Schéma simplifié du pipeline utilisé pour la détection des variants.

2.3. Protocoles de biologie moléculaire

Tous les résultats de confirmation des analyses bio-informatiques ont été basés sur de l'amplification des régions d'intérêt par PCR (Polymerase Chain Reaction) nécessitant le design d'amorces spécifiques, des conditions de PCR particulières et des analyses adéquates.

2.3.1. Design des amorces selon les expériences

2.3.1.1. Amorces de génotypage via la méthode PACE®

Le design des amorces a été réalisé par l'entreprise 3CR Biosciences, après sélection de 21 SNP. Un trio d'amorces (une amorce spécifique de chacun des allèles couplée à un fluorochrome et une amorce commune) est utilisé pour génotyper chacun des SNP. Les amorces utilisées pour la PACE® sur ADNg sont disponibles dans le Tableau 4.

Identifiant	Séquence
2_1384751_A1	GAAGGTGACCAAGTTCATGCTCCAGGAGCAGAGTGCGCAC
2_1384751_A2	GAAGGTTCGGAGTCAACGGATTATCCAGGAGCAGAGTGCGCAA
2_1384730_C1	TCCCTGCCCAGCCCCACT
2_1431183_A1	GAAGGTGACCAAGTTCATGCTGCACGGGCAGGAGGCAG
2_1431183_A2	GAAGGTTCGGAGTCAACGGATTCTGCACGGGCAGGAGGCAA
2_1431162_C1	GGCCCCGGCTCTGGGACT
2_1471484_A1	GAAGGTGACCAAGTTCATGCTCACACACACGCACGCACGCA
2_1471484_A2	GAAGGTTCGGAGTCAACGGATTACACACACGCACGCACGCG
2_1471462_C1	CGTGCGTGCGCGCGTGTGT
2_1558557_A1	GAAGGTGACCAAGTTCATGCTGAGGCCTTCCCCTGGAACC
2_1558557_A2	GAAGGTTCGGAGTCAACGGATTGAGGCCTTCCCCTGGAACG
2_1558519_C1	CCCTCCTCGTCAAGCCTCCAA
2_1598467_A1	GAAGGTGACCAAGTTCATGCTCACACCTCTGGTCGGATGTCC
2_1598467_A2	GAAGGTTCGGAGTCAACGGATTACACCTCTGGTCGGATGTCA
2_1598425_C1	GGACCAGGCCTTTCCTCCAT
2_1640689_A1	GAAGGTGACCAAGTTCATGCTAGAGATGGCGAAGTGCACAGG
2_1640689_A2	GAAGGTTCGGAGTCAACGGATTACAGAGATGGCGAAGTGCACAGT
2_1640627_C1	CCTGGTGAGGGAGTTCCCCAAA
2_1645662_A1	GAAGGTGACCAAGTTCATGCTAGCAGCCAGGCAGCCTCC
2_1645662_A2	GAAGGTTCGGAGTCAACGGATTGAGCAGCCAGGCAGCCTCT
2_1645598_C1	GCAGAGACCACTTCCGGAGCAA
2_1647007_A1	GAAGGTGACCAAGTTCATGCTAGCGGCAGGGAGCACCCA
2_1647007_A2	GAAGGTTCGGAGTCAACGGATTGCGGCAGGGAGCACCCG
2_1646969_C1	CACGTGGCTCAGGCGCTCGA
2_1859450_A1	GAAGGTGACCAAGTTCATGCTCGAGCGCCCGTCCCTGC
2_1859450_A2	GAAGGTTCGGAGTCAACGGATTGCGAGCGCCCGTCCCTGT
2_1859482_C1	CCCGAACTCCGGACGCGGT
2_2008026_A1	GAAGGTGACCAAGTTCATGCTGGAGGCCGTGGTGGGTGC
2_2008026_A2	GAAGGTTCGGAGTCAACGGATTGGAGGCCGTGGTGGGTGT
2_2007999_C1	CGCGAGCTGCTGCCTGCGT
2_2028578_A1	GAAGGTGACCAAGTTCATGCTAGTAATCACAACCCCGTCCCG
2_2028578_A2	GAAGGTTCGGAGTCAACGGATTACAGTAATCACAACCCCGTCCCA
2_2028624_C1	AGGGCCTGTGTCTGCCAGAT
2_2052041_A1	GAAGGTGACCAAGTTCATGCTCCTGACTCTGGCTCCTGGC
2_2052041_A2	GAAGGTTCGGAGTCAACGGATTGCCTGACTCTGGCTCCTGGA
2_2052063_C1	CTGGCCCTCCCCAGGGTCA
2_1453747_A1	GAAGGTGACCAAGTTCATGCTGCCGGAGCCAGGAGCTCC
2_1453747_A2	GAAGGTTCGGAGTCAACGGATTGCCGGAGCCAGGAGCTCT

2_1453768_C1	GGACCTGGGCAGCCCACCA
2_1464727_A1	GAAGGTGACCAAGTTCATGCTCCCAGGAAGGCGGTGTCTG
2_1464727_A2	GAAGGTTCGGAGTCAACGGATTGCCAGGAAGGCGGTGTCTA
2_1464702_C1	GCCAAGCTCTCCACCTGCCAT
2_1491015_A1	GAAGGTGACCAAGTTCATGCTGTGCAGGTTTCACCTGTGCTG
2_1491015_A2	GAAGGTTCGGAGTCAACGGATTTCGTGCAGGTTTCACCTGTGCTT
2_1490994_C1	GTGTGGCAGGGTGGCTGCCT
2_1629884_A1	GAAGGTGACCAAGTTCATGCTGCCACCACCCTCACCGC
2_1629884_A2	GAAGGTTCGGAGTCAACGGATTGCCACCACCCTCACCGT
2_1629858_C1	ACGAGCCCTGTAAGTGCCTCAAATT
2_1651143_A1	GAAGGTGACCAAGTTCATGCTGGCACCAACCTCGCCTCATTA
2_1651143_A2	GAAGGTTCGGAGTCAACGGATTGGCACCAACCTCGCCTCATTT
2_1651164_C1	CCTGACCGCGGCCACTGCA
2_1674918_A1	GAAGGTGACCAAGTTCATGCTGCAGGAGGGCGGGGTGG
2_1674918_A2	GAAGGTTCGGAGTCAACGGATTGGCAGGAGGGCGGGGTGA
2_1674893_C1	CAGAGGGGTCTGAGTGCCTAA
2_1730450_A1	GAAGGTGACCAAGTTCATGCTCCTGTGTCCAGAGGAAGCA
2_1730450_A2	GAAGGTTCGGAGTCAACGGATTCTCCTGTGTCCAGAGGAAGCG
2_1730483_C1	ACGACAGGCACGCACGCGTAAA
2_1886548_A1	GAAGGTGACCAAGTTCATGCTCGCGTCTTCGAGCTGCTGTAA
2_1886548_A2	GAAGGTTCGGAGTCAACGGATTTCGCGTCTTCGAGCTGCTGTAG
2_1886579_C1	GTTTAACTCCACTAAGGTTTGCGGGAA
2_1949307_A1	GAAGGTGACCAAGTTCATGCTGCGGGGAGCTGGTGAGT
2_1949307_A2	GAAGGTTCGGAGTCAACGGATTGCGGGGAGCTGGTGAGG
2_1949284_C1	CACCCTCTAGAAGCCCACCCTT

Tableau 4 : Amorces utilisées pour la PACE® sur ADNg.

L'identifiant de l'amorce correspond à : n°chromosome_position SNP_type d'amorce (A1 : amorce spécifique de l'allèle 1 ; A2 : amorce spécifique de l'allèle 2 ; C1 : amorce commune).

2.3.1.2. Amorces de génotypage via du séquençage Sanger

Le séquençage Sanger sur ADNg a été réalisé sur un SNP sélectionné pour l'épigénotypage et l'étude de l'expression mais dont les résultats de génotypage PACE sont incohérents avec le génotypage par analyse bioinformatique. Le design des amorces a été réalisé à l'aide du site Primer3Plus (<https://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>). Puis l'outil Basic Local Alignment Search Tool (BLAST) du site NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome) a permis de vérifier la spécificité des amorces. Les amorces utilisées sont disponibles dans le Tableau 5.

SNP	Identifiant	Séquence (5' - 3')	Taille amplicon (pb)
2_1859450	2_1859252_F	TAGACCTGAGCCCAGACTCG	483
	2_1859734_R	ATCGCTGCACCTGGTACG	

Tableau 5 : Amorces utilisées pour le séquençage Sanger sur ADNg pour le SNP n°9 (1 859 450 pb).

2.3.1.3. Amorces d'épigénotypage par MS-PCR

La méthode MS-PCR (Méthylation Spécifique PCR) a été utilisée pour évaluer l'origine parentale des allèles méthylés et ainsi conclure à une DMR parentale-spécifique. Cette méthode repose sur le design de deux couples d'amorces par région d'intérêt, un ciblant les CpG méthylés et l'autre les CpG non méthylés. Le design des amorces a été réalisé avec le site Urogene (<https://www.urogene.org/cgi-bin/methprimer/methprimer.cgi>). Puis l'outil Basic Local Alignment Search Tool (BLAST) du site NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome) a permis de vérifier la spécificité des amorces. Les amorces utilisées sont disponibles dans le Tableau 6.

SNP	Cible	Identifiant	Séquence (5' - 3')	Taille amplicon
2_1384751	CpG méthylés	2_1384751_left_M	AAAGAGTTTCGTTTTTTTAGGGCGG	201
		2_1384751_right_M	AAACCCTAACCCACACCCCTTACG	
	CpG non méthylés	2_1384751_left_U	AAAGAGTTTTGTTTTTTTAGGGTGG	201
		2_1384751_right_U	AAACCCTAACCCACACCCCTTACAAC	
2_1645662	CpG méthylés	2_1645662_left_M	TTCGGAGTAAGGGTTACGGTTTCGA	126
		2_1645662_right_M	GCCACCTAAAAACGACCCACGAA	
	CpG non méthylés	2_1645662_left_U	TTTGGAGTAAGGGTTATGGTTTTGA	128
		2_1645662_right_U	ACACCACCTAAAAACAACCCACAAA	
	CpG méthylés	2_1859450_left_M	ATGTGGAGTTTATTTTCGTTTAGTAGCGA	151
		2_1859450_right_M	TCCGCGACCAACTCAATACAACGTT	
	CpG non méthylés	2_1859450_left_U	ATGTGGAGTTTATTTTGTTTAGTAGTGA	151
		2_1859450_right_U	ACATCCACAACCAACTCAATACAACATT	

Tableau 6 : Amorces utilisées pour la PCR méthylation-spécifique.

2.3.2. Conditions PCR selon les expériences

2.3.2.1. PCR pour la méthode PACE®

Cette approche a été utilisée pour génotyper les ADNg mais aussi les ADN complémentaires (ADNc), issus des ARNs extraits de cerveau et de muscle pour la totalité des individus du croisement réciproque.

La PCR allèle-spécifique a été réalisée avec le kit PACE® GENOTYPING MASTER MIX (3cr bioscience) (<https://3crbio.com/wp-content/uploads/2023/01/PACE-User-Guide-v2.pdf>), avec le réactif PACE® Genotyping Master Mix (Low ROX – 25nM) : version PACE-IR ou PACE 2.0 selon les marqueurs. Dans chaque puit d'une plaque 384 puits, le volume total est de 5 µL, et sont déposés 10 ng (1µL) d'ADNg ou bien 1 µL des ADNc dilués au 1/5e ou au 1/10e. Tous les SNP sélectionnés sont ainsi génotypés pour les 14 individus du croisement réciproque avec le réactif PACE-IR et le programme PCR A (Figure 37) dans un premier temps. Les SNP pour lesquels les résultats étaient ininterprétables ou incohérents avec le génotype déterminé bioinformatiquement ont été de nouveau génotypés par PACE avec le réactif PACE 2.0 et le cycle PCR B. Les résultats ont été déterminés via une lecture en point final avec l'instrument QuantStudio™ 6 Flex System (référence : 4485691), et du logiciel QuantStudio™ Real-Time System PCR.

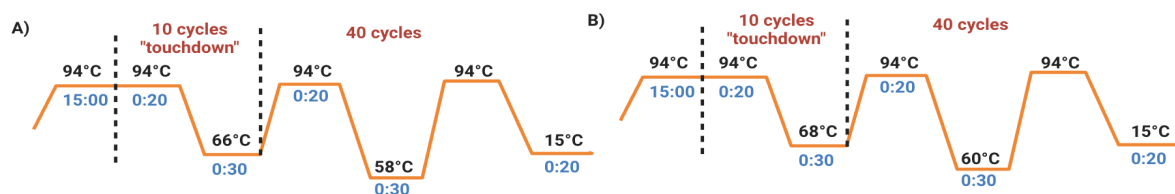


Figure 37 : Programmes PCR utilisés pour la PACE® sur ADNg.

2.3.2.2. PCR pour le séquençage Sanger

Cette approche a été utilisée à partir (i) d'ADNg pour le génotypage de variants ne fonctionnant pas avec la méthode PACE® et (ii) d'ADNg traité au bisulfite de sodium pour l'épigénotypage de variants d'intérêt.

Pour choisir la meilleure température d'hybridation des amorces (T_m) pour le cycle PCR, une PCR multi-températures a été réalisée en testant tous les couples d'amorces sur 6 températures de T_m (de 58°C à 68°C avec un pas de 2), en prenant un ADNg témoin positif et l'eau comme témoin négatif. La réalisation du mix PCR a été faite avec les concentrations finales suivantes : tampon avec $MgCl_2$ à 1 mM, dNTP à 0,2 mM, amorces à 0,5 µM, enzyme polymérase (Eurobio Scientific : PCR BIO Classic Taq, référence PB10.15-01) à 0,25 µM. Le volume final de la réaction (mélange PCR + ADNg) est de 15 µL, la quantité d'ADNg utilisée est de 10 ng par

puits (soit une concentration finale de 0,66 ng/μL). Le programme PCR utilisé pour le génotypage est présenté sur la Figure 38 et 5 cycles PCR ont été ajoutés pour l'épigénotypage. Une température de 60°C a été retenue pour toutes les réactions d'amplification. Les produits PCR sont ensuite observés après électrophorèse sur gel d'agarose (2%).

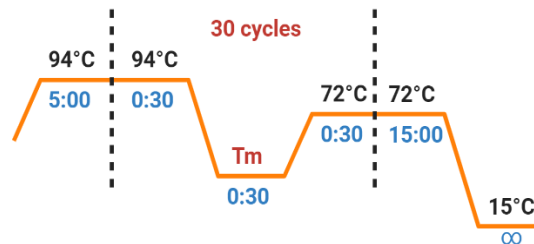


Figure 38 : Programme PCR classique utilisé.

Les produits PCR sont ensuite purifiés dans un mélange de volume final de 8 μL, avec 1 μL de produit PCR, et l'ajout des enzymes de purification : "Thermosensitive Alkaline Phosphatase" (F0654, ThermoScientific) à 0,06 U/μL final), et "Exonucléase I" (MO293L, NEB) à 1,25 U/μL final. Puis, la réaction enzymatique s'effectue à 37°C pendant 45 min et les enzymes sont inactivées 30 min à 80°C et la réaction de séquençage Sanger a été réalisée par l'entreprise Genewiz (<https://www.genewiz.com/en-GB/>) et les résultats des chromatogrammes sont analysés avec le logiciel Chromas.

2.4. Logiciels et sites utilisés

La liste des sites et logiciels utilisés est la suivante :

- QuantStudio Real-Time PCR : visualisation des résultats de PCR allèle spécifique
- <https://www.ncbi.nlm.nih.gov/gene>: BLAST pour le design d'amorces
- <https://www.ensembl.org/index.html>: visualisation de séquences, export de séquences, et BLAST pour le design d'amorces
- <https://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>: design d'amorces séquençage Sanger
- <https://www.urogene.org/cgi-bin/methprimer/methprimer.cgi>: design des amorces de PCR méthylation spécifique
- <https://technelysium.com.au/wp/chromas/>: analyse des chromatogrammes issus du séquençage Sanger

- IGV (Integrated Genome Viewer): logiciel de visualisation de séquences (Thorvaldsdóttir, Robinson, Mesirov 2013).

Le logiciel IGV a été largement utilisé dans cette étude. Il a été utilisé avec pour référence le génome Sscrofa 11.1 et le fichier d'annotation gtf 11.1.109. Il a tout d'abord été utilisé pour la détection visuelle de régions hémi méthylées (partie épigénotypage), puis pour la sélection de variants informatifs (parmi ceux détectés par la méthode bioinformatique) situés à moins de 100 pb d'une région hémi méthylée (en raison de l'utilisation d'un séquençage basé sur une longueur de lecture de 150 pb, et de la nécessité de pouvoir associer sur une même lecture les données de méthylation et de génotype dans le cadre de la question de l'empreinte parentale). Pour la détection visuelle des régions hémi méthylées, la sélection des régions a été subjective, basée sur l'aspect bicolore représentant l'hémi méthylation sur le logiciel IGV. Les positions de début et fin de toutes les régions pouvant s'apparenter à une région hémi-méthylée ont été collectées dans un tableau, puis toutes les régions séparées de moins de 500 pb ont été regroupées pour ne former qu'une région. Puis, il a été utilisé pour la visualisation des données de séquençage ARN.

III. RESULTATS

La région d'intérêt, localisée sur le chromosome 2 et d'environ 678 kb (1 381 838 bp - 2 059 694 bp), est représentée sur la Figure 39 avec les données disponibles sur le génome porcin de référence (Susscrofa 11.1) et le fichier d'annotation (gtf 11.1.109). Il est important de noter que les gènes *H19* et *IGF2* ne sont pas annotés sur cette version du génome du porc.

L'objectif de mon stage a été d'enrichir les connaissances sur cette région du génome porcin, dans le cadre de l'étude de l'empreinte parentale, via l'analyse de différentes approches -omics et leur confirmation par des approches de biologie moléculaire plus ciblées.

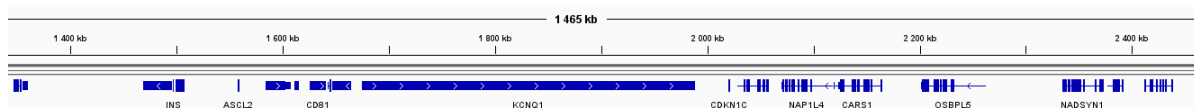


Figure 39 : Fichier d'annotation du génome porcin utilisé gtf Susscrofa 11.1.109.

3.1. Origine parentale des allèles

L'objectif a été d'identifier, sélectionner et valider des variants informatifs, permettant de discriminer l'origine parentale des deux allèles, sur la région d'étude. L'identification et la sélection ont été réalisées avec le fichier de sortie du workflow bioinformatique issu des données de séquençage épigénomiques. L'annexe 1 permet de comprendre comment exploiter les génotypes à partir des données épigénomiques. Puis, un génotypage a été développé pour (i) confirmer les génotypes déterminés par le workflow bioinformatique et (ii) déterminer le génotype des individus pour lesquels les données épigénomiques n'étaient pas disponibles.

3.1.1. Sélection bioinformatique de variants informatifs

Parmi les 678 kb de la région, 2968 variants informatifs ont été détectés par le workflow bioinformatique. Une sélection visuelle des variants situés à moins de 100 bp d'une région hémi-méthylée a été réalisée avec le logiciel IGV, permettant une sélection de 151 variants informatifs ciblant des régions hémi-méthylées. Puis, 21 ont été choisis selon les critères suivants : proche d'une grande région hémi-méthylée d'au moins 1200 bp (n=4), localisation dans un exon (n=8) et position vis-à-vis de gènes d'intérêts (n=9) (Figure 40 et Tableau 7).

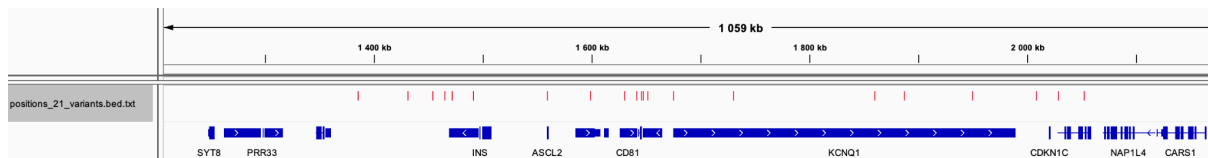


Figure 40 : Localisation des 21 variants sélectionnés.

En rose (1ère piste) : localisation sur le chromosome 2 des 21 variants sélectionnés. En bleu : gtf 11.1.109.

Nom Variant	Chromosome	Début région hémi-méthylée	Fin région hémi-méthylée	Taille région hémi-méthylée	Exon (gtf .109)	Position variant
1	2	1384526	1387094	2568	Non	1384751
2	2	1430777	1431277	500	Non	1431183
3	2	1471414	1471490	76	Oui	1471484
4	2	1558526	1558751	225	Oui	1558557
5	2	1598306	1598459	153	Oui	1598467
6	2	1640666	1640732	66	Oui	1640689
7	2	1645581	1645824	243	Oui	1645662
8	2	1646982	1647040	58	Oui	1647007
9	2	1858364	1860008	1644	Non	1859450
10	2	2006993	2008073	1080	Non	2008026

11	2	2028428	2028493	65	Oui	2028578
12	2	2051545	2051952	407	Oui	2052041
13	2	1453473	1453771	298	Non	1453747
14	2	1464632	1465948	1316	Non	1464727
15	2	1490633	1491381	748	Non	1491015
16	2	1629814	1630660	846	Non	1629884
17	2	1649915	1651186	1271	Non	1651143
18	2	1674971	1675090	119	Non	1674918
19	2	1730203	1730343	140	Non	1730450
20	2	1886266	1886628	362	Non	1886548
21	2	1949123	1949741	618	Non	1949307

Tableau 7 : Récapitulatif des 21 variants sélectionnés sur la région d'étude.

3.1.2. Validation des génotypes

Afin de vérifier que le génotypage interprété grâce aux données traitées par le workflow bioinformatique est correcte, et afin de connaître le génotype des individus non utilisés dans l'étude tout génome, tous les individus des trois familles (paragraphe 2.2.1 et Figure 35) ont été génotypés par PACE® ou par séquençage Sanger.

Parmi les 21 variants sélectionnés à partir des analyses bioinformatiques, 17 ont permis de confirmer les génotypes identifiés au préalable (Figure 41A et B), 2 étaient incohérents avec les génotypes déterminés via les données épigénomiques (variants en position 1 640 689 pb et 1 859 450 bp) et 2 n'ont pas permis une amplification correcte (Figure 41C) (variants en position 1 471 484 pb et 2 028 578 pb).

Le variant localisé 1859 450 pb présentait un intérêt particulier car il permettait de cibler une grande DMR. Le génotypage à l'aide de la méthode PACE® n'étant pas concluant, un séquençage en Sanger a été réalisé. Malheureusement, l'interprétation des chromatogrammes n'a pas permis de conclure et de génotyper sans ambiguïté les individus pour ce variant. Il semblerait qu'un problème dans le design des amorces soit la cause (résultats non montrés).

Ainsi, les génotypes des individus n'ayant pas été séquencés ont pu être déterminés pour 17 des 21 variants sélectionnés suite aux analyses bioinformatiques.

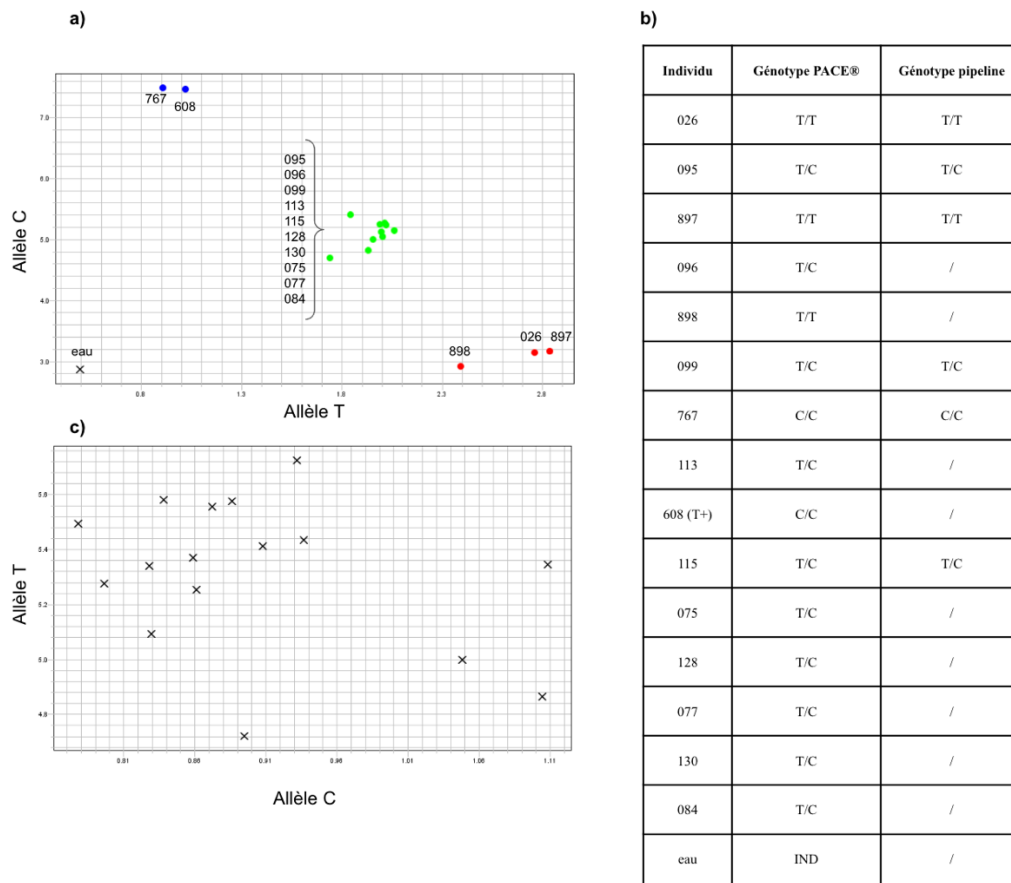


Figure 41 : Génotypage sur ADNg à l'aide de la méthode PACE®.

a) Exemple du génotypage pour le SNP 2 (position 1 431 183 pb) : en bleu : homozygotes C/C, en rouge : homozygotes T/T, en vert : hétérozygotes T/C ; b) Comparaison avec le génotypage issu du traitement bioinformatique des données de séquençage d'ADNg traité enzymatiquement (IND : indéterminé) ; c) Exemple de SNP dont les génotypes sont indéterminables par PACE® : SNP 3 (position 1 471 484 bp).

3.2. Identification de régions hémi-méthylées

L'objectif a été de (i) localiser les régions hémi-méthylées, représentant des DMR candidates, à partir des analyses bioinformatiques issues des données épigénomiques et (ii) confirmer certaines d'entre elles par des méthodes ciblées afin de déterminer si les DMR candidates présentaient une méthylation dépendante de l'origine parentale.

3.2.1. Analyse des données épigénomiques

Afin d'identifier les régions présentant une méthylation parentale-spécifique, une analyse bioinformatique et une analyse visuelle des données épigénomiques ont été comparées. Une région d'hémi-méthylation ciblant une DMR candidate est visualisée sur la Figure 42.

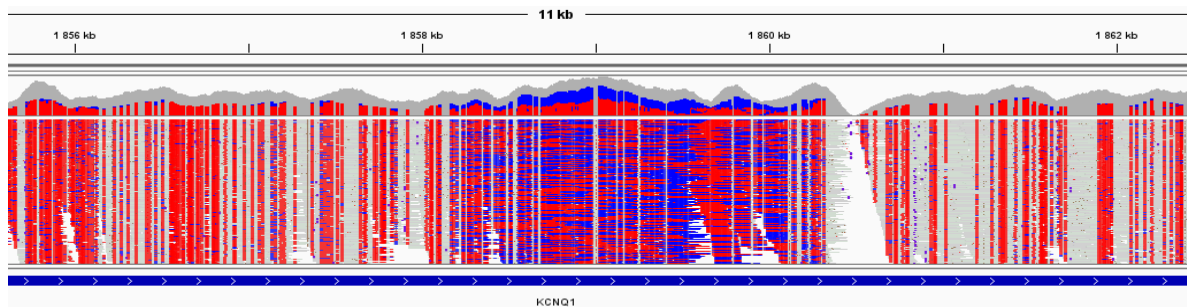


Figure 42 : Visualisation des données épigénomiques de l'individu 128 à l'aide de IGV.

Une région de 11 kb du gène *KCNQ1* a été choisie. La première piste représente la couverture de chacune des bases de la région à partir du séquençage NGS. La seconde piste représente toutes les lectures de 150 pb qui s'alignent à cet endroit du génome et qui sont donc capturées et spécifiques. Les couleurs rouge et bleu représentent respectivement les dinucléotides CpG méthylés et non méthylés. Ici au centre, une région hémiméthylée est représentée avec de part et d'autre des régions méthylées.

Pour l'identification des régions hémiméthylées de façon automatisée, deux critères de sélection ont été considérés : une région hémiméthylée correspond à l'accumulation de (i) 5 CpG hémiméthylés au taux 40%-60% dans un intervalle de 100 pb ou (ii) 20 CpG hémiméthylés au taux 35%-65% dans un intervalle de 350 pb. En regroupant les régions hémiméthylées séparées de moins de 100 pb, le critère utilisant le taux 40%-60% (respectivement 35%-65%) permet d'identifier 89 (respectivement 17) régions hémiméthylées présentant une taille moyenne de 60 (respectivement 689) pb dans la région de 677 kb via l'analyse bioinformatique.

L'analyse visuelle des données épigénomiques a permis de détecter 67 régions hémiméthylées. Parmi elles, seules 11 ne présentent aucun variant à moins de 100 pb. La taille moyenne des régions hémiméthylées détectées visuellement est de 314 paires de bases. La comparaison entre ces deux méthodes est représentée dans la Figure 43.



Figure 43 : Comparaison des détections bioinformatique et visuelle des régions hémiméthylées.

En orange (1ère piste) : régions hémiméthylées déterminées bioinformatiquement en considérant le taux de méthylation 35%-65% et un rassemblement des régions hémiméthylées localisées à moins de 100 pb. En vert (2de piste) : régions hémiméthylées déterminées bioinformatiquement en considérant le taux de méthylation 40%-60% et un rassemblement des régions hémiméthylées localisées à moins de 100 pb. En rose (3ème piste) : régions hémiméthylées déterminées visuellement. En bleu le *gtf* (version 11.1.109) du génome de l'espèce porcine (*Sscrofa*.)

3.2.2. Méthylation dépendante de l'origine parentale

3.2.2.1. Identification bioinformatique de DMR candidats

Parmi les 17 variants informatifs permettant de déterminer sans ambiguïté l'origine parentale des deux allèles : 3 semblent présenter une méthylation allèle-spécifique à minima (la vérification d'une spécificité parentale de la méthylation étant impossible pour ces variants en raison de l'homozygotie des descendants dans un sens de croisement), et 2 semblent présenter un profil de méthylation parent-spécifique chez les 7 animaux pour lesquels des données épigénomiques sont disponibles. L'exemple du variant localisé en 1 384 751 pb est montré sur la Figure 44. L'annexe 1 permet de comprendre comment exploiter les génotypes à partir des données épigénomiques.

Pour ce variant, dans la famille complète du croisement réciproque, le père est homozygote G/G, la mère est homozygote T/T et tous les descendants sont hétérozygotes G/T. Ainsi, la visualisation de cette région d'intérêt pour tous les individus permet de conclure à l'identification d'une DMR dépendante de l'origine parentale avec l'allèle paternel méthylé et l'allèle maternel non-méthylé.

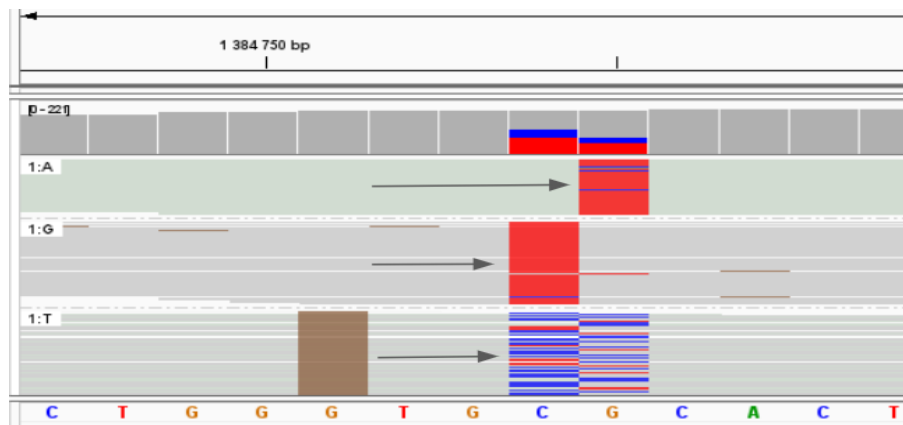


Figure 44 : Exemple du variant 1 384 751 pb discriminant les CpG méthylées et non méthylées.

Visualisation des données épigénomiques de l'individu 128 à l'aide de IGV. Les lectures de 150 pb alignées ont été triées selon les deux allèles à la position 1 384 751 pb : environ la moitié des lectures portent l'allèle T (marron) et l'autre moitié porte l'allèle G (gris ou vert). Au niveau du dinucléotide CpG positionné à proximité, les lectures incluant l'allèle T présentent majoritairement des CpG non méthylées (bleu) alors que les lectures avec l'allèle G présentent majoritairement des CpG méthylées (rouge).

3.2.2.2. Validation de DMR parentale-spécifique

Pour confirmer la méthylation dépendante de l'origine parentale identifiée au préalable, 3 régions dont la méthylation parentale-spécifique peut-être évaluée à l'aide des variants 1 384 751 pb, 1 645 662 pb et 1 859 450 pb ont été sélectionnées et analysées par MS-PCR puis

séquençage Sanger. La MS-PCR permet d'amplifier de façon indépendante l'allèle méthylé de l'allèle non-méthylé via le design de couples d'amorces spécifiques.

Pour ce faire, seuls 2 descendants hétérozygotes pour un des 3 variants ont été exploités et l'origine parentale de la méthylation dans le cerveau et le muscle a été évaluée (Figure 45). Les résultats d'amplification sont mitigés mais analysables pour certains d'entre eux. Alors que l'amplification des allèles méthylés semblent avoir fonctionnée pour les variants 1 384 751 pb et 1 859 450 pb, c'est l'amplification de l'allèle non-méthylé qui est correcte pour le variant 1 645 662 pb.

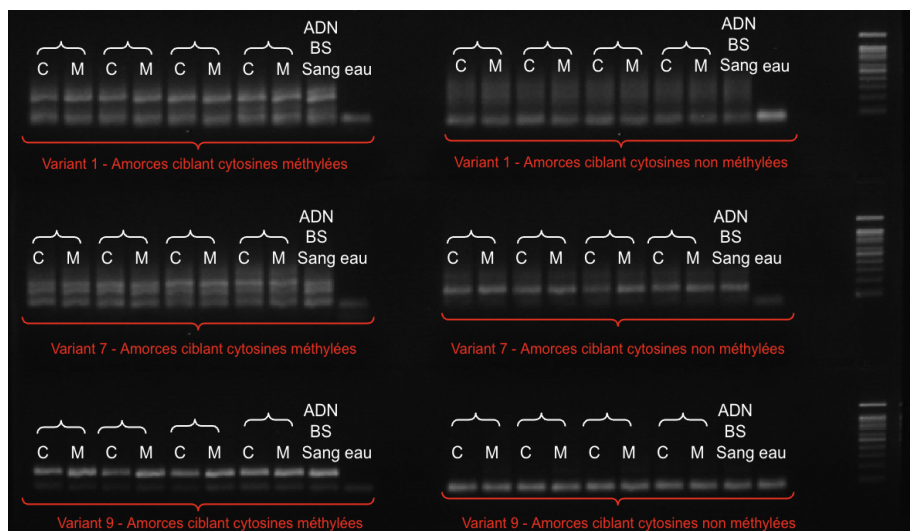


Figure 45 : Électrophorèse sur gel d'agarose des produits PCR obtenus avec les amorces de MS-PCR.

Variants 1 384 751 pb (variant 1), 1 645 662 pb (variant 7) et 1 859 450 pb (variant 9) pour les variants 1, 7 et 9. C : cerveau et M : muscle

Un séquençage Sanger a été réalisé pour les produits d'amplification correcte pour valider ou non la méthylation parentale-spécifique. L'exemple du variant localisé en 1 384 751 pb pour l'amplification de l'allèle méthylé est présenté sur la Figure 46. Celui permet de démontrer que la DMR localisée dans la région 1 384 751 pb est méthylée sur l'allèle paternel dans le cerveau et le muscle comme le suggérait les résultats issus des analyses bioinformatiques (paragraphe 3.2.2.1).

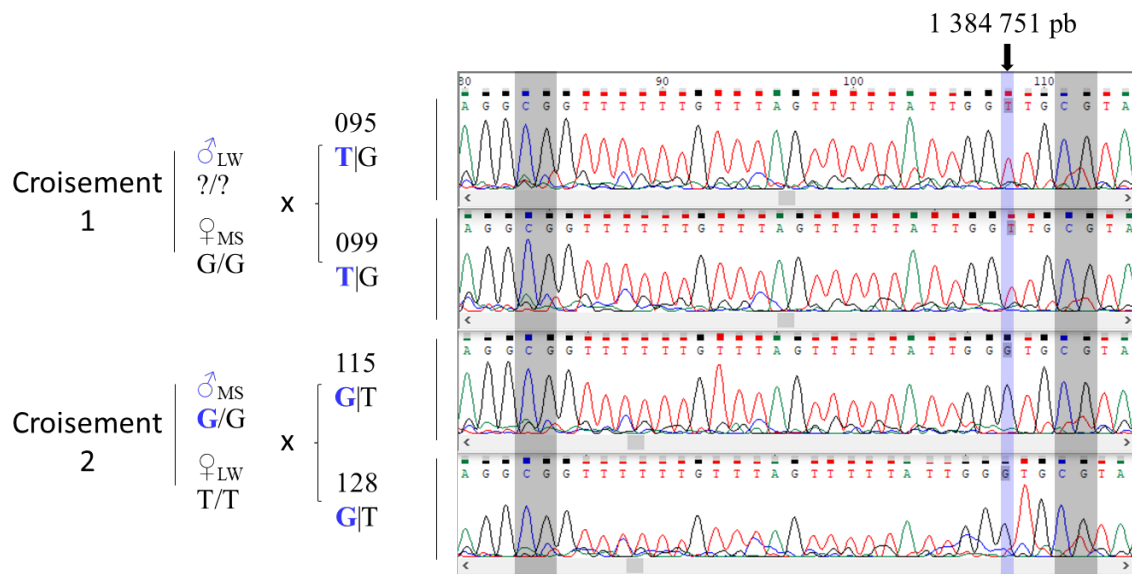


Figure 46 : Chromatogramme du fragment méthylé issu de muscle pour le variant 1 384 751 pb.

Les quatre individus séquencés des 2 croisements réciproques sont présentés sur la gauche et leurs génotypes est rappelé. La méthylation des dinucléotides CpG est encadrée en gris et la position du variant noté en bleu. Dans tous les cas, l'allèle paternel (T et G respectivement pour les croisements 1 et 2) est méthylé.

Ainsi, nous avons pu valider l'origine parentale des DMRs avoisinant les positions 1 384 751 pb et 1 859 450 pb avec une méthylation paternelle pour la région 1 384 751 pb et une méthylation maternelle pour la région 1 859 450 pb ; aucune différence n'a été observée entre le cerveau et le muscle. Aucune méthylation allèle ou parentale-spécifique n'a été observée pour la région 1 645 662 pb comme le suggérait les analyses bioinformatiques bien qu'une zone d'hémi-méthylation ait été observée.

3.3. Identification de l'expression parentale-spécifique

Indépendamment de la méthylation dépendante de l'origine parentale, l'objectif de l'étude de l'expression est d'identifier quels sont les gènes soumis à empreinte de la région d'intérêt dans le cerveau et dans le muscle. Pour ce faire, nous avons exploité les 17 variants informatifs afin d'observer s'ils permettaient de cibler des transcrits afin de potentiellement en déduire l'allèle parental exprimé.

L'étude de l'expression a été réalisée de deux manières : par inspection visuelle des données transcriptomiques par génotypage des ADNc issus des tissus cerveau et muscle pour 3 variants sélectionnés (2 étant exprimés et soumis à empreinte parentale et un étant à expression biallélique utilisé comme témoin).

3.3.1. Analyse des données transcriptomiques

Nous avons recherché une signature moléculaire de l’empreinte parentale à savoir que le génotype d’un individu à un variant informatif est hétérozygote sur l’ADNg et homozygote ou avec un biais allélique sur l’ADNc. Cette expression mono-allélique doit en plus dépendre de l’origine de l’allèle parental transcrit.

Parmi les 17 variants informatifs identifiés au préalable, 2 ciblent des transcrits exprimés de façon bi-allélique dans le muscle et le cerveau (Figure 47A), 3 ciblent des transcrits exprimés de façon bi-allélique dans le muscle et non-exprimés dans le cerveau, 1 ciblant le transcrit non-codant *HI9* semble soumis à empreinte parentale dans le muscle avec expression de l’allèle maternel et non détecté dans le cerveau (variant 1 384 751 pb) (Figure 47B), 1 ciblant le transcrit antisens long non-codant *KCNQIOT1* semble soumis à empreinte parentale dans le muscle et le cerveau avec expression de l’allèle paternel (1 859 450 pb) et 1 ciblant le transcrit *IGF2* semble soumis à empreinte parentale dans le muscle avec une expression paternelle mais bi-allélique dans le cerveau (1 471 484 pb).

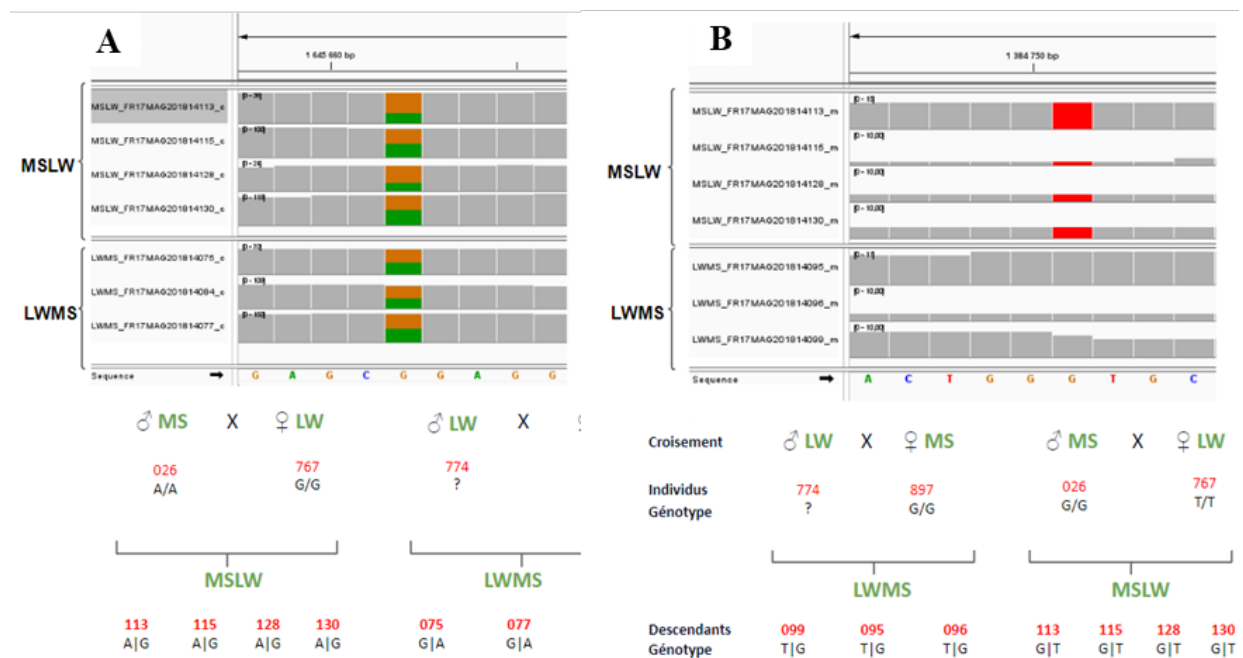


Figure 47 : Visualisation des données transcriptomiques avoisinant 2 variants d’intérêt.

A- Exemple des données de séquençage ARN provenant du cerveau dans la région du variant 1 645 662 pb. Le transcrit du gène *TSSC4* montre une expression bi-allélique : tous les individus expriment les allèles A (vert) et G (orange). B- Exemple des données de séquençage ARN provenant du muscle dans la région du variant 1 384 751 pb. Le transcrit du gène *HI9* montre une expression maternelle : les individus MSLW expriment uniquement l’allèle T (rouge), tandis que les 3 individus LWMS expriment uniquement l’allèle de référence (G, gris).

3.3.2. Premières expériences de validation de l'expression de gènes soumis à empreinte

Comme précédemment, nous avons souhaité confirmer ou non les résultats obtenus à partir des données -omiques et des analyses bio-informatiques. Pour ce faire, nous avons génotypé les ADNc issus de cerveau et de muscle pour tous les individus du croisement réciproque. Les mises au point faites au préalable pour le génotypage à l'aide de la méthode PACE® ont été réutilisées pour les 3 variants d'intérêt 1 384 751 pb, 1 645 662 pb et 1 859 450 pb.

Cette approche nouvelle n'a malheureusement pas permis de confirmer le statut d'empreinte pour les gènes *HI9* et *KCNQ1OT1* car les résultats sont ambigus et difficilement interprétables. Nous avons tout de même vérifié que les ADNc synthétisés n'étaient pas contaminés avec de l'ADNg. Il sera essentiel de poursuivre les mises au point ou d'utiliser une méthode plus classique comme la RT-qPCR pour évaluer l'expression parentale-spécifique des transcrits d'intérêt. En revanche, l'expression bi-allélique du transcrit TSSC4 ciblé par le variant 1 645 662 pb semble valider dans le muscle et le cerveau comme le montre la Figure 48.

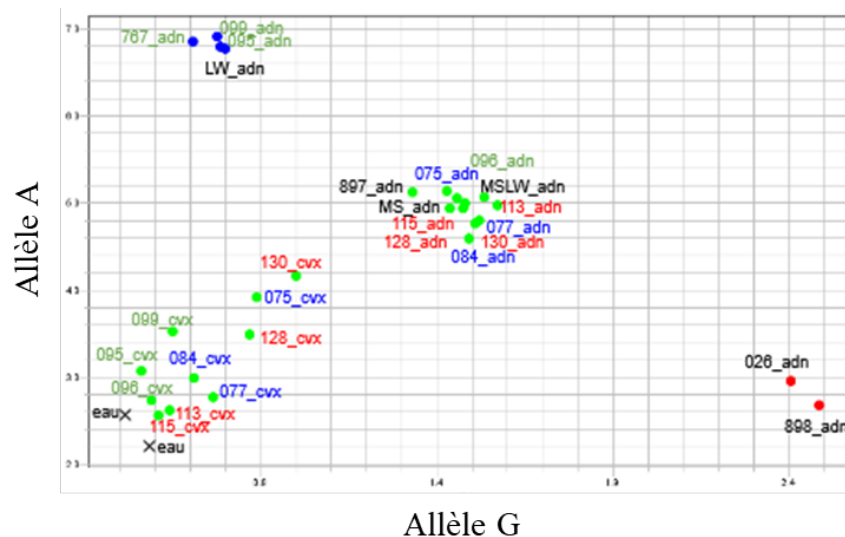


Figure 48 : Génotypage sur ADNg et ADNc du variant 1 645 662 pb à l'aide de la méthode PACE®.

Les individus homozygotes AA et GG sont respectivement colorés en bleu et rouge alors que les individus hétérozygotes A/G sont en vert. Les ADNg et ADNc semblent formés 2 groupes d'hétérozygotes composés des mêmes individus.

Ainsi, ces résultats préliminaires nécessitent d'être approfondis afin d'évaluer précisément quels sont les transcrits exprimés dans le cerveau et le muscle des porcelets pour améliorer l'annotation du génome du porc dans cette région. En effet, il est important de rappeler que les gènes *HI9* et *IGF2* ne sont pas annotés dans la version actuelle utilisée ici (fichier d'annotation gtf 11.1.109).

IV. DISCUSSION

4.1. Mise en évidence de deux régions soumises à empreinte parentale

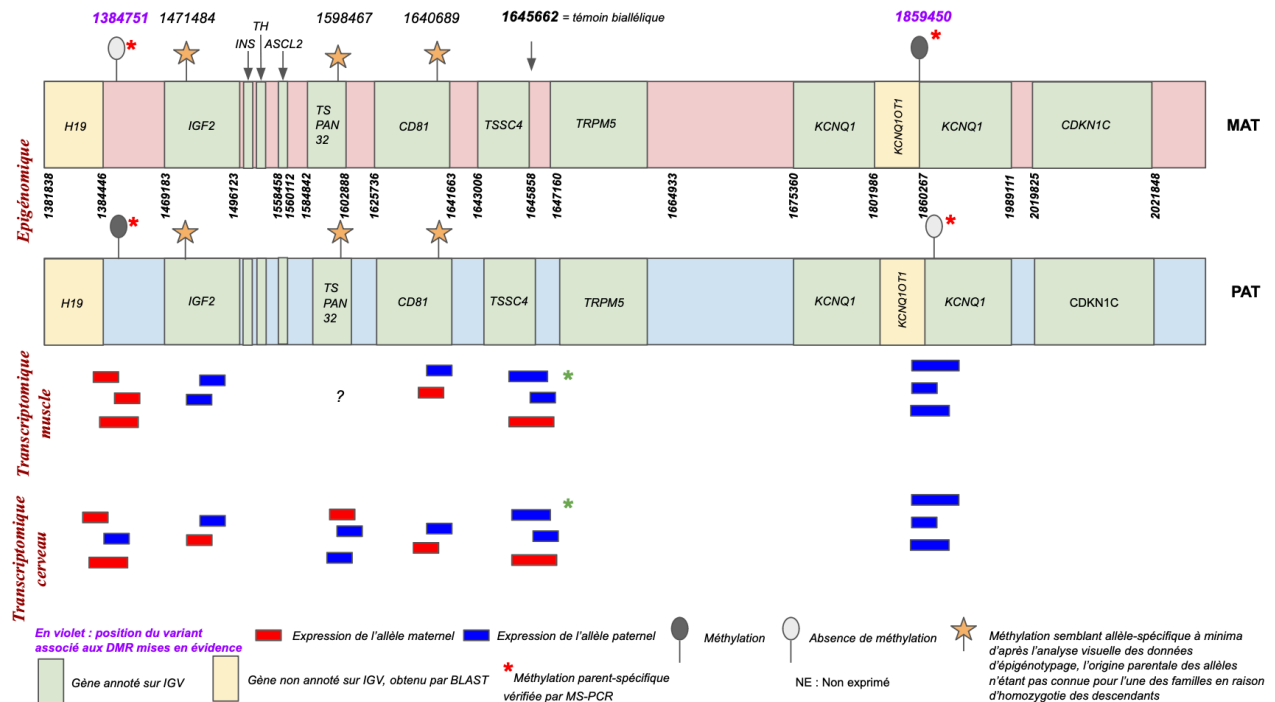


Figure 49 : Schématisation des résultats obtenus concernant l'empreinte parentale de la région *H19-IGF2/KCNQ1-CDKN1C* chez le porc (chromosome 2, entre les positions 1 381 838 bp et 2 059 694 bp).

En verts : les gènes ; en rouge : chromosome maternel ; en bleu : chromosome paternel.

La région soumise à empreinte *H19-IGF2/KCNQ1-CDKN1C* a été caractérisée finement par des approches combinant technologies -omiques et techniques de biologie moléculaire à partir de l'exploitation de croisements réciproques entre les races porcines LW et MS (Figure 49).

L'expression de l'origine parentale des transcrits a été réalisée au travers de l'étude des données transcriptomiques. Les résultats préliminaires confirment pour certains gènes le profil d'empreinte dans le cerveau et le muscle en comparaison à d'autres études menées chez le porc, l'Homme et la souris. A titre d'exemple, l'expression paternelle pour le gène *IGF2* semble spécifique du tissu musculaire puisqu'une expression bi-allélique est observé dans le cerveau. Ceci conforte les données obtenues chez le porc et l'Homme où le foie et le cerveau de fœtus et d'adultes montrent une expression bi-allélique d'*IGF2* (Pham et al. 1998; Wrzeska et al. 2006; Ahn et al. 2022). Par ailleurs, l'ARNinc *H19* présente une expression maternelle dans le muscle et ne semble pas exprimé dans le cerveau des porcelets contrairement au cerveau de fœtus humains (Pham et al. 1998) suggérant un changement important entre les stades pré- et post-natal. Il est cependant nécessaire d'analyser les données transcriptomiques plus finement

et globalement le long du cluster *H19-IGF2/KCNQ1-CDKN1C* afin de lever des ambiguïtés comme celle autour de l'expression de *KCNQ1*. En effet, Yin-Quao et *al.* suggèrent une expression paternelle de *KCNQ1* dans le cerveau chez le porc alors que celui-ci est à expression maternelle chez l'Homme et la souris (Yin-Qiao Wu et al. 2020). Mes résultats tendent à montrer que l'expression paternelle détectée au niveau de *KCNQ1* correspondrait en réalité à l'ARNlnc *KCNQ1OT1* non annoté chez le porc et hébergé dans le gène *KCNQ1*. Ainsi, en considérant *KCNQ1OT1*, mes résultats sont concordants avec les données humaine et bovine qui montrent une expression paternelle dans différents tissus (Robbins et al. 2012).

L'empreinte génomique parentale est un modèle original de régulation épigénétique puisque l'expression parentale-spécifique est régulée par l'apposition de marques épigénétiques (méthylation de l'ADN ou histones) dépendante de l'origine parentale. Ici, j'ai montré, via l'analyse des données épigénomiques et des approches ciblées, que deux régions proches de 1 384 751 pb et 1 859 450 pb s'apparentent respectivement à des DMR paternelle et maternelle. Bien que l'ensemble des gènes de la région ne soit pas annoté correctement dans la version utilisée du génome du porc, la DMR paternelle est localisée entre les gènes *H19* et *IGF2* alors que la DMR maternelle est positionnée dans le gène *KCNQ1* et en amont du gène qui héberge *KCNQ1OT1*. La taille des régions identifiées ainsi que leur localisation suggèrent fortement que nous avons identifié les centres de contrôle d'empreinte, appelés ICR1 et ICR2 chez l'Homme et la souris (Shmela, Gicquel 2013). Ces résultats sont concordants avec ceux obtenus chez le porc au moins pour ICR1 (Braunschweig, Owczarek-Lipska, Stahlberger-Saitbekova 2011). Par ailleurs, il sera important de poursuivre les analyses de méthylation afin d'identifier d'éventuels d'autres DMR de cette région d'intérêt comme les DMR somatiques en amont du gène *H19*, *H19* prom, ou *IGF2* DMR0 et DMR2 (Monk et al. 2019). En effet, l'exploitation d'échantillons de semence (non mentionné dans cette étude) présentera l'intérêt de distinguer l'origine germinale ou somatique des DMR identifiés (Akbari et al. 2022; Jima et al. 2022).

Le nombre de gènes soumis à empreinte représente environ 1 à 2% des gènes annotés dans le génome des mammifères. Leur mode de régulation atypique induit à ce titre un mode de transmission atypique connu sous le nom d'héritabilité génétique non-mendélienne. Des mutations dans certains de ces gènes sont d'ailleurs responsables de phénotypes agronomiques (Cockett et al. 1996; Van Laere et al. 2003). Même si peu nombreux, les effets des gènes soumis à empreinte sur les phénotypes peuvent être importants et représentent donc un domaine d'étude à ne pas négliger. La communauté scientifique continue d'ailleurs de fournir des preuves de

l'implication de l'empreinte génomique parentale dans la croissance et le développement des espèces de rente (Ibeagha-Awemu, Zhao 2015). Dans le cas des caractères influencés par des effets maternels qui sont mesurés et estimés à partir des performances des descendants comme le poids de naissance ou la croissance avant sevrage, les gènes soumis à empreinte génomique parentale représentent une cible d'intérêt. En effet, les gènes soumis à empreinte sont connus pour jouer un rôle majeur dans le développement et la croissance foétale et postnatale (Monk et al. 2019). Par exemple, une étude récente menée chez l'Homme a montré que la transmission d'un allèle paternel dans le gène *DLKI* était associée à une réduction significative du poids de naissance (Moore et al. 2015). Bien qu'à l'heure actuelle, l'empreinte parentale ne soit pas intégrée en routine dans les modèles d'évaluation des valeurs génétiques, adapter les modèles en tenant compte de ces phénomènes non-mendéliens permettrait d'améliorer l'estimation de la valeur des reproducteurs, et donc l'efficacité de la sélection (Goddard, Whitelaw 2014). Malgré tout, avec le développement de nouvelles analyses génomiques et épigénomiques, il est possible qu'à l'avenir, les éleveurs ne s'appuient pas seulement sur les informations génétiques mais utilisent également l'épigénome des candidats à la sélection.

4.2. Une approche multiomique locale posant des bases pour la caractérisation de l'imprintome porcin

L'approche multiomique qui a permis de mettre en évidence les ICR de la région porcine homologue à la région 11p15 de l'Homme contribue à des développements visant à identifier des signatures moléculaires de l'empreinte à plus large échelle chez le porc. Des auteurs influents dans le domaine de l'empreinte utilisent le terme « imprintome » pour désigner la portion de l'épigénome qui correspond spécifiquement aux ICR (Jima et al. 2022). Alors que des cartes détaillées de l'imprintome commencent à voir le jour chez les espèces modèles (Jima et al. 2022; Richard Albert et al. 2023; Andergassen et al. 2017), la poursuite et l'extension à d'autres régions du travail mené ici devrait aboutir à une première carte de l'imprintome porcin. Dans cette optique, mes résultats constituent un appui à l'identification d'options pertinentes pour la détection des motifs d'intérêt, de même que dans la mise au jour de certaines difficultés expérimentales.

Mon travail de détection *in silico* par exploration IGV des régions hémi-méthylées a confirmé qu'elles pouvaient s'étendre au-delà de 1 000 pb chez le porc (Ahn, Wu, et al. 2020), et que l'adoption de critères trop stricts d'hémi-méthylation ou de fusionnement de régions voisines risquerait d'aboutir à la détection de DMR incomplètes. Concernant la détection des variants,

mes résultats expérimentaux ont confirmé la validité des haplotypes identifiés bioinformatiquement dans la région *H19/IGF2-KCNQ1/CDKN1C*, ce qui confirme par la même occasion la pertinence des outils choisis, pour certains desquels il existe encore peu de travaux comparatifs des performances sur données biologiques (Lindner et al. 2022). Par ailleurs, une stratégie relativement stricte a été adoptée pour la découverte du polymorphisme, ce qui s'est traduit par la sélection de seulement deux variants informatifs (en positions 1 384 751 pb et 1 859 450 pb) pour les expériences de validation (PACE® sur ADNc et MS-PCR). Compte tenu du faible nombre d'incohérences relevées sur les génotypes (< 5%), certains critères utilisés lors de l'analyse bioinformatique pourraient être relâchés afin d'éviter d'écarter a priori des variants dont la localisation est intéressante.

Pour des raisons de chronologie des analyses et de disponibilité des échantillons, les données transcriptomiques n'ont été utilisées ici qu'a posteriori, afin de valider les modalités d'expression d'un sous-ensemble de variants présélectionnés grâce aux données épigénomiques. Il pourrait être intéressant d'effectuer l'analyse transcriptomique sur l'ensemble de la région sans a priori, pour identifier tous les variants à expression mono-allélique disponibles, et ensuite distinguer ceux présentant une expression parent-dépendante de ceux arborant une simple expression allèle-spécifique. Une approche de ce type est d'ores et déjà envisagée par mon équipe d'accueil sur plusieurs régions d'intérêt.

Ces données transcriptomiques sont d'autant plus intéressantes qu'il existe des marges d'amélioration de l'annotation du génome du porc, en particulier dans les régions soumises à empreinte qui présentent une certaine complexité transcriptionnelle (Ahn, Wu, et al. 2020; Ahn, Hwang, et al. 2020b). Une analyse plus approfondie et systématisée des données transcriptomiques disponibles pourrait permettre de résoudre une partie de cette complexité, et par conséquent d'identifier un plus grand nombre de variants exprimés.

Les travaux rapportés ont aussi permis de mesurer certaines difficultés pouvant affecter les expériences dans le champ de la génétique. Tout d'abord, la quantité de tissu disponible par animal pour amplifier le matériel génétique est limitée et certaines réplifications ou expériences sont donc impossibles en l'absence de tissu ou lorsque sa qualité est dégradée. Nous avons été confrontés à ce problème pour un verrat de l'une des familles, et l'absence de certaines données pour cet individu complique par conséquent la détermination de l'origine parentale des allèles de ses descendants. D'autre part, le temps nécessaire à la mise au point d'expériences allèle-spécifiques, incontournables quand on s'intéresse à l'empreinte parentale, dépend d'éléments qui ne sont pas a priori contrôlables. Nous avons essayé de faire les choix les plus parcimonieux

à la lumière des données disponibles, mais de nouveaux essais sont nécessaires afin de systématiser les expériences de PACE® ou de MS-PCR du type de celles que j'ai menées, afin de les utiliser sur un grand nombre de variants d'intérêt et ainsi étudier différentes régions du génome soumises à empreinte simultanément.

CONCLUSION

Cette étude a permis dans un premier temps de réaliser une synthèse générale sur la génétique en filière porcine selon trois axes : (i) un axe historique présentant l'organisation de la sélection au sein de la filière porcine, (ii) un axe technologique soulignant dans quelles mesures les évolutions technologiques dans le domaine de la génétique ont permis d'optimiser cette sélection, (iii) un axe plus prospectif proposant une synthèse concernant les gènes soumis à empreinte parentale dont une connaissance fine pourrait contribuer à améliorer les modèles de sélection génétique.

Dans un deuxième temps, la partie expérimentale de cette étude a proposé de rechercher des signatures moléculaires caractéristiques de l'empreinte parentale dans une région (homologue à la région 11p15 humaine) du génome porcin. Elle a notamment permis de mettre en évidence dans l'espèce porcine deux grandes régions (> 1200 pb) riches en CpG hémi-méthylés dont les données épigénomiques et transcriptomiques permettent de conclure à une méthylation et une expression parent-dépendante caractéristique de l'empreinte parentale. Cependant, pour étudier un plus grand nombre de régions, des mises points supplémentaires de méthodes pour l'étude de l'expression (comme la PACE® sur ADNc, afin d'utiliser une méthode moins coûteuse que le séquençage d'ARN) et de l'origine parentale de la méthylation (comme la MS-PCR) sont nécessaires.

Enfin, comme évoqué dans ce manuscrit, les cinquante dernières années se caractérisent par une véritable révolution technologique permettant une connaissance toujours plus riche du génome de différentes espèces. Or, les données génomiques (« facilement » accessibles aujourd'hui) ne permettant pas à elles seules d'expliquer en totalité le phénotype, c'est pourquoi le développement de technologies visant à étudier des données épigénomiques pourrait constituer la quatrième révolution technologique. En effet, que ce soit en sélection animale, ou (et a fortiori) en médecine humaine, l'étude de l'expression du génome et de l'interaction entre génome et environnement constitue un sujet suscitant l'intérêt de nombreux chercheurs et acteurs de ces domaines.

Bibliographie

ADRIAENSSENS, Eric, LOTTIN, Séverine, BERTEAUX, Nathalie, HORNEZ, Louis, FAUQUETTE, William, FAFEUR, Véronique, PEYRAT, Jean-Philippe, LE BOURHIS, Xuefen, HONDERMARCK, Hubert, COLL, Jean, DUGIMONT, Thierry et CURGY, Jean-Jacques, 2002. Cross-Talk between Mesenchyme and Epithelium Increases H19 Gene Expression during Scattering and Morphogenesis of Epithelial Cells. *Experimental Cell Research*. mai 2002. Vol. 275, n° 2, pp. 215-229. DOI 10.1006/excr.2002.5500.

AHN, Jinsoo, HWANG, In-Sul, PARK, Mi-Ryung, CHO, In-Cheol, HWANG, Seongsoo et LEE, Kichoon, 2020a. The Landscape of Genomic Imprinting at the Porcine *SGCE/PEG10* Locus from Methylome and Transcriptome of Parthenogenetic Embryos. *G3 (Bethesda, Md.)*. 5 novembre 2020. Vol. 10, n° 11, pp. 4037-4047. DOI 10.1534/g3.120.401425.

AHN, Jinsoo, HWANG, In-Sul, PARK, Mi-Ryung, CHO, In-Cheol, HWANG, Seongsoo et LEE, Kichoon, 2020b. The Landscape of Genomic Imprinting at the Porcine *SGCE / PEG10* Locus from Methylome and Transcriptome of Parthenogenetic Embryos. *G3 Genes|Genomes|Genetics*. 1 novembre 2020. Vol. 10, n° 11, pp. 4037-4047. DOI 10.1534/g3.120.401425.

AHN, Jinsoo, LEE, Joonbum, KIM, Dong-Hwan, HWANG, In-Sul, PARK, Mi-Ryung, CHO, In-Cheol, HWANG, Seongsoo et LEE, Kichoon, 2022. Loss of Monoallelic Expression of *IGF2* in the Adult Liver Via Alternative Promoter Usage and Chromatin Reorganization. *Frontiers in Genetics*. [en ligne]. 2022. Vol. 13. [Consulté le 17 juin 2023]. Disponible à l'adresse: <https://www.frontiersin.org/articles/10.3389/fgene.2022.920641>

AHN, Jinsoo, WU, Huiguang, LEE, Joonbum, HWANG, In-Sul, YU, Debing, AHN, Jin-Seop, LEE, Jeong-Woong, HWANG, Seongsoo et LEE, Kichoon, 2020. Identification of a Novel Imprinted Transcript in the Porcine *GNAS* Complex Locus Using Methylome and Transcriptome of Parthenogenetic Fetuses. *Genes*. 14 janvier 2020. Vol. 11, n° 1, pp. 96. DOI 10.3390/genes11010096.

AKBARI, Vahid, GARANT, Jean-Michel, O'NEILL, Kieran, PANDOH, Pawan, MOORE, Richard, MARRA, Marco A, HIRST, Martin et JONES, Steven JM, 2022. Genome-wide detection of imprinted differentially methylated regions using nanopore sequencing. TYLER, Jessica K et KELSEY, Gavin (éd.), *eLife*. 5 juillet 2022. Vol. 11, pp. e77898. DOI 10.7554/eLife.77898.

ALLENDORF, F. W., FUNK, W. C., AITKEN, S. N., BYRNE, M. et LUIKART, G. H., 2022. *Conservation and the genomics of populations*. . 3rd. Oxford University Press. ISBN 978-0-19-885657-3.

AMARGER, Valérie, NGUYEN, Minh, VAN LAERE, Anne-Sophie, BRAUNSCHWEIG, Martin, NEZER, Carine, GEORGES, Michel et ANDERSSON, Leif, 2002. Comparative sequence analysis of the *INS-IGF2-H19* gene cluster in pigs. *Mammalian Genome: Official Journal of the International Mammalian Genome Society*. juillet 2002. Vol. 13, n° 7, pp. 388-398. DOI 10.1007/s00335-001-3059-x.

ANDERGASSEN, Daniel, DOTTER, Christoph P, WENZEL, Daniel, SIGL, Verena, BAMMER, Philipp C, MUCKENHUBER, Markus, MAYER, Daniela, KULINSKI, Tomasz M, THEUSSL, Hans-Christian, PENNINGER, Josef M, BOCK, Christoph, BARLOW, Denise P, PAULER, Florian M et HUDSON, Quanah J, 2017. Mapping the mouse Allelome reveals tissue-specific regulation of allelic expression. *eLife*. 14 août 2017. Vol. 6, pp. e25125. DOI 10.7554/eLife.25125.

ARGÜELLO, Héctor, ESTELLÉ, Jordi, LEONARD, Finola C., CRISPIE, Fiona, COTTER, Paul D., O'SULLIVAN, Orla, LYNCH, Helen, WALIA, Kavita, DUFFY, Geraldine, LAWLOR, Peadar G. et GARDINER, Gillian E., 2019. Influence of the Intestinal Microbiota on Colonization Resistance to Salmonella and the Shedding Pattern of Naturally Exposed Pigs. *mSystems*. 2019. Vol. 4, n° 2, pp. e00021-19. DOI 10.1128/mSystems.00021-19.

ARNESON, Adriana, HAGHANI, Amin, THOMPSON, Michael J., PELLEGRINI, Matteo, KWON, Soo Bin, VU, Ha, MACIEJEWSKI, Emily, YAO, Mingjia, LI, Caesar Z., LU, Ake T., MORSELLI, Marco, RUBBI, Liudmilla, BARNES, Bret, HANSEN, Kasper D., ZHOU, Wanding, BREEZE, Charles E., ERNST, Jason et HORVATH, Steve, 2022. A mammalian methylation array for profiling methylation levels at conserved sequences. *Nature Communications*. 10 février 2022. Vol. 13, pp. 783. DOI 10.1038/s41467-022-28355-z.

ASSOCIATION FRANÇAISE DU SYNDROME D'ANGELMAN, 2021. *Veille scientifique concernant le syndrome d'Angelman* [en ligne]. 2021. Disponible à l'adresse: <https://www.angelman-afsa.org/fichiers/Veille%20scientifique%20%20Septembre%202021%20-1631539675.pdf>

AYKROYD, Bethany R. L., TUNSTER, Simon J. et SFERRUZZI-PERRI, Amanda N., 2022. Loss of imprinting of the Igf2-H19 ICR1 enhances placental endocrine capacity via sex-specific alterations in signalling pathways in the mouse. *Development (Cambridge, England)*. 1 janvier 2022. Vol. 149, n° 1, pp. dev199811. DOI 10.1242/dev.199811.

BARAN, Yael, SUBRAMANIAM, Meena, BITON, Anne, TUKIAINEN, Taru, TSANG, Emily K., RIVAS, Manuel A., PIRINEN, Matti, GUTIERREZ-ARCELUS, Maria, SMITH, Kevin S., KUKURBA, Kim R., ZHANG, Rui, ENG, Celeste, TORGERSON, Dara G., URBANEK, Cydney, CONSORTIUM, the GTEx, LI, Jin Billy, RODRIGUEZ-SANTANA, Jose R., BURCHARD, Esteban G., SEIBOLD, Max A., MACARTHUR, Daniel G., MONTGOMERY, Stephen B., ZAITLEN, Noah A. et LAPPALAINEN, Tuuli, 2015. The landscape of genomic imprinting across diverse adult human tissues. *Genome Research*. 7 janvier 2015. Vol. 25, n° 7, pp. 927-936. DOI 10.1101/gr.192278.115.

BARIL, Claire et BERGONZINI, Jean-Claude, 1994. *De la Génétique des Populations à la Génétique Quantitative*. . Département Forestier du Centre de Coopération Internationale en Recherche Agronomique pour le Développement. Département Forestier du Centre de Coopération.

BARKO, P.C., MCMICHAEL, M.A., SWANSON, K.S. et WILLIAMS, D.A., 2018. The Gastrointestinal Microbiome: A Review. *Journal of Veterinary Internal Medicine*. 2018. Vol. 32, n° 1, pp. 9-25. DOI 10.1111/jvim.14875.

BARLOW, D. P., STÖGER, R., HERRMANN, B. G., SAITO, K. et SCHWEIFER, N., 1991. The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the Tme locus. *Nature*. 3 janvier 1991. Vol. 349, n° 6304, pp. 84-87. DOI 10.1038/349084a0.

BARLOW, Denise P. et BARTOLOMEI, Marisa S., 2014. Genomic imprinting in mammals. *Cold Spring Harbor Perspectives in Biology*. 1 février 2014. Vol. 6, n° 2, pp. a018382. DOI 10.1101/cshperspect.a018382.

BARTOLOMEI, M. S., ZEMEL, S. et TILGHMAN, S. M., 1991. Parental imprinting of the mouse H19 gene. *Nature*. 9 mai 1991. Vol. 351, n° 6322, pp. 153-155. DOI 10.1038/351153a0.

BARTOLOMEI, Marisa S. et FERGUSON-SMITH, Anne C., 2011. Mammalian genomic imprinting. *Cold Spring Harbor Perspectives in Biology*. 1 juillet 2011. Vol. 3, n° 7, pp. a002592. DOI 10.1101/cshperspect.a002592.

- BATTRAM, Thomas, GAUNT, Tom R., RELTON, Caroline L., TIMPSON, Nicholas J. et HEMANI, Gibran, 2022. A comparison of the genes and genesets identified by GWAS and EWAS of fifteen complex traits. *Nature Communications*. 19 décembre 2022. Vol. 13, pp. 7816. DOI 10.1038/s41467-022-35037-3.
- BECKMAN, J. S. et WEBER, J. L., 1992. Survey of human and rat microsatellites. *Genomics*. avril 1992. Vol. 12, n° 4, pp. 627-631. DOI 10.1016/0888-7543(92)90285-z.
- BIDANEL, J. P., BOICHARD, D. et CHEVALET, C., 2008. De la génétique à la génomique. *INRAE Productions Animales*. 20 avril 2008. Vol. 21, n° 1, pp. 15-32. DOI 10.20870/productions-animales.2008.21.1.3372.
- BIDANEL, J. P., ROY, P. Le, OLLIVIER, L., BONNEAU, M., CHARDON, P., ELSÉN, J. M., GELLIN, J. et MILAN, D., 1996. Etablissement et utilisation de la carte génétique porcine. *INRAE Productions Animales*. 17 août 1996. Vol. 9, n° 4, pp. 299-310. DOI 10.20870/productions-animales.1996.9.4.4063.
- BIDANEL, Jean-Pierre, SILALAH, Parsaoran, TRIBOUT, Thierry, CANARIO, Laurianne, DUCOS, Alain, GARREAU, Hervé, GILBERT, Hélène, LARZUL, Catherine, MILAN, Denis, RIQUET, Juliette, SCHWOB, Sandrine, MERCAT, Marie-José, HASSENFRAZ, Claire, BOUQUET, Alain, BAZIN, Christophe et BIDANEL, Joel, 2020. Cinquante années d'amélioration génétique du porc en France : bilan et perspectives. *INRAE Productions Animales*. [en ligne]. 4 mai 2020. Vol. 33, n° 1. [Consulté le 3 août 2022]. DOI 10.20870/productions-animales.2020.33.1.3092.
- BIDANEL, Jean-Pierre, SILALAH, Parsaoran, TRIBOUT, Thierry, CANARIO, Laurianne, DUCOS, Alain, GARREAU, Hervé, GILBERT, Hélène, LARZUL, Catherine, MILAN, Denis, RIQUET, Juliette, SCHWOB, Sandrine, MERCAT, Marie-José, HASSENFRAZ, Claire, BOUQUET, Alain, BAZIN, Christophe et BIDANEL, Joel, 2020. Cinquante années d'amélioration génétique du porc en France : bilan et perspectives. *INRAE Productions Animales*. [en ligne]. 29 janvier 2020. Vol. 33, n° 1. [Consulté le 17 août 2022]. DOI 10.20870/productions-animales.2020.33.1.3092.
- BOICHARD, D, FRITZ, S, ROSSIGNOL, M N, BOSCHER, M Y, MALAFOSSE, A et COLLEAU, J J, 2002. Implementation of marker-assisted selection in french dairy cattle. . 2002.
- BOICHARD, D., GUILLAUME, F., BAUR, A., CROISEAU, P., ROSSIGNOL, M. N., BOSCHER, M. Y., DRUET, T., GENESTOUT, L., COLLEAU, J. J., JOURNAUX, L., DUCROCQ, V. et FRITZ, S., 2012. Genomic selection in French dairy cattle. *Animal Production Science*. 2012. Vol. 52, n° 3, pp. 115. DOI 10.1071/AN11119.
- BOICHARD, ROY, P. Le, LEVÉZIEL, H. et ELSÉN, J. M., 1998a. Utilisation des marqueurs moléculaires en génétique animale. *INRAE Productions Animales*. 2 février 1998. Vol. 11, n° 1, pp. 67-80. DOI 10.20870/productions-animales.1998.11.1.3918.
- BOICHARD, ROY, P. Le, LEVÉZIEL, H. et ELSÉN, J. M., 1998b. Utilisation des marqueurs moléculaires en génétique animale. *INRAE Productions Animales*. 2 février 1998. Vol. 11, n° 1, pp. 67-80. DOI 10.20870/productions-animales.1998.11.1.3918.
- BOUQUET, Alban, 2017. Mise en place de la sélection génomique dans le schéma de sélection de la population Landrace Français. *Journées de la Recherche Porcine*. 2017. N° 2017, pp. 31-36.
- BRATKOVIČ, Tomaž, BOŽIČ, Janja et ROGELJ, Boris, 2020. Functional diversity of small nucleolar RNAs. *Nucleic Acids Research*. 28 février 2020. Vol. 48, n° 4, pp. 1627-1651.

DOI 10.1093/nar/gkz1140.

BRAUNSCHWEIG, Martin H., OWCZAREK-LIPSKA, Marta et STAHLBERGER-SAITBEKOVA, Nasikhat, 2011. Relationship of porcine IGF2 imprinting status to DNA methylation at the H19 DMD and the IGF2 DMRs 1 and 2. *BMC Genetics*. 17 mai 2011. Vol. 12, n° 1, pp. 47. DOI 10.1186/1471-2156-12-47.

BRAUNSCHWEIG, Martin, JAGANNATHAN, Vidhya, GUTZWILLER, Andreas et BEE, Giuseppe, 2012. Investigations on transgenerational epigenetic response down the male line in F2 pigs. *PLoS One*. 2012. Vol. 7, n° 2, pp. e30583. DOI 10.1371/journal.pone.0030583.

BREITWIESER, Florian P., LU, Jennifer et SALZBERG, Steven L., 2019. A review of methods and databases for metagenomic classification and assembly. *Briefings in Bioinformatics*. 19 juillet 2019. Vol. 20, n° 4, pp. 1125-1136. DOI 10.1093/bib/bbx120.

BRENAUT, Pauline, 2021. Nouveaux critères de sélection des aptitudes maternelles des truies. *Bilan d'activité IFIP*. [en ligne]. 2021. [Consulté le 17 mai 2023]. Disponible à l'adresse: <https://ifip.asso.fr/documentations/41480-nouveaux-criteres-de-selection-des-aptitudes-maternelles-des-truies/>

BRINK, R. A., 1956. A Genetic Change Associated with the R Locus in Maize Which Is Directed and Potentially Reversible. *Genetics*. novembre 1956. Vol. 41, n° 6, pp. 872-889. DOI 10.1093/genetics/41.6.872.

BRISKEN, Cathrin, AYYANNAN, Ayyakkannu, NGUYEN, Cuc, HEINEMAN, Anna, REINHARDT, Ferenc, TAN, Jian, DEY, S. K., DOTTO, G. Paolo et WEINBERG, Robert A., 2002. IGF-2 is a mediator of prolactin-induced morphogenesis in the breast. *Developmental Cell*. décembre 2002. Vol. 3, n° 6, pp. 877-887. DOI 10.1016/s1534-5807(02)00365-9.

BUTLER, Merlin G., MILLER, Jennifer L. et FORSTER, Janice L., 2019. Prader-Willi Syndrome - Clinical Genetics, Diagnosis and Treatment Approaches: An Update. *Current Pediatric Reviews*. 2019. Vol. 15, n° 4, pp. 207-244. DOI 10.2174/1573396315666190716120925.

CALDWELL, Blake A. et BARTOLOMEI, Marisa S., 2022. DNA methylation reprogramming of genomic imprints in the mammalian germline: A TET-centric view. *Andrology*. 23 septembre 2022. DOI 10.1111/andr.13303.

CALENGE, F., MARTIN, C., FLOCH, N. Le, PHOCAS, F., MORGAVI, D., ROGEL-GAILLARD, C. et QUÉRÉ, P., 2014. Intégrer la caractérisation du microbiote digestif dans le phénotypage de l'animal de rente : vers un nouvel outil de maîtrise de la santé en élevage ? *INRAE Productions Animales*. 28 août 2014. Vol. 27, n° 3, pp. 209-222. DOI 10.20870/productions-animales.2014.27.3.3068.

CANARIO, Laurianne, 2021. Eléments de génétique quantitative et application aux populations animales. *Journées de la Recherche Porcine*. [en ligne]. 2021. N° 53. [Consulté le 1 mai 2023]. Disponible à l'adresse: <https://hal.inrae.fr/hal-03135223/document>

CANIBE, Nuria, O'DEA, Mark et ABRAHAM, Sam, 2019. Potential relevance of pig gut content transplantation for production and research. *Journal of Animal Science and Biotechnology*. 2019. Vol. 10, pp. 55. DOI 10.1186/s40104-019-0363-4.

CASSIDY, Féaron C. et CHARALAMBOUS, Marika, 2018. Genomic imprinting, growth and maternal-fetal interactions. SUAREZ, Raul K. et HOPPELER, Hans H. (éd.), *Journal of Experimental Biology*. 7 mars 2018. Vol. 221, n° Suppl_1, pp. jeb164517. DOI 10.1242/jeb.164517.

- CASSIDY, Suzanne B., SCHWARTZ, Stuart, MILLER, Jennifer L. et DRISCOLL, Daniel J., 2012. Prader-Willi syndrome. *Genetics in Medicine: Official Journal of the American College of Medical Genetics*. janvier 2012. Vol. 14, n° 1, pp. 10-26. DOI 10.1038/gim.0b013e31822bead0.
- CASTILLO-FERNANDEZ, Juan E., SPECTOR, Tim D. et BELL, Jordana T., 2014. Epigenetics of discordant monozygotic twins: implications for disease. *Genome Medicine*. 31 juillet 2014. Vol. 6, n° 7, pp. 60. DOI 10.1186/s13073-014-0060-z.
- CATTANACH, B. M. et KIRK, M., 1985. Differential activity of maternally and paternally derived chromosome regions in mice. *Nature*. 6 juin 1985. Vol. 315, n° 6019, pp. 496-498. DOI 10.1038/315496a0.
- CHAO, Wendy et D'AMORE, Patricia A., 2008. IGF2: epigenetic regulation and role in development and disease. *Cytokine & Growth Factor Reviews*. avril 2008. Vol. 19, n° 2, pp. 111-120. DOI 10.1016/j.cytogfr.2008.01.005.
- CHARALAMBOUS, Marika, COWLEY, Michael, GEOGHEGAN, Fleur, SMITH, Florentia M., RADFORD, Elizabeth J., MARLOW, Benjamin P., GRAHAM, Christopher F., HURST, Laurence D. et WARD, Andrew, 2010. Maternally-inherited Grb10 reduces placental size and efficiency. *Developmental Biology*. 1 janvier 2010. Vol. 337, n° 1, pp. 1-8. DOI 10.1016/j.ydbio.2009.10.011.
- CHAVATTE-PALMER, Pascale et TARRADE, Anne, 2016. Placentation in different mammalian species. *Annales d'Endocrinologie*. juin 2016. Vol. 77, n° 2, pp. 67-74. DOI 10.1016/j.ando.2016.04.006.
- CHELH, I., RODRIGUEZ, J., BONNIEU, A., CASSAR-MALEK, I., COTTIN, P., GABILLARD, J. C., LEIBOVITCH, S., SASSI, A. Hadj, SEILIEZ, I. et PICARD, B., 2009. La myostatine : un régulateur négatif de la masse musculaire chez les vertébrés. *INRAE Productions Animales*. 9 décembre 2009. Vol. 22, n° 5, pp. 397-408. DOI 10.20870/productions-animales.2009.22.5.3364.
- CHEMINEAU, Philippe, 2014. *La Reproduction Animale et Humaine*. . Versailles : Quae. Syntheses. ISBN 978-2-7592-2208-7. absent
- CHEN, Congying, ZHOU, Yunyan, FU, Hao, XIONG, Xinwei, FANG, Shaoming, JIANG, Hui, WU, Jinyuan, YANG, Hui, GAO, Jun et HUANG, Lusheng, 2021. Expanded catalog of microbial genes and metagenome-assembled genomes from the pig gut microbiome. *Nature Communications*. 17 février 2021. Vol. 12, n° 1, pp. 1106. DOI 10.1038/s41467-021-21295-0.
- CHEN, Haiying, SUN, Manni, LIU, Jing, TONG, Chunxiao et MENG, Tao, 2015. Silencing of Paternally Expressed Gene 10 Inhibits Trophoblast Proliferation and Invasion. *PLoS ONE*. 17 décembre 2015. Vol. 10, n° 12, pp. e0144845. DOI 10.1371/journal.pone.0144845.
- CHEN, Yinwei, ZHOU, Jinghua et WANG, Li, 2021. Role and Mechanism of Gut Microbiota in Human Disease. *Frontiers in Cellular and Infection Microbiology*. 17 mars 2021. Vol. 11, pp. 625913. DOI 10.3389/fcimb.2021.625913.
- CHEVALET, C. et BOICHARD, D., 1992. Sélection assistée par marqueurs. *INRAE Productions Animales*. 2 décembre 1992. Vol. 5, n° HS, pp. 291-294. DOI 10.20870/productions-animales.1992.5.HS.4307.
- CNRTL, [sans date]. Définition de la parthénogénèse. [en ligne]. Disponible à l'adresse: <https://cnrtl.fr/definition/academie9/parth%C3%A9nogen%C3%A8se>

COCKETT, Noelle E, JACKSON, Samp P, SHAY, Tracy L, FARNIR, Frédéric, BERGHMANS, Stéphane, SNOWDER, Dary D, NIELSEN, Dahlia M et GEORGES, Michel, 1996. Polar Overdominance at the Ovine callipyge Locus. *Science*. [en ligne]. 1996. Vol. 273. [Consulté le 17 juin 2023]. DOI 10.1126/science.273.5272.236.

CONGRAS, Annabelle, YERLE-BOUISSOU, Martine, PINTON, Alain, VIGNOLES, Florence, LIAUBET, Laurence, FERCHAUD, Stéphane et ACLOQUE, Hervé, 2014. Sperm DNA methylation analysis in swine reveals conserved and species-specific methylation patterns and highlights an altered methylation at the GNAS locus in infertile boars. *Biology of Reproduction*. décembre 2014. Vol. 91, n° 6, pp. 137. DOI 10.1095/biolreprod.114.119610.

CORTIJO, Sandra, WARDENAAR, René, COLOMÉ-TATCHÉ, Maria, GILLY, Arthur, ETCHEVERRY, Mathilde, LABADIE, Karine, CAILLIEUX, Erwann, HOSPITAL, Frédéric, AURY, Jean-Marc, WINCKER, Patrick, ROUDIER, François, JANSEN, Ritsert C., COLOT, Vincent et JOHANNES, Frank, 2014. Mapping the epigenetic basis of complex traits. *Science (New York, N.Y.)*. 7 mars 2014. Vol. 343, n° 6175, pp. 1145-1148. DOI 10.1126/science.1248127.

CUI, Hengmi, 2007. Loss of imprinting of IGF2 as an epigenetic marker for the risk of human cancer. *Disease Markers*. 2007. Vol. 23, n° 1-2, pp. 105-112. DOI 10.1155/2007/363464.

CURLEY, James P., BARTON, Sheila, SURANI, Azim et KEVERNE, Eric B., 2004. Coadaptation in mother and infant regulated by a paternally expressed imprinted gene. *Proceedings of the Royal Society B: Biological Sciences*. 22 juin 2004. Vol. 271, n° 1545, pp. 1303-1309. DOI 10.1098/rspb.2004.2725.

DAVID, Ingrid, CANARIO, Laurianne, COMBES, Sylvie et DEMARS, Julie, 2019. Intergenerational Transmission of Characters Through Genetics, Epigenetics, Microbiota, and Learning in Livestock. *Frontiers in Genetics*. 31 octobre 2019. Vol. 10, pp. 1058. DOI 10.3389/fgene.2019.01058.

DAVID, Ingrid et RICARD, Anne, 2019. Unified Model for Inclusive Inheritance in Livestock Species. . 2019. Vol. 212, pp. 1075-1099.

DAY, Troy et BONDURIANSKY, Russell, 2004. Intralocus sexual conflict can drive the evolution of genomic imprinting. *Genetics*. août 2004. Vol. 167, n° 4, pp. 1537-1546. DOI 10.1534/genetics.103.026211.

DEAMER, David, AKESON, Mark et BRANTON, Daniel, 2016. Three decades of nanopore sequencing. *Nature Biotechnology*. 6 mai 2016. Vol. 34, n° 5, pp. 518-524. DOI 10.1038/nbt.3423.

DECHIARA, T. M., ROBERTSON, E. J. et EFSTRATIADIS, A., 1991. Parental imprinting of the mouse insulin-like growth factor II gene. *Cell*. 22 février 1991. Vol. 64, n° 4, pp. 849-859. DOI 10.1016/0092-8674(91)90513-x.

DEKKER, Job et MISTELI, Tom, 2015. Long-Range Chromatin Interactions. *Cold Spring Harbor Perspectives in Biology*. octobre 2015. Vol. 7, n° 10, pp. a019356. DOI 10.1101/cshperspect.a019356.

DE KONING, D. J., RATTINK, A. P., HARLIZIUS, B., VAN ARENDONK, J. A., BRASCAMP, E. W. et GROENEN, M. A., 2000. Genome-wide scan for body composition in pigs reveals important role of imprinting. *Proceedings of the National Academy of Sciences of the United States of America*. 5 juillet 2000. Vol. 97, n° 14, pp. 7947-7950. DOI 10.1073/pnas.140216397.

DELANOUE, Elsa, 2017. Points de vue et attentes des consommateurs et citoyens vis-à-vis de l'élevage. Une étude quantitative auprès de 2 000 personnes en France. *Journées de la Recherche Porcine*. 2017. N° 49, pp. 295-300.

DEMARS, Julie, LABRUNE, Yann, IANNUCELLI, Nathalie, DESHAYES, Alice, LEROUX, Sophie, GILBERT, Hélène, AYMARD, Patrick, BENITEZ, Florence et RIQUET, Juliette, 2022. A genome-wide epistatic network underlies the molecular architecture of continuous color variation of body extremities. *Genomics*. mai 2022. Vol. 114, n° 3, pp. 110361. DOI 10.1016/j.ygeno.2022.110361.

DEPOUDENT, Caroline, PONET, Philippe et BERNARD, Marie-Lou, 2023. Impact des conditions de travail sur l'attractivité du métier de salarié en élevage porcin. *Journées de la Recherche Porcine*. 2023. N° 55, pp. 7-12.

D'ERASMO, Laura, MINICOCCI, Ilenia, DI COSTANZO, Alessia, PIGNA, Giovanni, COMMODARI, Daniela, CECI, Fabrizio, MONTALI, Anna, BRANCATO, Francesca, STANCA, Ilaria, NICOLUCCI, Antonio, ASCIONE, Andrea, GALEA, Nicola, CARBONE, Iacopo, FRANCONI, Marco, MARANGHI, Marianna et ARCA, Marcello, 2021. Clinical Implications of Monogenic Versus Polygenic Hypercholesterolemia: Long-Term Response to Treatment, Coronary Atherosclerosis Burden, and Cardiovascular Events. *Journal of the American Heart Association*. 4 mai 2021. Vol. 10, n° 9, pp. e018932. DOI 10.1161/JAHA.120.018932.

DOUBLET, Anna-Charlotte, CROISEAU, Pascal, FRITZ, Sébastien, MICHENET, Alexis, HOZÉ, Chris, DANCHIN-BURGE, Coralie, LALOË, Denis et RESTOUX, Gwendal, 2019. The impact of genomic selection on genetic diversity and genetic gain in three French dairy cattle breeds. *Genetics Selection Evolution*. 23 septembre 2019. Vol. 51, n° 1, pp. 52. DOI 10.1186/s12711-019-0495-1.

DUCOS, Alain, DOUHARD, Frédéric, SAVIETTO, Davi, SAUTIER, Marion, FILLON, Valérie, GUNIA, Mélanie, RUPP, Rachel, MORENO-ROMIEUX, Carole, MIGNON-GRASTEAU, Sandrine, GILBERT, Hélène et FORTUN-LAMOTHE, Laurence, 2021. Contributions de la génétique animale à la transition agroécologique des systèmes d'élevage. *INRAE Productions Animales*. 18 octobre 2021. Vol. 34, n° 2, pp. 79-96. DOI 10.20870/productions-animales.2021.34.2.4773.

DUCROCQ, Vincent, 1992. Les bases de la génétique quantitative : Du modèle génétique au modèle statistique. *INRAE Productions Animales*. 30 décembre 1992. Vol. 5, n° HS, pp. 75-8. DOI 10.20870/productions-animales.1992.5.HS.4266.

EL KHATTABI, Laïla, 2015. *Identification et caractérisation de nouveaux gènes soumis à l'empreinte parentale dans le placenta humain*. [en ligne]. Thèse de doctorat. Sorbonne Paris Cité. [Consulté le 25 avril 2023]. Disponible à l'adresse: <https://www.theses.fr/2015USPCB227>

ELLEGREN, H., JOHANSSON, M., CHOWDHARY, B. P., MARKLUND, S., RUYTER, D., MARKLUND, L., BRÄUNER-NIELSEN, P., EDFORS-LILJA, I., GUSTAVSSON, I. et JUNEJA, R. K., 1993. Assignment of 20 microsatellite markers to the porcine linkage map. *Genomics*. mai 1993. Vol. 16, n° 2, pp. 431-439. DOI 10.1006/geno.1993.1207.

ELSEN, J. M., 2011. Sélection sur gènes et sélection assistée par marqueurs. *INRAE Productions Animales*. 8 septembre 2011. Vol. 24, n° 4, pp. 357-362. DOI 10.20870/productions-animales.2011.24.4.3267.

ENGEL, Nora, RAVAL, Anjali K., THORVALDSEN, Joanne L. et BARTOLOMEI, S.

Marisa, 2008. Three-dimensional conformation at the H19/Igf2 locus supports a model of enhancer tracking. *Human Molecular Genetics*. 1 octobre 2008. Vol. 17, n° 19, pp. 3021-3029. DOI 10.1093/hmg/ddn200.

ESTELLÉ, Jordi, 2019. Benefits from the joint analysis of host genomes and metagenomes: Select the holobiont. *Journal of Animal Breeding and Genetics = Zeitschrift Fur Tierzucht Und Zuchtungsbiologie*. mars 2019. Vol. 136, n° 2, pp. 75-76. DOI 10.1111/jbg.12383.

FALCONER, D. S. et MACKAY, T.F.C, 1996. *Introduction to Quantitative Genetics*. . 4. Addison Wesley Longman.

FERGUSON-SMITH, Anne C. et BOURC'HIS, Deborah, 2018. The discovery and importance of genomic imprinting. *eLife*. 22 octobre 2018. Vol. 7, pp. e42368. DOI 10.7554/eLife.42368.

FERNÁNDEZ, A. I., ALVES, E., FERNÁNDEZ, A., DE PEDRO, E., LÓPEZ-GARCÍA, M. A., OVILO, C., RODRÍGUEZ, M. C. et SILIÓ, L., 2008. Mitochondrial genome polymorphisms associated with longissimus muscle composition in Iberian pigs. *Journal of Animal Science*. juin 2008. Vol. 86, n° 6, pp. 1283-1290. DOI 10.2527/jas.2007-0568.

FITZ-JAMES, Maximilian H. et CAVALLI, Giacomo, 2022. Molecular mechanisms of transgenerational epigenetic inheritance. *Nature Reviews Genetics*. juin 2022. Vol. 23, n° 6, pp. 325-341. DOI 10.1038/s41576-021-00438-5.

FORNERIS, Natalia S., VITEZICA, Zulma G., LEGARRA, Andres et PÉREZ-ENCISO, Miguel, 2017. Influence of epistasis on response to genomic selection using complete sequence data. *Genetics, Selection, Evolution : GSE*. 25 août 2017. Vol. 49, pp. 66. DOI 10.1186/s12711-017-0340-3.

FRAGA, Mario F., BALLESTAR, Esteban, PAZ, Maria F., ROPER, Santiago, SETIEN, Fernando, BALLESTAR, Maria L., HEINE-SUÑER, Damia, CIGUDOSA, Juan C., URIOSTE, Miguel, BENITEZ, Javier, BOIX-CHORNET, Manuel, SANCHEZ-AGUILERA, Abel, LING, Charlotte, CARLSSON, Emma, POULSEN, Pernille, VAAG, Allan, STEPHAN, Zarko, SPECTOR, Tim D., WU, Yue-Zhong, PLASS, Christoph et ESTELLER, Manel, 2005. Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences of the United States of America*. 26 juillet 2005. Vol. 102, n° 30, pp. 10604-10609. DOI 10.1073/pnas.0500398102.

FRÉSARD, Laure, LEROUX, Sophie, SERVIN, Bertrand, GOURICHON, David, DEHAIS, Patrice, CRISTOBAL, Magali San, MARSAUD, Nathalie, VIGNOLES, Florence, BED'HOM, Bertrand, COVILLE, Jean-Luc, HORMOZDIARI, Farhad, BEAUMONT, Catherine, ZERJAL, Tatiana, VIGNAL, Alain, MORISSON, Mireille, LAGARRIGUE, Sandrine et PITEL, Frédérique, 2014. Transcriptome-wide investigation of genomic imprinting in chicken. *Nucleic Acids Research*. avril 2014. Vol. 42, n° 6, pp. 3768-3782. DOI 10.1093/nar/gkt1390.

FUJII, J., OTSU, K., ZORZATO, F., DE LEON, S., KHANNA, V. K., WEILER, J. E., O'BRIEN, P. J. et MACLENNAN, D. H., 1991. Identification of a mutation in porcine ryanodine receptor associated with malignant hyperthermia. *Science (New York, N.Y.)*. 26 juillet 1991. Vol. 253, n° 5018, pp. 448-451. DOI 10.1126/science.1862346.

GAHUROVA, Lenka, TOMIZAWA, Shin-ichi, SMALLWOOD, Sébastien A., STEWART-MORGAN, Kathleen R., SAADEH, Heba, KIM, Jeesun, ANDREWS, Simon R., CHEN, Taiping et KELSEY, Gavin, 2017. Transcription and chromatin determinants of de novo DNA methylation timing in oocytes. *Epigenetics & Chromatin*. 12 mai 2017. Vol. 10, pp. 25. DOI 10.1186/s13072-017-0133-5.

GARG, Shilpa, MARTIN, Marcel et MARSCHALL, Tobias, 2016. Read-based phasing of

related individuals. *Bioinformatics (Oxford, England)*. 15 juin 2016. Vol. 32, n° 12, pp. i234-i242. DOI 10.1093/bioinformatics/btw276.

GAYON, Jean, 2016. From Mendel to epigenetics: History of genetics. *Comptes Rendus Biologies*. juillet 2016. Vol. 339, n° 7-8, pp. 225-230. DOI 10.1016/j.crvi.2016.05.009.

GDS MAYENNE, 2019. Les races porcines élevées en Mayenne. | M-Elevage. [en ligne]. 9 mai 2019. [Consulté le 17 mai 2023]. Disponible à l'adresse: [https://www.m-elevage.fr/porcins/races/Obtenir des informations sur les races porcines élevées en Mayenne \(Caractéristiques, histoire, chiffres...\)](https://www.m-elevage.fr/porcins/races/Obtenir%20des%20informations%20sur%20les%20races%20porcines%20%C3%A9lev%C3%A9es%20en%20Mayenne%20(Caract%C3%A9ristiques,%20histoire,%20chiffres...)).

GENEIMPRINT, 2023. Imprinted gene database. [en ligne]. 2023. Disponible à l'adresse: <https://www.geneimprint.com/site/genes-by-species>

GENOTOUL, 2017. Séquençage Haut-Débit sur Roche 454 GS FLX. [en ligne]. Poster. 2017. Disponible à l'adresse: https://get.genotoul.fr/wp-content/uploads/2017/06/101210_genotoul_poster_454.pdf

GHEZZI, Daniele, MARELLI, Cecilia, ACHILLI, Alessandro, GOLDWURM, Stefano, PEZZOLI, Gianni, BARONE, Paolo, PELLECCIA, Maria Teresa, STANZIONE, Paolo, BRUSA, Livia, BENTIVOGLIO, Anna Rita, BONUCCELLI, Ubaldo, PETROZZI, Lucia, ABBRUZZESE, Giovanni, MARCHESE, Roberta, CORTELLI, Pietro, GRIMALDI, Daniela, MARTINELLI, Paolo, FERRARESE, Carlo, GARAVAGLIA, Barbara, SANGIORGI, Simonetta, CARELLI, Valerio, TORRONI, Antonio, ALBANESE, Alberto et ZEVIANI, Massimo, 2005. Mitochondrial DNA haplogroup K is associated with a lower risk of Parkinson's disease in Italians. *European journal of human genetics: EJHG*. juin 2005. Vol. 13, n° 6, pp. 748-752. DOI 10.1038/sj.ejhg.5201425.

GIBNEY, E. R. et NOLAN, C. M., 2010. Epigenetics and gene expression. *Heredity*. juillet 2010. Vol. 105, n° 1, pp. 4-13. DOI 10.1038/hdy.2010.54.

GIRARDOT, Michael, CAVAILLÉ, Jérôme et FEIL, Robert, 2012. Small regulatory RNAs controlled by genomic imprinting and their contribution to human disease. *Epigenetics*. 17 décembre 2012. Vol. 7, n° 12, pp. 1341-1348. DOI 10.4161/epi.22884.

GLASER, Juliane, IRANZO, Julian, BORENSZTEIN, Maud, MARINUCCI, Mattia, GUALTIERI, Angelica, JOUHANNEAU, Colin, TEISSANDIER, Aurélie, GASTON-MASSUET, Carles et BOURC'HIS, Deborah, 2022. The imprinted Zdbf2 gene finely tunes control of feeding and growth in neonates. BRICKMAN, Joshua M, ISALES, Carlos, WARD, Andrew, CHAMBERLAIN, Stormy et KELSEY, Gavin (éd.), *eLife*. 20 janvier 2022. Vol. 11, pp. e65641. DOI 10.7554/eLife.65641.

GODDARD, Michael E. et WHITELAW, Emma, 2014. The use of epigenetic phenomena for the improvement of sheep and cattle. *Frontiers in Genetics*. [en ligne]. 2014. Vol. 5. [Consulté le 21 mai 2023]. Disponible à l'adresse: <https://www.frontiersin.org/articles/10.3389/fgene.2014.00247>

GOODWIN, Sara, MCPHERSON, John D. et MCCOMBIE, W. Richard, 2016. Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews. Genetics*. 17 mai 2016. Vol. 17, n° 6, pp. 333-351. DOI 10.1038/nrg.2016.49.

GREGG, Christopher, ZHANG, Jiangwen, BUTLER, James E., HAIG, David et DULAC, Catherine, 2010. Sex-specific parent-of-origin allelic expression in the mouse brain. *Science (New York, N.Y.)*. 6 août 2010. Vol. 329, n° 5992, pp. 682-685. DOI 10.1126/science.1190831.

GROENEN, Martien A. M., ARCHIBALD, Alan L., UENISHI, Hirohide, TUGGLE,

Christopher K., TAKEUCHI, Yasuhiro, ROTHSCHILD, Max F., ROGEL-GAILLARD, Claire, PARK, Chankyu, MILAN, Denis, MEGENS, Hendrik-Jan, LI, Shengting, LARKIN, Denis M., KIM, Heebal, FRANTZ, Laurent A. F., CACCAMO, Mario, AHN, Hyeonju, AKEN, Bronwen L., ANSELMO, Anna, ANTHON, Christian, AUVIL, Loretta, BADAoui, Bouabid, BEATTIE, Craig W., BENDIXEN, Christian, BERMAN, Daniel, BLECHA, Frank, BLOMBERG, Jonas, BOLUND, Lars, BOSSE, Mirte, BOTTI, Sara, BUJIE, Zhan, BYSTROM, Megan, CAPITANU, Boris, CARVALHO-SILVA, Denise, CHARDON, Patrick, CHEN, Celine, CHENG, Ryan, CHOI, Sang-Haeng, CHOW, William, CLARK, Richard C., CLEE, Christopher, CROOIJMANS, Richard P. M. A., DAWSON, Harry D., DEHAIS, Patrice, DE SAPIO, Fioravante, DIBBITS, Bert, DROU, Nizar, DU, Zhi-Qiang, EVERSOLE, Kellye, FADISTA, João, FAIRLEY, Susan, FARAUT, Thomas, FAULKNER, Geoffrey J., FOWLER, Katie E., FREDHOLM, Merete, FRITZ, Eric, GILBERT, James G. R., GIUFFRA, Elisabetta, GORODKIN, Jan, GRIFFIN, Darren K., HARROW, Jennifer L., HAYWARD, Alexander, HOWE, Kerstin, HU, Zhi-Liang, HUMPHRAY, Sean J., HUNT, Toby, HORNSHØJ, Henrik, JEON, Jin-Tae, JERN, Patric, JONES, Matthew, JURKA, Jerzy, KANAMORI, Hiroyuki, KAPETANOVIC, Ronan, KIM, Jaebum, KIM, Jae-Hwan, KIM, Kyu-Won, KIM, Tae-Hun, LARSON, Greger, LEE, Kyooyeol, LEE, Kyung-Tai, LEGGETT, Richard, LEWIN, Harris A., LI, Yingrui, LIU, Wansheng, LOVELAND, Jane E., LU, Yao, LUNNEY, Joan K., MA, Jian, MADSEN, Ole, MANN, Katherine, MATTHEWS, Lucy, MCLAREN, Stuart, MOROZUMI, Takeya, MURTAUGH, Michael P., NARAYAN, Jitendra, NGUYEN, Dinh Truong, NI, Peixiang, OH, Song-Jung, ONTERU, Suneel, PANITZ, Frank, PARK, Eung-Woo, PARK, Hong-Seog, PASCAL, Geraldine, PAUDEL, Yogesh, PEREZ-ENCISO, Miguel, RAMIREZ-GONZALEZ, Ricardo, REECY, James M., RODRIGUEZ-ZAS, Sandra, ROHRER, Gary A., RUND, Laretta, SANG, Yongming, SCHACHTSCHNEIDER, Kyle, SCHRAIBER, Joshua G., SCHWARTZ, John, SCOBIE, Linda, SCOTT, Carol, SEARLE, Stephen, SERVIN, Bertrand, SOUTHEY, Bruce R., SPERBER, Goran, STADLER, Peter, SWEEDLER, Jonathan V., TAFER, Hakim, THOMSEN, Bo, WALI, Rashmi, WANG, Jian, WANG, Jun, WHITE, Simon, XU, Xun, YERLE, Martine, ZHANG, Guojie, ZHANG, Jianguo, ZHANG, Jie, ZHAO, Shuhong, ROGERS, Jane, CHURCHER, Carol et SCHOOK, Lawrence B., 2012. Analyses of pig genomes provide insight into porcine demography and evolution. *Nature*. 15 novembre 2012. Vol. 491, n° 7424, pp. 393-398. DOI 10.1038/nature11622. F

GU, Tian-Peng, GUO, Fan, YANG, Hui, WU, Hai-Ping, XU, Gui-Fang, LIU, Wei, XIE, Zhi-Guo, SHI, Linyu, HE, Xinyi, JIN, Seung-gi, IQBAL, Khursheed, SHI, Yujiang, GENO, DENG, Zixin, SZABÓ, Piroška E., PFEIFER, Gerd P., LI, Jinsong et XU, Guo-Liang, 2011. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature*. 4 septembre 2011. Vol. 477, n° 7366, pp. 606-610. DOI 10.1038/nature10443.

GUO, Jing, HE, Hongjuan, LIU, Hui, LIU, Qi, ZHANG, Lili, LIU, Boqi, SUGIMOTO, Kenkichi et WU, Qiong, 2016. Aquaporin-1, a New Maternally Expressed Gene, Regulates Placental Development in the Mouse. *Biology of Reproduction*. 1 août 2016. Vol. 95, n° 2, pp. 40, 1-9. DOI 10.1095/biolreprod.116.138636.

HACKETT, Jamie A. et SURANI, M. Azim, 2013. DNA methylation dynamics during the mammalian life cycle. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 5 janvier 2013. Vol. 368, n° 1609, pp. 20110328. DOI 10.1098/rstb.2011.0328.

HACKETT, Jamie A., ZYLICZ, Jan J. et SURANI, M. Azim, 2012. Parallel mechanisms of epigenetic reprogramming in the germline. *Trends in Genetics: TIG*. avril 2012. Vol. 28, n° 4, pp. 164-174. DOI 10.1016/j.tig.2012.01.005.

HAGER, Reinmar, CHEVERUD, James M, LEAMY, Larry J et WOLF, Jason B, 2008. Sex

dependent imprinting effects on complex traits in mice. *BMC Evolutionary Biology*. 31 octobre 2008. Vol. 8, pp. 303. DOI 10.1186/1471-2148-8-303.

HAIG, D, 1997. Parental antagonism, relatedness asymmetries, and genomic imprinting. *Proceedings of the Royal Society B: Biological Sciences*. 22 novembre 1997. Vol. 264, n° 1388, pp. 1657-1662.

HAIG, D., 2014. Coadaptation and conflict, misconception and muddle, in the evolution of genomic imprinting. *Heredity*. août 2014. Vol. 113, n° 2, pp. 96-103. DOI 10.1038/hdy.2013.97.

HAJKOVA, Petra, ERHARDT, Sylvia, LANE, Natasha, HAAF, Thomas, EL-MAARRI, Osman, REIK, Wolf, WALTER, Jörn et SURANI, M.Azim, 2002. Epigenetic reprogramming in mouse primordial germ cells. *Mechanisms of Development*. septembre 2002. Vol. 117, n° 1-2, pp. 15-23. DOI 10.1016/S0925-4773(02)00181-8.

HALABIAN, Raheleh, VALIZADEH ARSHAD, AHMADI, Ali, SAEEDI, Pardis, AZIMZADEH JAMALKANDI, Sadegh et ALIVAND, Mohammad Reza, 2021. Laboratory methods to decipher epigenetic signatures: a comparative review. *Cellular & Molecular Biology Letters*. 11 novembre 2021. Vol. 26, n° 1, pp. 46. DOI 10.1186/s11658-021-00290-9.

HEBEL, Pascale, 2013. Evolution, facteurs de changement et perspectives de la consommation de viande de porc en France. *Journées de la Recherche Porcine*. 2013. N° 45, pp. 25-30.

HEBERT, Hélène, LURETTE, Amandine, FOURICHON, Christine, SEEGERS, Henri et BELLOC, Catherine, 2007. Modalités de conduite en bandes en élevage porcin : effet sur les contacts entre animaux. *Journées de la Recherche Porcine*. 2007. N° 39, pp. 345-350.

HENDERSON, C R, 1963. Selection Index and Expected Genetic Advance. In : *Statistical genetics and plant breeding*. pp. 141-163.

HENDERSON, C. R., 1975. Best linear unbiased estimation and prediction under a selection model. *Biometrics*. juin 1975. Vol. 31, n° 2, pp. 423-447.

HEYN, Holger, CARMONA, F. Javier, GOMEZ, Antonio, FERREIRA, Humberto J., BELL, Jordana T., SAYOLS, Sergi, WARD, Kirsten, STEFANSSON, Olafur A., MORAN, Sebastian, SANDOVAL, Juan, EYFJORD, Jorunn E., SPECTOR, Tim D. et ESTELLER, Manel, 2013. DNA methylation profiling in breast cancer discordant identical twins identifies DOK7 as novel epigenetic biomarker. *Carcinogenesis*. janvier 2013. Vol. 34, n° 1, pp. 102-108. DOI 10.1093/carcin/bgs321.

HIBOOST, 2018. Les OSP françaises réaffirment leur volonté d'unir leurs forces dans un nouveau travail collectif ... *IFIP*. [en ligne]. 8 janvier 2018. [Consulté le 19 mars 2023]. Disponible à l'adresse: <https://ifip.asso.fr/actualites/les-osp-francaises-reaffirment-leur-volonte-dunir-leurs-forces-dans-un-nouveau-travail-collectif/Bioporc>

HIGGS, M. J., HILL, M. J., JOHN, R. M. et ISLES, A. R., 2022. Systematic investigation of imprinted gene expression and enrichment in the mouse brain explored at single-cell resolution. *BMC Genomics*. 17 novembre 2022. Vol. 23, n° 1, pp. 754. DOI 10.1186/s12864-022-08986-8.

HILL, William G., GODDARD, Michael E. et VISSCHER, Peter M., 2008. Data and Theory Point to Mainly Additive Genetic Variance for Complex Traits. *PLoS Genetics*. 29 février 2008. Vol. 4, n° 2, pp. e1000008. DOI 10.1371/journal.pgen.1000008.

HIRASAWA, Ryutaro, CHIBA, Hatsune, KANEDA, Masahiro, TAJIMA, Shoji, LI, En,

JAENISCH, Rudolf et SASAKI, Hiroyuki, 2008. Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. *Genes & Development*. 15 juin 2008. Vol. 22, n° 12, pp. 1607-1616. DOI 10.1101/gad.1667008.

HOLM, B., BAKKEN, M., KLEMETSDAL, G. et VANGEN, O., 2004. Genetic correlations between reproduction and production traits in swine. *Journal of Animal Science*. décembre 2004. Vol. 82, n° 12, pp. 3458-3464. DOI 10.2527/2004.82123458x.

HORE, Timothy A., RAPKINS, Robert W. et GRAVES, Jennifer A. Marshall, 2007. Construction and evolution of imprinted loci in mammals. *Trends in genetics: TIG*. septembre 2007. Vol. 23, n° 9, pp. 440-448. DOI 10.1016/j.tig.2007.07.003.

HUBERT, Jean-Noël, IANNUCELLI, N., CABAU, Cédric, JACOMET, Eva, BILLON, Yvon, SERRE, Rémy-Félix, VANDECASTEELE, Céline, DONNADIEU, Cécile et DEMARS, Julie, 2023. *A methyl-seq tool to capture genomic imprinted loci* [en ligne]. 21 mars 2023. bioRxiv. [Consulté le 12 mai 2023].

HUBERT, Jean-Noël, IANNUCELLI, Nathalie, CABAU, Cédric, SERRE, Rémy-Félix, VANDECASTEELE, Céline, JACOMET, Eva, BRENOT, Sylvie, BILLON, Yvon, HEBRARD, William, DONNADIEU, Cécile et DEMARS, Julie, 2022a. Vers une carte de l'imprintome porcin: comparaison de 2 méthodes de méthyl-seq. . 2022.

HUBERT, Jean-Noël, IANNUCELLI, Nathalie, CABAU, Cédric, SERRE, Rémy-Félix, VANDECASTEELE, Céline, JACOMET, Eva, BRENOT, Sylvie, BILLON, Yvon, HEBRARD, William, DONNADIEU, Cécile et DEMARS, Julie, 2022b. Targeting the pig imprintome : a comparative approach using two Methyl-seq technologies. . 2022.

HUBERT, Jean-Noël, PERRET, Mathilde, IANNUCELLI, Nathalie, JACOMET, Eva, CABAU, Cédric et DEMARS, Julie, 2023. A bioinformatics pipeline to detect imprinting from methyl-seq trio data. 2023.

HUFFMAN, Michael A. et HIRATA, Satoshi, 2004. An experimental study of leaf swallowing in captive chimpanzees: insights into the origin of a self-medicative behavior and the role of social learning. *Primates; Journal of Primatology*. avril 2004. Vol. 45, n° 2, pp. 113-118. DOI 10.1007/s10329-003-0065-5.

HYDE, Lara, FRISO, Simonetta et CHOI, Sang-Woon, 2020. Introduction to Epigenetics. In : . pp. 129-139. ISBN 978-0-12-804572-5.

IBEAGHA-AWEMU, Eveline M. et ZHAO, Xin, 2015. Epigenetic marks: regulators of livestock phenotypes and conceivable sources of missing variation in livestock improvement programs. *Frontiers in Genetics*. 2015. Vol. 6, pp. 302. DOI 10.3389/fgene.2015.00302.

ILLUMINA, [sans date]. Trusted bead-based technology. [en ligne]. Disponible à l'adresse: <https://www.illumina.com/science/technology/microarray.html>

ILLUMINA, 2021. *Système MiSeq* [en ligne]. 2021. Disponible à l'adresse: <https://www.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/miseq-data-sheet-m-gl-00006-translations/miseq-system-specification-sheet-m-gl-00006-fra-view.pdf>

ILLUMINA, 2023. Illumina Sequencing Platforms. [en ligne]. 2023. Disponible à l'adresse: <https://www.illumina.com/systems/sequencing-platforms.html>

INOUE, Azusa, JIANG, Lan, LU, Falong, SUZUKI, Tsukasa et ZHANG, Yi, 2017. Maternal

H3K27me3 controls DNA methylation-independent imprinting. *Nature*. juillet 2017. Vol. 547, n° 7664, pp. 419-424. DOI 10.1038/nature23262.

INRAE, 2020. Un holobionte qu'est-ce-que c'est ? Et pourquoi INRAE s'y intéresse-t-il ? *INRAE Institutionnel*. [en ligne]. 2020. [Consulté le 27 avril 2023]. Disponible à l'adresse: <https://www.inrae.fr/actualites/holobionte>

JACOB, Erin et HEGELE, Robert A., 2022. Monogenic Versus Polygenic Forms of Hypercholesterolemia and Cardiovascular Risk: Are There Any Differences? *Current Atherosclerosis Reports*. juin 2022. Vol. 24, n° 6, pp. 419-426. DOI 10.1007/s11883-022-01018-6.

JAIN, Miten, OLSEN, Hugh E., PATEN, Benedict et AKESON, Mark, 2016. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biology*. 25 novembre 2016. Vol. 17, n° 1, pp. 239. DOI 10.1186/s13059-016-1103-0.

JAMMES, H. et RENARD, J.P., 2010. Epigénétique et construction du phénotype, un enjeu pour les productions animales ? *INRAE Productions Animales*. 8 février 2010. Vol. 23, n° 1, pp. 23-42. DOI 10.20870/productions-animales.2010.23.1.3283.

JEON, J. T., CARLBORG, O., TÖRNSTEN, A., GIUFFRA, E., AMARGER, V., CHARDON, P., ANDERSSON-EKLUND, L., ANDERSSON, K., HANSSON, I., LUNDSTRÖM, K. et ANDERSSON, L., 1999. A paternally expressed QTL affecting skeletal and cardiac muscle mass in pigs maps to the IGF2 locus. *Nature Genetics*. février 1999. Vol. 21, n° 2, pp. 157-158. DOI 10.1038/5938.

JIMA, Dereje D., SKAAR, David A., PLANCHART, Antonio, MOTSINGER-REIF, Alison, CEVIK, Sebnem E., PARK, Sarah S., COWLEY, Michael, WRIGHT, Fred, HOUSE, John, LIU, Andy, JIRTLE, Randy L. et HOYO, Cathrine, 2022. Genomic map of candidate human imprint control regions: the imprintome. *Epigenetics*. 9 décembre 2022. Vol. 17, n° 13, pp. 1920-1943. DOI 10.1080/15592294.2022.2091815.

JOHN, Rosalind M., 2017. Imprinted genes and the regulation of placental endocrine function: Pregnancy and beyond. *Placenta*. 1 août 2017. Vol. 56, pp. 86-90. DOI 10.1016/j.placenta.2017.01.099.

JORDAN, Bertrand, 2017. Séquençage d'ADN : l'offensive des nanopores: Chroniques génomiques. *médecine/sciences*. août 2017. Vol. 33, n° 8-9, pp. 801-804. DOI 10.1051/medsci/20173308028.

KALISH, Jennifer M., JIANG, Connie et BARTOLOMEI, Marisa S., 2014. Epigenetics and imprinting in human disease. *International Journal of Developmental Biology*. 10 juillet 2014. Vol. 58, n° 2-3-4, pp. 291-298. DOI 10.1387/ijdb.140077mb.

KAMINSKY, Zachary A., TANG, Thomas, WANG, Sun-Chong, PTAK, Carolyn, OH, Gabriel H. T., WONG, Albert H. C., FELDCAMP, Laura A., VIRTANEN, Carl, HALFVARSON, Jonas, TYSK, Curt, MCRAE, Allan F., VISSCHER, Peter M., MONTGOMERY, Grant W., GOTTESMAN, Irving I., MARTIN, Nicholas G. et PETRONIS, Art, 2009. DNA methylation profiles in monozygotic and dizygotic twins. *Nature Genetics*. février 2009. Vol. 41, n° 2, pp. 240-245. DOI 10.1038/ng.286.

KAPPELER, Peter (éd.), 2010. *Animal Behaviour: Evolution and Mechanisms*. [en ligne]. Berlin, Heidelberg : Springer. [Consulté le 4 mai 2023]. ISBN 978-3-642-02623-2.

KASPRZAK, Aldona et ADAMEK, Agnieszka, 2019. Insulin-Like Growth Factor 2 (IGF2) Signaling in Colorectal Cancer—From Basic Research to Potential Clinical Applications.

International Journal of Molecular Sciences. 3 octobre 2019. Vol. 20, n° 19, pp. 4915. DOI 10.3390/ijms20194915.

KAWAHARA, Manabu, WU, Qiong, TAKAHASHI, Nozomi, MORITA, Shinnosuke, YAMADA, Kaori, ITO, Mitsuteru, FERGUSON-SMITH, Anne C. et KONO, Tomohiro, 2007. High-frequency generation of viable mice from engineered bi-maternal embryos. *Nature Biotechnology*. septembre 2007. Vol. 25, n° 9, pp. 1045-1050. DOI 10.1038/nbt1331.

KELLY, Derek E., RAMDAS, Shweta, MA, Rong, RAWLINGS-GOSS, Renata A., GRANT, Gregory R., RANCIARO, Alessia, HIRBO, Jibril B., BEGGS, William, YEAGER, Meredith, CHANOCK, Stephen, NYAMBO, Thomas B., OMAR, Sabah A., WOLDEMESKEL, Dawit, BELAY, Gurja, LI, Hongzhe, BROWN, Christopher D. et TISHKOFF, Sarah A., 2023. The genetic and evolutionary basis of gene expression variation in East Africans. *Genome Biology*. 24 février 2023. Vol. 24, n° 1, pp. 35. DOI 10.1186/s13059-023-02874-4.

KEVERNE, E. B., 2013. Importance of the matriline for genomic imprinting, brain development and behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 5 janvier 2013. Vol. 368, n° 1609, pp. 20110327. DOI 10.1098/rstb.2011.0327.

KHATIB, Hasan, 2021. Transgenerational epigenetic inheritance in farm animals: How substantial is the evidence? *Livestock Science*. 1 août 2021. Vol. 250, pp. 104557. DOI 10.1016/j.livsci.2021.104557.

KINGDOM, Rebecca et WRIGHT, Caroline F., 2022. Incomplete Penetrance and Variable Expressivity: From Clinical Studies to Population Cohorts. *Frontiers in Genetics*. 25 juillet 2022. Vol. 13, pp. 920390. DOI 10.3389/fgene.2022.920390.

KOBAYASHI, Hisato, 2021. Canonical and Non-canonical Genomic Imprinting in Rodents. *Frontiers in Cell and Developmental Biology*. [en ligne]. 2021. Vol. 9. [Consulté le 9 mai 2023]. Disponible à l'adresse: <https://www.frontiersin.org/articles/10.3389/fcell.2021.713878>

KOERNER, Martha V., PAULER, Florian M., HUANG, Ru et BARLOW, Denise P., 2009. The function of non-coding RNAs in genomic imprinting. *Development (Cambridge, England)*. juin 2009. Vol. 136, n° 11, pp. 1771-1783. DOI 10.1242/dev.030403.

KOHLI, Rahul M. et ZHANG, Yi, 2013. TET enzymes, TDG and the dynamics of DNA demethylation. *Nature*. 24 octobre 2013. Vol. 502, n° 7472, pp. 472-479. DOI 10.1038/nature12750.

KRESSMANN, Gil, 2022. Un marché bio en crise profonde. *Paysans & société*. 2022. Vol. 393, n° 3, pp. 34-40. DOI 10.3917/pes.393.0034.

KRUEGER, Felix et ANDREWS, Simon R., 2011. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics (Oxford, England)*. 1 juin 2011. Vol. 27, n° 11, pp. 1571-1572. DOI 10.1093/bioinformatics/btr167.

LAFRAMBOISE, Thomas, 2009. Single nucleotide polymorphism arrays: a decade of biological, computational and technological advances. *Nucleic Acids Research*. juillet 2009. Vol. 37, n° 13, pp. 4181-4193. DOI 10.1093/nar/gkp552. A

LANDE, R. et THOMPSON, R., 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics*. mars 1990. Vol. 124, n° 3, pp. 743-756. DOI 10.1093/genetics/124.3.743.

LANDER, E. S., LINTON, L. M., BIRREN, B., NUSBAUM, C., ZODY, M. C., BALDWIN,

J., DEVON, K., DEWAR, K., DOYLE, M., FITZHUGH, W., FUNKE, R., GAGE, D., HARRIS, K., HEAFORD, A., HOWLAND, J., KANN, L., LEHOCZKY, J., LEVINE, R., MCEWAN, P., MCKERNAN, K., MELDRIM, J., MESIROV, J. P., MIRANDA, C., MORRIS, W., NAYLOR, J., RAYMOND, C., ROSETTI, M., SANTOS, R., SHERIDAN, A., SOUGNEZ, C., STANGE-THOMANN, Y., STOJANOVIC, N., SUBRAMANIAN, A., WYMAN, D., ROGERS, J., SULSTON, J., AINSCOUGH, R., BECK, S., BENTLEY, D., BURTON, J., CLEE, C., CARTER, N., COULSON, A., DEADMAN, R., DELOUKAS, P., DUNHAM, A., DUNHAM, I., DURBIN, R., FRENCH, L., GRAFHAM, D., GREGORY, S., HUBBARD, T., HUMPHRAY, S., HUNT, A., JONES, M., LLOYD, C., MCMURRAY, A., MATTHEWS, L., MERCER, S., MILNE, S., MULLIKIN, J. C., MUNGALL, A., PLUMB, R., ROSS, M., SHOWNKEEN, R., SIMS, S., WATERSTON, R. H., WILSON, R. K., HILLIER, L. W., MCPHERSON, J. D., MARRA, M. A., MARDIS, E. R., FULTON, L. A., CHINWALLA, A. T., PEPIN, K. H., GISH, W. R., CHISSOE, S. L., WENDL, M. C., DELEHAUNTY, K. D., MINER, T. L., DELEHAUNTY, A., KRAMER, J. B., COOK, L. L., FULTON, R. S., JOHNSON, D. L., MINX, P. J., CLIFTON, S. W., HAWKINS, T., BRANSCOMB, E., PREDKI, P., RICHARDSON, P., WENNING, S., SLEZAK, T., DOGGETT, N., CHENG, J. F., OLSEN, A., LUCAS, S., ELKIN, C., UBERBACHER, E., FRAZIER, M., GIBBS, R. A., MUZNY, D. M., SCHERER, S. E., BOUCK, J. B., SODERGRÉN, E. J., WORLEY, K. C., RIVES, C. M., GORRELL, J. H., METZKER, M. L., NAYLOR, S. L., KUCHERLAPATI, R. S., NELSON, D. L., WEINSTOCK, G. M., SAKAKI, Y., FUJIYAMA, A., HATTORI, M., YADA, T., TOYODA, A., ITOH, T., KAWAGOE, C., WATANABE, H., TOTOKI, Y., TAYLOR, T., WEISSENBACH, J., HEILIG, R., SAURIN, W., ARTIGUENAVE, F., BROTTIER, P., BRULS, T., PELLETIER, E., ROBERT, C., WINCKER, P., SMITH, D. R., DOUCETTE-STAMM, L., RUBENFIELD, M., WEINSTOCK, K., LEE, H. M., DUBOIS, J., ROSENTHAL, A., PLATZER, M., NYAKATURA, G., TAUDIEN, S., RUMP, A., YANG, H., YU, J., WANG, J., HUANG, G., GU, J., HOOD, L., ROWEN, L., MADAN, A., QIN, S., DAVIS, R. W., FEDERSPIEL, N. A., ABOLA, A. P., PROCTOR, M. J., MYERS, R. M., SCHMUTZ, J., DICKSON, M., GRIMWOOD, J., COX, D. R., OLSON, M. V., KAUL, R., RAYMOND, C., SHIMIZU, N., KAWASAKI, K., MINOSHIMA, S., EVANS, G. A., ATHANASIOU, M., SCHULTZ, R., ROE, B. A., CHEN, F., PAN, H., RAMSER, J., LEHRACH, H., REINHARDT, R., MCCOMBIE, W. R., DE LA BASTIDE, M., DEDHIA, N., BLÖCKER, H., HORNISCHER, K., NORDSIEK, G., AGARWALA, R., ARAVIND, L., BAILEY, J. A., BATEMAN, A., BATZOGLOU, S., BIRNEY, E., BORK, P., BROWN, D. G., BURGE, C. B., CERUTTI, L., CHEN, H. C., CHURCH, D., CLAMP, M., COPLEY, R. R., DOERKS, T., EDDY, S. R., EICHLER, E. E., FUREY, T. S., GALAGAN, J., GILBERT, J. G., HARMON, C., HAYASHIZAKI, Y., HAUSSLER, D., HERMJAKOB, H., HOKAMP, K., JANG, W., JOHNSON, L. S., JONES, T. A., KASIF, S., KASPRYZK, A., KENNEDY, S., KENT, W. J., KITTS, P., KOONIN, E. V., KORF, I., KULP, D., LANCET, D., LOWE, T. M., MCLYSAGHT, A., MIKKELSEN, T., MORAN, J. V., MULDER, N., POLLARA, V. J., PONTING, C. P., SCHULER, G., SCHULTZ, J., SLATER, G., SMIT, A. F., STUPKA, E., SZUSTAKOWKI, J., THIERRY-MIEG, D., THIERRY-MIEG, J., WAGNER, L., WALLIS, J., WHEELER, R., WILLIAMS, A., WOLF, Y. I., WOLFE, K. H., YANG, S. P., YEH, R. F., COLLINS, F., GUYER, M. S., PETERSON, J., FELSENFELD, A., WETTERSTRAND, K. A., PATRINOS, A., MORGAN, M. J., DE JONG, P., CATANESE, J. J., OSOEGAWA, K., SHIZUYA, H., CHOI, S., CHEN, Y. J., SZUSTAKOWKI, J., et INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM, 2001. Initial sequencing and analysis of the human genome. *Nature*. 15 février 2001. Vol. 409, n° 6822, pp. 860-921. DOI 10.1038/35057062.

LARZUL, Catherine, BOREY, Marion, BILLON, Yvon, ROSSIGNOL, Marie-Noelle, LEMONNIER, Gaëtan, ESTELLÉ, Jordi et GAILLARD, Claire Rogel, 2021. Response to

selection on fecal microbiota composition in Large White piglets. In : *ISAG*. [en ligne]. 26 juillet 2021. [Consulté le 27 avril 2023]. Disponible à l'adresse: <https://hal.science/hal-03464516>

LATOS, Paulina A., PAULER, Florian M., KOERNER, Martha V., ŞENERGIN, H. Başak, HUDSON, Quanah J., STOCSITS, Roman R., ALLHOFF, Wolfgang, STRICKER, Stefan H., KLEMENT, Ruth M., WARCZOK, Katarzyna E., AUMAYR, Karin, PASIERBEK, Pawel et BARLOW, Denise P., 2012. Airn Transcriptional Overlap, But Not Its lncRNA Products, Induces Imprinted Igf2r Silencing. *Science*. 2012. Vol. 338, n° 6113, pp. 1469-1472.

LAVER, T., HARRISON, J., O'NEILL, P. A., MOORE, K., FARBOS, A., PASZKIEWICZ, K. et STUDHOLME, D. J., 2015. Assessing the performance of the Oxford Nanopore Technologies MinION. *Biomolecular Detection and Quantification*. mars 2015. Vol. 3, pp. 1-8. DOI 10.1016/j.bdq.2015.02.001.

LE ROUX, Anne, 2002. *La réforme des verrats de centres d'insémination artificielle pour baisse de qualité de semence : approche anatomopathologique et histologique*. [en ligne]. other. [Consulté le 17 mai 2023]. Disponible à l'adresse: <https://oatao.univ-toulouse.fr/1052/>

LEGARRA, Andres, CHRISTENSEN, Ole F., AGUILAR, Ignacio et MISZTAL, Ignacy, 2014. Single Step, a general approach for genomic selection. *Livestock Science*. 1 août 2014. Vol. 166, pp. 54-65. DOI 10.1016/j.livsci.2014.04.029.

LEROUX, Sophie, GOURICHON, David, LETERRIER, Christine, LABRUNE, Yann, COUSTHAM, Vincent, RIVIÈRE, Sandrine, ZERJAL, Tatiana, COVILLE, Jean-Luc, MORISSON, Mireille, MINVIELLE, Francis et PITEL, Frédérique, 2017. Embryonic environment and transgenerational effects in quail. *Genetics, Selection, Evolution : GSE*. 26 janvier 2017. Vol. 49, pp. 14. DOI 10.1186/s12711-017-0292-7.

LEWIS, Annabelle, MITSUYA, Kohzoh, UMLAUF, David, SMITH, Paul, DEAN, Wendy, WALTER, Joern, HIGGINS, Michael, FEIL, Robert et REIK, Wolf, 2004. Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of DNA methylation. *Nature Genetics*. décembre 2004. Vol. 36, n° 12, pp. 1291-1295. DOI 10.1038/ng1468.

LI, E., BEARD, C. et JAENISCH, R., 1993. Role for DNA methylation in genomic imprinting. *Nature*. 25 novembre 1993. Vol. 366, n° 6453, pp. 362-365. DOI 10.1038/366362a0.

LI, Rui, MONTPETIT, Alexandre, ROUSSEAU, Marylène, WU, Si Yu Margaret, GREENWOOD, Celia M. T., SPECTOR, Timothy D., POLLAK, Michael, POLYCHRONAKOS, Constantin et RICHARDS, J. Brent, 2014. Somatic point mutations occurring early in development: a monozygotic twin study. *Journal of Medical Genetics*. 1 janvier 2014. Vol. 51, n° 1, pp. 28-34. DOI 10.1136/jmedgenet-2013-101712.

LI, Xiajun, ITO, Mitsuteru, ZHOU, Fen, YOUNGSON, Neil, ZUO, Xiaopan, LEDER, Philip et FERGUSON-SMITH, Anne C., 2008. A Maternal-Zygotic Effect Gene, Zfp57, Maintains Both Maternal and Paternal Imprints. *Developmental Cell*. 14 octobre 2008. Vol. 15, n° 4, pp. 547-557. DOI 10.1016/j.devcel.2008.08.014.

LI, Yuanyuan, 2021. Modern epigenetics methods in biological research. *Methods (San Diego, Calif.)*. mars 2021. Vol. 187, pp. 104-113. DOI 10.1016/j.ymeth.2020.06.022.

LI, Yufeng et SASAKI, Hiroyuki, 2011. Genomic imprinting in mammals: its life cycle, molecular mechanisms and reprogramming. *Cell Research*. mars 2011. Vol. 21, n° 3, pp. 466-473. DOI 10.1038/cr.2011.15.

LINDNER, Melanie, GAWEHNS, Fleur, TE MOLDER, Sebastiaan, VISSER, Marcel E., OERS, Kees et LAINE, Veronika N., 2022. Performance of methods to detect genetic variants from bisulphite sequencing data in a non-model species. *Molecular Ecology Resources*. février

2022. Vol. 22, n° 2, pp. 834-846. DOI 10.1111/1755-0998.13493.

LIU, Chia-Wei, CHEN, Jin-Bor, TIAO, Mao-Meng, WENG, Shao-Wen, HUANG, Tiao-Lai, CHUANG, Jiin-Haur, CHEN, Shang-Der, CHUANG, Yao-Chung, LEE, Wen-Chin, LIN, Tsu-Kung et WANG, Pei-Wen, 2012. Mitochondrial DNA coding and control region variants as genetic risk factors for type 2 diabetes. *Diabetes*. octobre 2012. Vol. 61, n° 10, pp. 2642-2651. DOI 10.2337/db11-1369.

LOYAU, Thomas, HENNEQUET-ANTIER, Christelle, COUSTHAM, Vincent, BERRI, Cécile, LEDUC, Marie, CROCHET, Sabine, SANNIER, Mélanie, DUCLOS, Michel Jacques, MIGNON-GRASTEAU, Sandrine, TESSERAUD, Sophie, BRIONNE, Aurélien, MÉTAYER-COUSTARD, Sonia, MOROLDO, Marco, LECARDONNEL, Jérôme, MARTIN, Patrice, LAGARRIGUE, Sandrine, YAHAV, Shlomo et COLLIN, Anne, 2016. Thermal manipulation of the chicken embryo triggers differential gene expression in response to a later heat challenge. *BMC genomics*. 4 mai 2016. Vol. 17, pp. 329. DOI 10.1186/s12864-016-2661-y.

MALTECCA, Christian, BERGAMASCHI, Matteo et TIEZZI, Francesco, 2020. The interaction between microbiome and pig efficiency: A review. *Journal of Animal Breeding and Genetics = Zeitschrift Fur Tierzucht Und Zuchtungsbiologie*. janvier 2020. Vol. 137, n° 1, pp. 4-13. DOI 10.1111/jbg.12443.

MANOLIO, Teri A., COLLINS, Francis S., COX, Nancy J., GOLDSTEIN, David B., HINDORFF, Lucia A., HUNTER, David J., MCCARTHY, Mark I., RAMOS, Erin M., CARDON, Lon R., CHAKRAVARTI, Aravinda, CHO, Judy H., GUTTMACHER, Alan E., KONG, Augustine, KRUGLYAK, Leonid, MARDIS, Elaine, ROTIMI, Charles N., SLATKIN, Montgomery, VALLE, David, WHITTEMORE, Alice S., BOEHNKE, Michael, CLARK, Andrew G., EICHLER, Evan E., GIBSON, Greg, HAINES, Jonathan L., MACKAY, Trudy F. C., MCCARROLL, Steven A. et VISSCHER, Peter M., 2009. Finding the missing heritability of complex diseases. *Nature*. 8 octobre 2009. Vol. 461, n° 7265, pp. 747-753. DOI 10.1038/nature08494.

MARANGA, Carina, FERNANDES, Tiago G., BEKMAN, Evguenia et DA ROCHA, Simão Teixeira, 2020. Angelman syndrome: a journey through the brain. *The FEBS Journal*. 2020. Vol. 287, n° 11, pp. 2154-2175. DOI 10.1111/febs.15258.

MARGOLIS, Seth S., SELL, Gabrielle L., ZBINDEN, Mark A. et BIRD, Lynne M., 2015. Angelman Syndrome. *Neurotherapeutics: The Journal of the American Society for Experimental NeuroTherapeutics*. juillet 2015. Vol. 12, n° 3, pp. 641-650. DOI 10.1007/s13311-015-0361-y.

MARTIN, David I. K., CROPLEY, Jennifer E. et SUTER, Catherine M., 2008. Environmental influence on epigenetic inheritance at the Avy allele. *Nutrition Reviews*. août 2008. Vol. 66 Suppl 1, pp. S12-14. DOI 10.1111/j.1753-4887.2008.00057.x.

MCGRATH, J. et SOLTER, D., 1984. Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell*. mai 1984. Vol. 37, n° 1, pp. 179-183. DOI 10.1016/0092-8674(84)90313-1.

MEHRMOHAMADI, Mahya, SEPEHRI, Mohammad Hossein, NAZER, Naghme et NOROUZI, Mohammad Reza, 2021. A Comparative Overview of Epigenomic Profiling Methods. *Frontiers in Cell and Developmental Biology*. [en ligne]. 2021. Vol. 9. [Consulté le 21 mai 2023]. Disponible à l'adresse: <https://www.frontiersin.org/articles/10.3389/fcell.2021.714687>

MENDIOLA, Aron Judd P. et LASALLE, Janine M., 2021. Epigenetics in Prader-Willi Syndrome. *Frontiers in Genetics*. 15 février 2021. Vol. 12, pp. 624581. DOI 10.3389/fgene.2021.624581.

MERKEL, Angelika, FERNÁNDEZ-CALLEJO, Marcos, CASALS, Eloi, MARCO-SOLA,

- Santiago, SCHUYLER, Ronald, GUT, Ivo G. et HEATH, Simon C., 2019. gemBS: high throughput processing for DNA methylation data from bisulfite sequencing. *Bioinformatics (Oxford, England)*. 1 mars 2019. Vol. 35, n° 5, pp. 737-742. DOI 10.1093/bioinformatics/bty690.
- METZKER, Michael L., 2010. Sequencing technologies — the next generation. *Nature Reviews Genetics*. janvier 2010. Vol. 11, n° 1, pp. 31-46. DOI 10.1038/nrg2626.
- MEUWISSEN, T. H., HAYES, B. J. et GODDARD, M. E., 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*. avril 2001. Vol. 157, n° 4, pp. 1819-1829. DOI 10.1093/genetics/157.4.1819.
- MILAN, Denis, JEON, Jin-Tae, LOOFT, Christian, AMARGER, Valerie, ROBIC, Annie, THELANDER, Mattias, ROGEL-GAILLARD, Claire, PAUL, Sven, IANNUCELLI, Nathalie, RASK, Lars, RONNE, Hans, LUNDSTRÖM, Kerstin, REINSCH, Norbert, GELLIN, Joel, KALM, Ernst, ROY, Pascale Le, CHARDON, Patrick et ANDERSSON, Leif, 2000. A Mutation in PRKAG3 Associated with Excess Glycogen Content in Pig Skeletal Muscle. *Science*. 19 mai 2000. Vol. 288, n° 5469, pp. 1248-1251. DOI 10.1126/science.288.5469.1248.
- MINISTÈRE DE L'AGRICULTURE DE L'ALIMENTATION DE LA PÊCHE ET DES AFFAIRES RURALES, 2004. *Rapport sur la filière porcine française* [en ligne]. [Consulté le 12 août 2022]. Disponible à l'adresse: https://medias.vie-publique.fr/data_storage_s3/rapport/pdf/034000221.pdf
- MINISTÈRE DE L'AGRICULTURE ET DE LA SOUVERAINETÉ ALIMENTAIRE, 2022. Agréments dans le domaine zootechnique. *Ministère de l'Agriculture et de la Souveraineté alimentaire*. [en ligne]. 4 juillet 2022. [Consulté le 17 mai 2023]. Disponible à l'adresse: <https://agriculture.gouv.fr/agrements-dans-le-domaine-zootechnique>
- MINVIELLE, F., 1992. Les bases de la génétique quantitative : Les modèles à un et deux locus. *INRAE Productions Animales*. 29 décembre 1992. Vol. 5, n° HS, pp. 61-67. DOI 10.20870/productions-animales.1992.5.HS.4263.
- MIRI, Kamelia, LATHAM, Keith, PANNING, Barbara, ZHONG, Zhisheng, ANDERSEN, Angela et VARMUZA, Susannah, 2013. The imprinted polycomb group gene Sfbmt2 is required for trophoblast maintenance and placenta development. *Development (Cambridge, England)*. novembre 2013. Vol. 140, n° 22, pp. 4480-4489. DOI 10.1242/dev.096511.
- MONK, David, 2015. Genomic imprinting in the human placenta. *American Journal of Obstetrics and Gynecology*. octobre 2015. Vol. 213, n° 4 Suppl, pp. S152-162. DOI 10.1016/j.ajog.2015.06.032.
- MONK, David, ARNAUD, Philippe, FROST, Jennifer, HILLS, Frank A., STANIER, Philip, FEIL, Robert et MOORE, Gudrun E., 2009. Reciprocal imprinting of human GRB10 in placental trophoblast and brain: evolutionary conservation of reversed allelic expression. *Human Molecular Genetics*. 15 août 2009. Vol. 18, n° 16, pp. 3066-3074. DOI 10.1093/hmg/ddp248.
- MONK, David, MACKAY, Deborah J. G., EGGERMANN, Thomas, MAHER, Eamonn R. et RICCIO, Andrea, 2019. Genomic imprinting disorders: lessons on how genome, epigenome and environment interact. *Nature Reviews Genetics*. avril 2019. Vol. 20, n° 4, pp. 235-248. DOI 10.1038/s41576-018-0092-0.
- MONTGOMERY, Sean A. et BERGER, Frédéric, 2021. The evolution of imprinting in plants: beyond the seed. *Plant Reproduction*. 2021. Vol. 34, n° 4, pp. 373-383. DOI 10.1007/s00497-021-00410-7.
- MOORE, Gudrun E., ISHIDA, Miho, DEMETRIOU, Charalambos, AL-OLABI, Lara, LEON, Lydia J., THOMAS, Anna C., ABU-AMERO, Sayeda, FROST, Jennifer M., STAFFORD, Jaime L., CHAOQUN, Yao, DUNCAN, Andrew J., BAIGEL, Rachel, BRIMIOULLE, Marina,

IGLESIAS-PLATAS, Isabel, APOSTOLIDOU, Sophia, AGGARWAL, Reena, WHITTAKER, John C., SYNGELAKI, Argyro, NICOLAIDES, Kypros H., REGAN, Lesley, MONK, David et STANIER, Philip, 2015. The role and interaction of imprinted genes in human fetal growth. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 5 mars 2015. Vol. 370, n° 1663, pp. 20140074. DOI 10.1098/rstb.2014.0074.

MOORE, S. S., SARGEANT, L. L., KING, T. J., MATTICK, J. S., GEORGES, M. et HETZEL, D. J., 1991. The conservation of dinucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species. *Genomics*. juillet 1991. Vol. 10, n° 3, pp. 654-660. DOI 10.1016/0888-7543(91)90448-n.

MOYES, C., BATTERSBY, B. et S, null, 1998. Regulation of muscle mitochondrial design. *The Journal of Experimental Biology*. 1998. Vol. 201, n° 3, pp. 299-307.

NAGANO, Takashi, MITCHELL, Jennifer A., SANZ, Lionel A., PAULER, Florian M., FERGUSON-SMITH, Anne C., FEIL, Robert et FRASER, Peter, 2008. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science (New York, N.Y.)*. 12 décembre 2008. Vol. 322, n° 5908, pp. 1717-1720. DOI 10.1126/science.1163802.

NATIONAL HUMAN GENOME RESEARCH INSTITUTE, 2023. DNA Sequencing Costs : Data. [en ligne]. 2023. Disponible à l'adresse: <https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data>

NERY DA SILVA, Arthur, SILVA ARAUJO, Michelle, PÉRTILLE, Fábio et ZANELLA, Adroaldo José, 2021. How Epigenetics Can Enhance Pig Welfare? *Animals : an Open Access Journal from MDPI*. 24 décembre 2021. Vol. 12, n° 1, pp. 32. DOI 10.3390/ani12010032.

NEZER, C., MOREAU, L., BROUWERS, B., COPPIETERS, W., DETILLEUX, J., HANSET, R., KARIM, L., KVASZ, A., LEROY, P. et GEORGES, M., 1999. An imprinted QTL with major effect on muscle mass and fat deposition maps to the IGF2 locus in pigs. *Nature Genetics*. février 1999. Vol. 21, n° 2, pp. 155-156. DOI 10.1038/5935.

NURK, Sergey, KOREN, Sergey, RHIE, Arang, RAUTIAINEN, Mikko, BZIKADZE, Andrey V., MIKHEENKO, Alla, VOLLGER, Mitchell R., ALTEMOSE, Nicolas, URALSKY, Lev, GERSHMAN, Ariel, AGANEZOV, Sergey, HOYT, Savannah J., DIEKHANS, Mark, LOGSDON, Glennis A., ALONGE, Michael, ANTONARAKIS, Stylianos E., BORCHERS, Matthew, BOUFFARD, Gerard G., BROOKS, Shelise Y., CALDAS, Gina V., CHEN, Nae-Chyun, CHENG, Haoyu, CHIN, Chen-Shan, CHOW, William, DE LIMA, Leonardo G., DISHUCK, Philip C., DURBIN, Richard, DVORKINA, Tatiana, FIDDES, Ian T., FORMENTI, Giulio, FULTON, Robert S., FUNGTAMMASAN, Arkarachai, GARRISON, Erik, GRADY, Patrick G. S., GRAVES-LINDSAY, Tina A., HALL, Ira M., HANSEN, Nancy F., HARTLEY, Gabrielle A., HAUKNES, Marina, HOWE, Kerstin, HUNKAPILLER, Michael W., JAIN, Chirag, JAIN, Miten, JARVIS, Erich D., KERPEDJIEV, Peter, KIRSCH, Melanie, KOLMOGOROV, Mikhail, KORLACH, Jonas, KREMITZKI, Milinn, LI, Heng, MADURO, Valerie V., MARSCHALL, Tobias, MCCARTNEY, Ann M., MCDANIEL, Jennifer, MILLER, Danny E., MULLIKIN, James C., MYERS, Eugene W., OLSON, Nathan D., PATEN, Benedict, PELUSO, Paul, PEVZNER, Pavel A., PORUBSKY, David, POTAPOVA, Tamara, ROGAEV, Evgeny I., ROSENFELD, Jeffrey A., SALZBERG, Steven L., SCHNEIDER, Valerie A., SEDLAZECK, Fritz J., SHAFIN, Kishwar, SHEW, Colin J., SHUMATE, Alaina, SIMS, Ying, SMIT, Arian F. A., SOTO, Daniela C., SOVIĆ, Ivan, STORER, Jessica M., STREETS, Aaron, SULLIVAN, Beth A., THIBAUD-NISSEN, Françoise, TORRANCE, James, WAGNER, Justin, WALENZ, Brian P., WENGER, Aaron, WOOD, Jonathan M. D., XIAO, Chunlin, YAN, Stephanie M., YOUNG, Alice C., ZARATE, Samantha, SURTI, Urvashi, MCCOY, Rajiv C., DENNIS, Megan Y., ALEXANDROV, Ivan A., GERTON, Jennifer L., O'NEILL, Rachel J., TIMP, Winston, ZOOK, Justin M., SCHATZ, Michael C., EICHLER, Evan E., MIGA, Karen H. et PHILLIPPY, Adam M., 2022. The

complete sequence of a human genome. *Science*. avril 2022. Vol. 376, n° 6588, pp. 44-53. DOI 10.1126/science.abj6987.

ODLING-SMEE, F., LALA, Kevin et FELDMAN, Marcus, 2003. *Niche Construction: The Neglected Process in Evolution*. . ISBN 978-1-4008-4726-6. ces.journalAbbreviation: Niche Construction: The Neglected Process in Evolution

OKAE, Hiroaki, MATOBA, Shogo, NAGASHIMA, Takeshi, MIZUTANI, Eiji, INOUE, Kimiko, OGONUKI, Narumi, CHIBA, Hatsune, FUNAYAMA, Ryo, TANAKA, Satoshi, YAEGASHI, Nobuo, NAKAYAMA, Keiko, SASAKI, Hiroyuki, OGURA, Atsuo et ARIMA, Takahiro, 2014. RNA sequencing-based identification of aberrant imprinting in cloned mice. *Human Molecular Genetics*. 15 février 2014. Vol. 23, n° 4, pp. 992-1001. DOI 10.1093/hmg/ddt495..

OLLIVIER, L., 1986. La loi sur l'élevage de 1966 : un bilan de 20 années de sélection rationnelle du porc en France. *Journées de la Recherche Porcine*. 1986. pp. 203-236.

OLLIVIER, Louis, RUNAVOT, J. P., DAGORN, Jean-Cyrille, GUEBLEZ, R., JEHANNO, J., KERISIT, R., LEGAULT, Christian, MOLENAT, M. et SELLIER, Pierre P., 1986. La loi sur l'élevage de 1966 : un bilan de 20 années de sélection rationnelle du porc en France. *Annales de zootechnie*. 1986. Vol. 35, pp. 302.

OLLIVIER, Louis, SANCRISTOBAL, Magali, FOULLEY, Jean-Louis, LEGAULT, Christian, LAVAL, Guillaume, MILAN, Denis, CARITEZ, Jean-Claude, BURGAUD, Germaine, GRUAND, Joseph, BOSCHER, Marie-Yvonne, AMIGUES, Yves, BARRET, Julien, HASSENFRAZ, Claire, LUQUET, Michel, LABROUE, Florence et CHEVALET, Claude, 2004. Analyse de la diversité de quelques races et lignées porcines françaises à l'aide de marqueurs génétiques dans le cadre d'un vaste projet européen. *Journées de la Recherche Porcine*. 2004. N° 36, pp. 265-274. Denis MILAN (2), Jean-Claude CARITEZ (3), Germaine BURGAUD (3), Joseph GRUAND (4), Marie-Yvonne BOSCHER (5), Yves AMIGUES (5), Julien BARRET (6), Claire HASSENFRAZ (6), Michel LUQUET (7), Florence LABROUE (8), Claude CHEVALET (2)

ONO, Ryuichi, NAKAMURA, Kenji, INOUE, Kimiko, NARUSE, Mie, USAMI, Takako, WAKISAKA-SAITO, Noriko, HINO, Toshiaki, SUZUKI-MIGISHIMA, Rika, OGONUKI, Narumi, MIKI, Hiromi, KOHDA, Takashi, OGURA, Atsuo, YOKOYAMA, Minesuke, KANEKO-ISHINO, Tomoko et ISHINO, Fumitoshi, 2006. Deletion of Peg10, an imprinted gene acquired from a retrotransposon, causes early embryonic lethality. *Nature Genetics*. janvier 2006. Vol. 38, n° 1, pp. 101-106. DOI 10.1038/ng1699.

OXFORD NANOPORE TECHNOLOGIES, [sans date]. MinION. [en ligne]. Disponible à l'adresse: <https://nanoporetech.com/products/minion>

OXFORD NANOPORE TECHNOLOGIES, [sans date]. PromethION. [en ligne]. Disponible à l'adresse: <https://nanoporetech.com/products/promethion>

PANKEY, C. L., ODHIAMBO, J. F., SMITH, A. M. et FORD, S. P., 2021. Effects of maternal obesity in an ovine model on metabolic outcomes in F2 adults and F3 neonates. *Domestic animal endocrinology*. juillet 2021. Vol. 76, pp. 106628. DOI 10.1016/j.domaniend.2021.106628.

PARLEMENT EUROPÉEN, 2021. Le Parlement Européen approuve l'appel citoyen à la disparition progressive de l'élevage en cage. [en ligne]. 10 juin 2021. [Consulté le 17 mai 2023]. Disponible à l'adresse: <https://www.europarl.europa.eu/news/fr/press-room/20210604IPR05532/le-pe-approuve-l-appel-citoyen-a-la-disparition-progressive-de-l-elevage-en-cage>L'élevage en cage pourrait progressivement disparaître en Europe d'ici 2027, affirment les députés.

PAROIS, Séverine, BONNEAU, Michel, CHEVILLON, Patrick, LARZUL, Catherine,

- QUINIOU, Nathalie, ROBIC, Annie et PRUNIER, Armelle, 2018. Odeurs indésirables de la viande de porcs mâles non castrés : problèmes et solutions potentielles. *INRAE Productions Animales*. 11 juin 2018. Vol. 31, n° 1, pp. 23-36. DOI 10.20870/productions-animales.2018.31.1.2206.
- PATTEN, M. M., ROSS, L., CURLEY, J. P., QUELLER, D. C., BONDURIANSKY, R. et WOLF, J. B., 2014. The evolution of genomic imprinting: theories, predictions and empirical tests. *Heredity*. août 2014. Vol. 113, n° 2, pp. 119-128. DOI 10.1038/hdy.2014.29.
- PEREZ, Julio D., RUBINSTEIN, Nimrod D. et DULAC, Catherine, 2016. New Perspectives on Genomic Imprinting, an Essential and Multifaceted Mode of Epigenetic Control in the Developing and Adult Brain. *Annual Review of Neuroscience*. 2016. Vol. 39, n° 1, pp. 347-384. DOI 10.1146/annurev-neuro-061010-113708.
- PHAM, Nga V, NGUYEN, Mai T, HU, Ji-Fan, VU, Thanh H et HOFFMAN, Andrew R, 1998. Dissociation of IGF2 and H19 imprinting in human brain. *Brain Research*. 1998. Vol. 810, pp. 1-8.
- PHOCAS, Florence, 2012. Pour mieux comprendre les index BLUP - Modèle animal. *Institut de l'Élevage*. [en ligne]. 24 septembre 2012. [Consulté le 17 mai 2023]. Disponible à l'adresse: <https://idele.fr/detail-article/pour-mieux-comprendre-les-index-blup-modele-animal>Cette note IBOVAL s'adresse a ceux qui souhaitent mieux comprendre le calcul des index bovins allaitants : de la construction et l'apport des différentes origines de données dans ces estimations à l'interprétation des index, notamment des femelles.
- PINEL, Manuela, 2023. Le Porc par les Chiffres : les données les plus récentes de la filière porcine. [en ligne]. 22 mars 2023. [Consulté le 17 mai 2023]. Disponible à l'adresse: <https://ifip.asso.fr/actualites/le-porc-par-les-chiffres-les-chiffres-cles-les-plus-recents-de-la-filiere/>
- PITEL, Frédérique, CALENGE, Fanny, AIGUEPERSE, Nadège, ESTELLÉ-FABRELLAS, Jordi, COUSTHAM, Vincent, CALANDREAU, Ludovic, MORISSON, Mireille, CHAVATTE-PALMER, Pascale et GINANE, Cécile, 2019. Rôle de l'environnement précoce dans la variabilité des phénotypes et l'adaptation des animaux d'élevage à leur milieu. *INRAE Productions Animales*. 15 avril 2019. Vol. 32, n° 2, pp. 247-262. DOI 10.20870/productions-animales.2019.32.2.2467.
- PRICKETT, Adam R. et OAKEY, Rebecca J., 2012. A survey of tissue-specific genomic imprinting in mammals. *Molecular Genetics and Genomics*. 1 août 2012. Vol. 287, n° 8, pp. 621-630. DOI 10.1007/s00438-012-0708-6.
- RASMUSSEN, Erik M. K. et AMDAM, Gro V., 2015. Cytosine modifications in the honey bee (*Apis mellifera*) worker genome. *Frontiers in Genetics*. 2015. Vol. 6, pp. 8. DOI 10.3389/fgene.2015.00008.
- RASSOULZADEGAN, Minoo, GRANDJEAN, Valérie, GOUNON, Pierre, VINCENT, Stéphane, GILLOT, Isabelle et CUZIN, François, 2006. RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. *Nature*. 25 mai 2006. Vol. 441, n° 7092, pp. 469-474. DOI 10.1038/nature04674.
- REAMON-BUETTNER, Stella Marie, MUTSCHLER, Vanessa et BORLAK, Juergen, 2008. The next innovation cycle in toxicogenomics: environmental epigenetics. *Mutation Research*. 2008. Vol. 659, n° 1-2, pp. 158-165. DOI 10.1016/j.mrrev.2008.01.003.
- REIK, W. et WALTER, J., 2001. Genomic imprinting: parental influence on the genome. *Nature Reviews. Genetics*. janvier 2001. Vol. 2, n° 1, pp. 21-32. DOI 10.1038/35047554.
- REVILLA, M., FRIGGENS, N. C., BROUDISCOU, L. P., LEMONNIER, G., BLANC, F.,

RAVON, L., MERCAT, M. J., BILLON, Y., ROGEL-GAILLARD, C., LE FLOCH, N., ESTELLÉ, J. et MUÑOZ-TAMAYO, R., 2019. Towards the quantitative characterisation of piglets' robustness to weaning: a modelling approach. *Animal: An International Journal of Animal Bioscience*. novembre 2019. Vol. 13, n° 11, pp. 2536-2546. DOI 10.1017/S1751731119000843.

RICHARD ALBERT, Julien, KOBAYASHI, Toshihiro, INOUE, Azusa, MONTEAGUDO-SÁNCHEZ, Ana, KUMAMOTO, Soichiro, TAKASHIMA, Tomoya, MIURA, Asuka, OIKAWA, Mami, MIURA, Fumihito, TAKADA, Shuji, HIRABAYASHI, Masumi, KORTHAUER, Keegan, KURIMOTO, Kazuki, GREENBERG, Maxim V. C., LORINCZ, Matthew et KOBAYASHI, Hisato, 2023. Conservation and divergence of canonical and non-canonical imprinting in murids. *Genome Biology*. 14 mars 2023. Vol. 24, n° 1, pp. 48. DOI 10.1186/s13059-023-02869-1.

RIDGE, Perry G., MAXWELL, Taylor J., CORCORAN, Christopher D., NORTON, Maria C., TSCHANZ, Joann T., O'BRIEN, Elizabeth, KERBER, Richard A., CAWTHON, Richard M., MUNGER, Ronald G. et KAUWE, John S. K., 2012. Mitochondrial genomic analysis of late onset Alzheimer's disease reveals protective haplogroups H6A1A/H6A1B: the Cache County Study on Memory in Aging. *PloS One*. 2012. Vol. 7, n° 9, pp. e45134. DOI 10.1371/journal.pone.0045134.

RIQUET, J., ROUSSEAU, S., MERCAT, M. J., PAILHOUS, E. et LARZUL, C., 2016. Les anomalies congénitales héréditaires chez le porc. *INRAE Productions Animales*. 28 décembre 2016. Vol. 29, n° 5, pp. 329-338. DOI 10.20870/productions-animales.2016.29.5.3001.

ROBBINS, Katherine Marie, CHEN, Zhiyuan, WELLS, Kevin Dale et RIVERA, Rocio Melissa, 2012. Expression of KCNQ1OT1, CDKN1C, H19, and PLAGL1 and the methylation patterns at the KvDMR1 and H19/IGF2 imprinting control regions is conserved between human and bovine. *Journal of Biomedical Science*. 15 novembre 2012. Vol. 19, n° 1, pp. 95. DOI 10.1186/1423-0127-19-95.

ROBERT-GRANIÉ, Christèle C., LEGARRA, Andres et DUCROCQ, Vincent V., 2011. Principes de base de la sélection génomique. MULSANT, Philippe, BODIN, Loys, COUDURIER, Bernard, DERETZ, Séverine, ROY, Pascale Le, QUILLET, Edwige et PEREZ, Jean-Marc (éd.), *Productions Animales*. 2011. Vol. 24, n° 4, pp. 331-340.

ROGEL-GAILLARD, Claire, ESTELLÉ, Jordi, BOREY, Marion et LARZUL, Catherine, 2021. Interplay between host genetics and the dynamics of the gut microbiota in pigs. In : *72nd Annual Meeting of the European Federation of Animal Science (EAAP)*. [en ligne]. 30 août 2021. [Consulté le 27 avril 2023]. Disponible à l'adresse: <https://hal.science/hal-03786280>

ROGUET, Christine, 2023. Evolution, diversité et typologie des exploitations porcines en France : enseignements du recensement agricole de 2020, comparaison aux recensements de 2010 et 2000. *Journées de la Recherche Porcine*. 2023. N° 55, pp. 1-6.

ROGUET, Christine et DELANOUE, Elsa, 2015. Perception de l'élevage par de jeunes adultes en France en 2014. *Journées de la Recherche Porcine*. 2015. N° 47, pp. 227-228.

ROUGIER, N. et WERB, Z., 2001. Minireview: Parthenogenesis in mammals. *Molecular Reproduction and Development*. août 2001. Vol. 59, n° 4, pp. 468-474. DOI 10.1002/mrd.1054.

ROY, Pascale LE, DUCOS, Alain et PHOCAS, Florence, 2019. Quelles performances pour les animaux de demain ? Objectifs et méthodes de sélection. *INRAE Productions Animales*. 11 avril 2019. Vol. 32, n° 2, pp. 233-246. DOI 10.20870/productions-animales.2019.32.2.2466.

- SALAS, Ramon Cesar D. et MINGALA, Claro N., 2017. Genetic Factors Affecting Pork Quality: Halothane and Rendement Napole Genes. *Animal Biotechnology*. 3 avril 2017. Vol. 28, n° 2, pp. 148-155. DOI 10.1080/10495398.2016.1243550.
- SANCHEZ, Maris-Pierre, RIQUET, Juliette, FÈVE, Katia, GILBERT, Hélène, LE ROY, Pascal, IANNUCELLI, Nathalie, GOGUÉ, Jean, PÉRY, Christophe, BIDANEL, Jean-Pierre et MILAN, Denis, 2003. Effets de la région du gène IGF2 et du gène halothane sur la composition corporelle et la qualité de la viande dans une population F2 Piétrain x Large White. . 2003. N° 35, pp. 269-276.
- SANDOVAL-MOTTA, Santiago, ALDANA, Maximino, MARTÍNEZ-ROMERO, Esperanza et FRANK, Alejandro, 2017. The Human Microbiome and the Missing Heritability Problem. *Frontiers in Genetics*. 2017. Vol. 8, pp. 80. DOI 10.3389/fgene.2017.00080.
- SANGER, F., NICKLEN, S. et COULSON, A. R., 1977. DNA Sequencing with Chain-Terminating Inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*. 1977. Vol. 74, n° 12, pp. 5463-5467.
- SANMIGUEL, Jennifer M et BARTOLOMEI, Marisa S, 2018. DNA methylation dynamics of genomic imprinting in mouse development. *Biology of Reproduction*. juillet 2018. Vol. 99, n° 1, pp. 252-262. DOI 10.1093/biolre/iroy036.
- SAUGÈRE, Dominique, 1989. Un premier bilan du programme de sélection contre le gène de sensibilité à l'halothane chez le porc Landrace français. *Journées de la Recherche Porcine*. 1989. N° 21, pp. 335-344.
- SCHENA, M., SHALON, D., DAVIS, R. W. et BROWN, P. O., 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science (New York, N.Y.)*. 20 octobre 1995. Vol. 270, n° 5235, pp. 467-470. DOI 10.1126/science.270.5235.467.
- SCHLOSS, Jeffery A., 2008. How to get genomes at one ten-thousandth the cost. *Nature Biotechnology*. octobre 2008. Vol. 26, n° 10, pp. 1113-1115. DOI 10.1038/nbt1008-1113.
- SCHWARTZ, Lauren, CAIXÀS, Assumpta, DIMITROPOULOS, Anastasia, DYKENS, Elisabeth, DUIS, Jessica, EINFELD, Stewart, GALLAGHER, Louise, HOLLAND, Anthony, RICE, Lauren, ROOF, Elizabeth, SALEHI, Parisa, STRONG, Theresa, TAYLOR, Bonnie et WOODCOCK, Kate, 2021. Behavioral features in Prader-Willi syndrome (PWS): consensus paper from the International PWS Clinical Trial Consortium. *Journal of Neurodevelopmental Disorders*. 21 juin 2021. Vol. 13, n° 1, pp. 25. DOI 10.1186/s11689-021-09373-2.
- SÉLÉNOU, Céline, BRIOUDE, Frédéric, GIABICANI, Eloïse, SOBRIER, Marie-Laure et NETCHINE, Irène, 2022. IGF2: Development, Genetic and Epigenetic Abnormalities. *Cells*. 10 juin 2022. Vol. 11, n° 12, pp. 1886. DOI 10.3390/cells11121886.
- SEMAGN, Kassa, BABU, Raman, HEARNE, Sarah et OLSEN, Michael, 2014. Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. *Molecular Breeding*. 1 janvier 2014. Vol. 33, n° 1, pp. 1-14. DOI 10.1007/s11032-013-9917-x.
- SHA, Ky, 2008. A Mechanistic View of Genomic Imprinting. *Annual Review of Genomics and Human Genetics*. 2008. Vol. 9, n° 1, pp. 197-216. DOI 10.1146/annurev.genom.122007.110031.
- SHEN, Lijun, WEI, Jia, CHEN, Tao, HE, Jing, QU, Jianchun, HE, Xiumei, JIANG, Luxi, QU, Yemin, FANG, Hezhi, CHEN, Guorong, LU, Jianxin et BAI, Yidong, 2011. Evaluating mitochondrial DNA in patients with breast cancer and benign breast disease. *Journal of Cancer*

Research and Clinical Oncology. avril 2011. Vol. 137, n° 4, pp. 669-675. DOI 10.1007/s00432-010-0912-x.

SHI, Yu, WANG, Guoping, LAU, Harry Cheuk-Hay et YU, Jun, 2022. Metagenomic Sequencing for Microbial DNA in Human Samples: Emerging Technological Advances. *International Journal of Molecular Sciences*. 16 février 2022. Vol. 23, n° 4, pp. 2181. DOI 10.3390/ijms23042181.

SHMELA, Mansur Ennuri et GICQUEL, Christine F., 2013. Human diseases versus mouse models: insights into the regulation of genomic imprinting at the human 11p15/mouse distal chromosome 7 region. *Journal of Medical Genetics*. 1 janvier 2013. Vol. 50, n° 1, pp. 11-20. DOI 10.1136/jmedgenet-2012-101321.

SIBLEY, C. P., COAN, P. M., FERGUSON-SMITH, A. C., DEAN, W., HUGHES, J., SMITH, P., REIK, W., BURTON, G. J., FOWDEN, A. L. et CONSTÂNCIA, M., 2004. Placental-specific insulin-like growth factor 2 (Igf2) regulates the diffusional exchange characteristics of the mouse placenta. *Proceedings of the National Academy of Sciences*. 25 mai 2004. Vol. 101, n° 21, pp. 8204-8208. DOI 10.1073/pnas.0402508101.

SKINNER, Michael K., 2011. Environmental epigenetic transgenerational inheritance and somatic epigenetic mitotic stability. *Epigenetics*. juillet 2011. Vol. 6, n° 7, pp. 838-842. DOI 10.4161/epi.6.7.16537.

SLATKO, Barton E., GARDNER, Andrew F. et AUSUBEL, Frederick M., 2018. Overview of Next Generation Sequencing Technologies. *Current protocols in molecular biology*. avril 2018. Vol. 122, n° 1, pp. e59. DOI 10.1002/cpmb.59.

SPENCER, H. G. et CLARK, A. G., 2014. Non-conflict theories for the evolution of genomic imprinting. *Heredity*. août 2014. Vol. 113, n° 2, pp. 112-118. DOI 10.1038/hdy.2013.129.

ST JOHN, Justin C., OKADA, Takashi, ANDREAS, Eryk et PENN, Alexander, 2022. The role of mtDNA in oocyte quality and embryo development. *Molecular Reproduction and Development*. 20 août 2022. DOI 10.1002/mrd.23640.

ST JOHN, Justin C. et TSAI, Te-Sha, 2018. The association of mitochondrial DNA haplotypes and phenotypic traits in pigs. *BMC genetics*. 6 juillet 2018. Vol. 19, n° 1, pp. 41. DOI 10.1186/s12863-018-0629-4.

STRINGER, Jessica M., SUZUKI, Shunsuke, PASK, Andrew J., SHAW, Geoff et RENFREE, Marilyn B., 2012. Promoter-Specific Expression and Imprint Status of Marsupial IGF2. *PLoS ONE*. 25 juillet 2012. Vol. 7, n° 7, pp. e41690. DOI 10.1371/journal.pone.0041690.

SURANI, M. A., BARTON, S. C. et NORRIS, M. L., 1984. Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature*. 5 avril 1984. Vol. 308, n° 5959, pp. 548-550. DOI 10.1038/308548a0.

TAKAHASHI, Nozomi, COLUCCIO, Andrea, THORBALL, Christian W., PLANET, Evarist, SHI, Hui, OFFNER, Sandra, TURELLI, Priscilla, IMBEAULT, Michael, FERGUSON-SMITH, Anne C. et TRONO, Didier, 2019. ZNF445 is a primary regulator of genomic imprinting. *Genes & Development*. 1 janvier 2019. Vol. 33, n° 1-2, pp. 49-54. DOI 10.1101/gad.320069.118.

THORVALDSDÓTTIR, Helga, ROBINSON, James T. et MESIROV, Jill P., 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics*. mars 2013. Vol. 14, n° 2, pp. 178-192. DOI 10.1093/bib/bbs017.

- TIFFON, Céline, 2018. The Impact of Nutrition and Environmental Epigenetics on Human Health and Disease. *International Journal of Molecular Sciences*. 1 novembre 2018. Vol. 19, n° 11, pp. 3425. DOI 10.3390/ijms19113425.
- TREROTOLA, Marco, RELI, Valeria, SIMEONE, Pasquale et ALBERTI, Saverio, 2015. Epigenetic inheritance and the missing heritability. *Human Genomics*. 28 juillet 2015. Vol. 9, n° 1, pp. 17. DOI 10.1186/s40246-015-0041-3.
- TRIBOUT, T., 2011. Perspectives d'application de la sélection génomique dans les schémas d'amélioration génétique porcins. *INRAE Productions Animales*. 8 septembre 2011. Vol. 24, n° 4, pp. 369-376. DOI 10.20870/productions-animales.2011.24.4.3270.
- TSAI, Te-Sha, RAJASEKAR, Sriram et ST JOHN, Justin C., 2016. The relationship between mitochondrial DNA haplotype and the reproductive capacity of domestic pigs (*Sus scrofa domestica*). *BMC genetics*. 18 mai 2016. Vol. 17, n° 1, pp. 67. DOI 10.1186/s12863-016-0375-4.
- TUCCI, Valter, ISLES, Anthony R., KELSEY, Gavin, FERGUSON-SMITH, Anne C., TUCCI, Valter, BARTOLOMEI, Marisa S., BENVENISTY, Nissim, BOURC'HIS, Déborah, CHARALAMBOUS, Marika, DULAC, Catherine, FEIL, Robert, GLASER, Juliane, HUELSMANN, Lisa, JOHN, Rosalind M., MCNAMARA, Gráinne I., MOORWOOD, Kim, MUSCATELLI, Françoise, SASAKI, Hiroyuki, STRASSMANN, Beverly I., VINCENZ, Claudius, WILKINS, Jon, ISLES, Anthony R., KELSEY, Gavin et FERGUSON-SMITH, Anne C., 2019. Genomic Imprinting and Physiological Processes in Mammals. *Cell*. 21 février 2019. Vol. 176, n° 5, pp. 952-965. DOI 10.1016/j.cell.2019.01.043.
- UMLAUF, David, GOTO, Yuji, CAO, Ru, CERQUEIRA, Frédérique, WAGSCHAL, Alexandre, ZHANG, Yi et FEIL, Robert, 2004. Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nature Genetics*. décembre 2004. Vol. 36, n° 12, pp. 1296-1300. DOI 10.1038/ng1467.
- URSING, B. M. et ARNASON, U., 1998. The complete mitochondrial DNA sequence of the pig (*Sus scrofa*). *Journal of Molecular Evolution*. septembre 1998. Vol. 47, n° 3, pp. 302-306. DOI 10.1007/pl00006388.
- VAIMAN, D., 2000. Etablissement des cartes génétiques. *INRAE Productions Animales*. 22 décembre 2000. Vol. 13, n° HS, pp. 73-78. DOI 10.20870/productions-animales.2000.13.HS.3814.
- VAN LAERE, Anne-Sophie, NGUYEN, Minh, BRAUNSCHWEIG, Martin, NEZER, Carine, COLLETTE, Catherine, MOREAU, Laurence, ARCHIBALD, Alan L., HALEY, Chris S., BUYS, Nadine, TALLY, Michael, ANDERSSON, Göran, GEORGES, Michel et ANDERSSON, Leif, 2003. A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature*. 23 octobre 2003. Vol. 425, n° 6960, pp. 832-836. DOI 10.1038/nature02064.
- VAN DIJK, Erwin L., AUGER, Hélène, JASZCZYSZYN, Yan et THERMES, Claude, 2014. Ten years of next-generation sequencing technology. *Trends in Genetics*. 1 septembre 2014. Vol. 30, n° 9, pp. 418-426. DOI 10.1016/j.tig.2014.07.001.
- VAN DIJK, Erwin L., JASZCZYSZYN, Yan, NAQUIN, Delphine et THERMES, Claude, 2018. The Third Revolution in Sequencing Technology. *Trends in Genetics*. septembre 2018. Vol. 34, n° 9, pp. 666-681. DOI 10.1016/j.tig.2018.05.008.
- VAN DIJK, Erwin et THERMES, Claude, 2021. La révolution de la génomique : les nouvelles

méthodes de séquençage et leurs applications. *Planet Vie ENS*. [en ligne]. 2021. Disponible à l'adresse: [https://planet-vie.ens.fr/thematiques/manipulations-en-laboratoire/la-revolution-de-la-genomique-les-nouvelles-methodes-de#:~:text=Les%20m%C3%A9thodes%20de%20s%C3%A9quen%C3%A7age%20de,micro%2Dr%C3%A9seaux%20\(microarrays\)](https://planet-vie.ens.fr/thematiques/manipulations-en-laboratoire/la-revolution-de-la-genomique-les-nouvelles-methodes-de#:~:text=Les%20m%C3%A9thodes%20de%20s%C3%A9quen%C3%A7age%20de,micro%2Dr%C3%A9seaux%20(microarrays).).

VARMUZA, S. et MANN, M., 1994. Genomic imprinting--defusing the ovarian time bomb. *Trends in genetics: TIG*. avril 1994. Vol. 10, n° 4, pp. 118-123. DOI 10.1016/0168-9525(94)90212-7.

VERRIER, Etienne, MILAN, Denis et ROGEL-GAILLARD, Claire, 2020. *Génétique des animaux d'élevage: diversité et adaptation dans un monde changeant*. . Versailles : Éditions Quae. Savoir faire. ISBN 978-2-7592-3099-0. 636.082 GEN, 636 GEN, 636.082 1 VER

VISSCHER, Peter M., HILL, William G. et WRAY, Naomi R., 2008. Heritability in the genomics era--concepts and misconceptions. *Nature Reviews. Genetics*. avril 2008. Vol. 9, n° 4, pp. 255-266. DOI 10.1038/nrg2322.

WADDINGTON, C., 2009. The Pupal Contraction as an Epigenetic Crisis in *Drosophila*. *Proceedings of the Zoological Society of London*. 20 août 2009. Vol. A111, pp. 181-188. DOI 10.1111/j.1469-7998.1942.tb08480.x. 1.

WATANABE, Toshiaki, TOMIZAWA, Shin-ichi, MITSUYA, Kohzoh, TOTOKI, Yasushi, YAMAMOTO, Yasuhiro, KURAMOCHI-MIYAGAWA, Satomi, IIDA, Naoko, HOKI, Yuko, MURPHY, Patrick J., TOYODA, Atsushi, GOTOH, Kengo, HIURA, Hitoshi, ARIMA, Takahiro, FUJIYAMA, Asao, SADO, Takashi, SHIBATA, Tatsuhiro, NAKANO, Toru, LIN, Haifan, ICHIYANAGI, Kenji, SOLOWAY, Paul D. et SASAKI, Hiroyuki, 2011. Role for piRNAs and noncoding RNA in de novo DNA methylation of the imprinted mouse *Rasgrfl* locus. *Science (New York, N.Y.)*. 13 mai 2011. Vol. 332, n° 6031, pp. 848-852. DOI 10.1126/science.1203919.

WEAVER, Jamie R., SARKISIAN, Garnik, KRAPP, Christopher, MAGER, Jesse, MANN, Mellissa R. W. et BARTOLOMEI, Marisa S., 2010. Domain-specific response of imprinted genes to reduced DNMT1. *Molecular and Cellular Biology*. août 2010. Vol. 30, n° 16, pp. 3916-3928. DOI 10.1128/MCB.01278-09.

WEBSTER, Matthew T., 2020. Population Genomics: How Do Cape Honey Bees Do Without Sex? *Current Biology*. juillet 2020. Vol. 30, n° 14, pp. R820-R821. DOI 10.1016/j.cub.2020.05.014.

WILKINS, Jon F., 2014. Genomic Imprinting of *Grb10* : Coadaptation or Conflict? *PLoS Biology*. 25 février 2014. Vol. 12, n° 2, pp. e1001800. DOI 10.1371/journal.pbio.1001800.

WILKINSON, Lawrence S., DAVIES, William et ISLES, Anthony R., 2007. Genomic imprinting effects on brain development and function. *Nature Reviews. Neuroscience*. novembre 2007. Vol. 8, n° 11, pp. 832-843. DOI 10.1038/nrn2235.

WOLF, Jason B et HAGER, Reinmar, 2006. A Maternal-Offspring Coadaptation Theory for the Evolution of Genomic Imprinting. *PLoS Biology*. décembre 2006. Vol. 4, n° 12, pp. e380. DOI 10.1371/journal.pbio.0040380.

WRZESKA, M., ŻYGA, A., REJDUCH, B. et SŁOTA, E., 2006. A note on biallelic expression of the *IGF2* gene in the liver and brain of adult pigs. *Journal of Animal and Feed Sciences*. 6 janvier 2006. Vol. 15, n° 1, pp. 57-60. DOI 10.22358/jafs/66840/2006.

WU, Xin, GALBRAITH, David A., CHATTERJEE, Paramita, JEONG, Hyeonsoo, GROZINGER, Christina M. et YI, Soojin V., 2020. Lineage and Parent-of-Origin Effects in DNA Methylation of Honey Bees (*Apis mellifera*) Revealed by Reciprocal Crosses and Whole-Genome Bisulfite Sequencing. *Genome Biology and Evolution*. 1 août 2020. Vol. 12, n° 8,

pp. 1482-1492. DOI 10.1093/gbe/evaa133.

WU, Yin-Qiao, ZHAO, Heng, LI, Ying-Ju, KHEDERZADEH, Saber, WEI, Hong-Jiang, ZHOU, Zhong-Yin et ZHANG, Ya-Ping, 2020. Genome-wide identification of imprinted genes in pigs and their different imprinting status compared with other mammals. *Zoological Research*. 18 novembre 2020. Vol. 41, n° 6, pp. 721-725. DOI 10.24272/j.issn.2095-8137.2020.072.

XIAO, Liang, ESTELLÉ, Jordi, KIILERICH, Pia, RAMAYO-CALDAS, Yulixis, XIA, Zhongkui, FENG, Qiang, LIANG, Suisha, PEDERSEN, Anni Øyan, KJELDSEN, Niels Jørgen, LIU, Chuan, MAGUIN, Emmanuelle, DORÉ, Joël, PONS, Nicolas, LE CHATELIER, Emmanuelle, PRIFTI, Edi, LI, Junhua, JIA, Huijue, LIU, Xin, XU, Xun, EHRLICH, Stanislav D., MADSEN, Lise, KRISTIANSSEN, Karsten, ROGEL-GAILLARD, Claire et WANG, Jun, 2016. A reference gene catalogue of the pig gut microbiome. *Nature Microbiology*. 19 septembre 2016. Vol. 1, pp. 16161. DOI 10.1038/nmicrobiol.2016.161.

YEN, N.-T., LIN, C.-S., JU, C.-C., WANG, S.-C. et HUANG, M.-C., 2007. Mitochondrial DNA polymorphism and determination of effects on reproductive trait in pigs. *Reproduction in Domestic Animals = Zuchthygiene*. août 2007. Vol. 42, n° 4, pp. 387-392. DOI 10.1111/j.1439-0531.2006.00797.x.

YONEKURA, Shinichi, OHATA, Masaki, TSUCHIYA, Megumi, TOKITA, Hitomi, MIZUSAWA, Moeko et TOKUTAKE, Yukako, 2019. Peg1/Mest, an imprinted gene, is involved in mammary gland maturation. *Journal of Cellular Physiology*. février 2019. Vol. 234, n° 2, pp. 1080-1087. DOI 10.1002/jcp.27219.

ZENG, Yang et CHEN, Taiping, 2019. DNA Methylation Reprogramming during Mammalian Development. *Genes*. 29 mars 2019. Vol. 10, n° 4, pp. 257. DOI 10.3390/genes10040257.

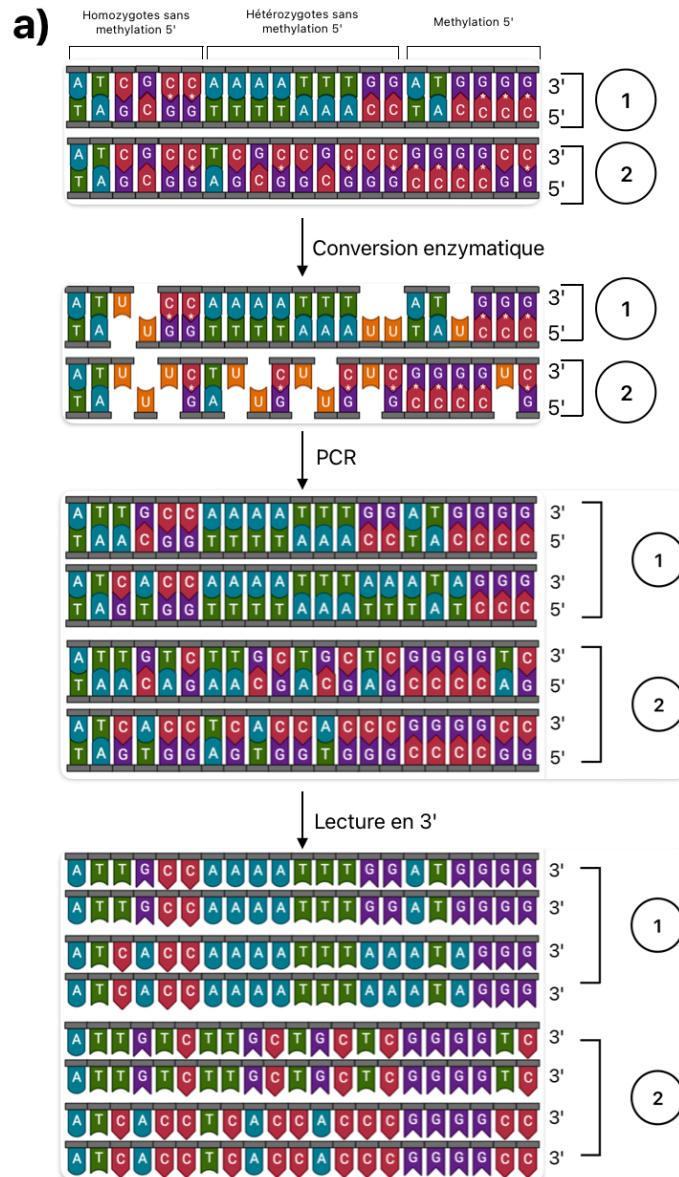
ZHANG, Jun, CHIODINI, Rod, BADR, Ahmed et ZHANG, Genfa, 2011. The impact of next-generation sequencing on genomics. *Journal of Genetics and Genomics = Yi Chuan Xue Bao*. 20 mars 2011. Vol. 38, n° 3, pp. 95-109. DOI 10.1016/j.jgg.2011.02.003.

ZHOU, Hao, YANG, Lingyu, DING, Jinmei, DAI, Ronghua, HE, Chuan, XU, Ke, LUO, Lingxiao, XIAO, Lu, ZHENG, Yuming, HAN, Chengxiao, AKINYEMI, Fisayo T., HONAKER, Christa F., ZHANG, Yan, SIEGEL, Paul B. et MENG, He, 2022. Intestinal Microbiota and Host Cooperate for Adaptation as a Hologenome. *mSystems*. 22 février 2022. Vol. 7, n° 1, pp. e0126121. DOI 10.1128/msystems.01261-21.

Annexe

Annexe 1 : Lien entre lectures de séquençage après conversion enzymatique et génotype

a) Conséquence de la conversion enzymatique (toutes les combinaisons étant représentées) ; b) Lien entre génotype et lectures de séquençage après conversion enzymatique.



b)

Génotype		Brin 3'	Brin 5'	A	T	G		C	C*
Brin 3'	Brin 5'	Brin 5'	T	A	C	C*	G	G	G
A	T	Séquençage	A	AT	A/G	A/G	AT/C	A/C	
		%	100%	50/50	75/25	50/50	50/25/25	50/50	
T	A	Séquençage		T	T/A/G	T/G	T/C	T/C	
		%		100%	50/25/25	50/50	75/25	50/50	
G	C	Séquençage			G/A	G/A	A/C/G/T	G/A/C	
		%			50/50	75/25	25/25/25/25	25/25/50	
C	C*	Séquençage				G	G/T/C	G/C	
		%				100%	50/25/25	50/50	
C	G	Séquençage					T/C	T/C	
		%					50/50	25/75	
C*	G	Séquençage						C	
		%						100%	

Signatures moléculaires des gènes soumis à empreinte parentale chez le porc

Chez les mammifères placentaires, un sous-ensemble de gènes est exprimé de façon mono-allélique et parent-dépendante. Ce phénomène héréditaire, régulé au travers de méthylations de l'ADN, est nommé empreinte génomique parentale. Moins étudié chez le porc que chez l'Homme ou la souris, il est pourtant associé à des caractères d'intérêt en filière porcine. Ce manuscrit propose une synthèse (i) de l'utilisation de l'information biologique pour l'amélioration des populations porcines d'élevage et (ii) des connaissances acquises sur l'empreinte parentale et son influence sur le phénotype. De plus, il rapporte un travail expérimental de caractérisation chez le porc de la région *H19/IGF2-KCNQ1/CDKN1C*, impliquée dans la croissance et la musculature des animaux. Une approche multi-omique originale montre ainsi que celle-ci héberge des signatures moléculaires de l'empreinte parentale et apporte des données précieuses pour comprendre plus finement les régulations dont elle fait l'objet.

Mots-clés : sélection génomique, épigénétique, séquençage de méthylome, séquençage d'ARN, croisement réciproque, région 11p15.

Molecular signatures of imprinted genes in pigs

In placental mammals, a subset of genes is expressed in a mono-allelic and parent-dependent manner. This hereditary phenomenon, regulated through DNA methylation, is called parental genomic imprinting. While less studied in pigs compared to humans or mice, it is nevertheless associated with important traits in the swine industry. This manuscript provides a synthesis of (i) the use of biological information for the improvement of commercial pig populations and (ii) the knowledge acquired on parental imprinting and its influence on phenotype. Additionally, it reports an experimental study characterizing the *H19/IGF2-KCNQ1/CDKN1C* region in pigs, which is involved in growth and muscular development. An original multi-omics approach demonstrates that this region harbors molecular signatures of parental imprinting and provides valuable data for a more refined understanding of the involved regulatory mechanisms.

Keywords : genomic selection, epigenetics, methylome sequencing, RNA sequencing, reciprocal crossing, 11p15 region.