

Mars 2025

Rapport final



Analyse du comportement alimentaire des porcs en croissance et de ses relations avec l'efficacité énergétique et les interactions sociales

Auteur (s)

Patrick Gagnon, Ph. D.

Laurence Maignel, M. Sc.

Luis Andrés Gonzalez-Gracia, candidat au Ph. D.

Anne Marthe Sophie Ngo-Bibinbe, candidate au Ph. D.

Jamie Ahloy-Dallaire, Ph. D.

Éric Paquet, Ph. D.

©Centre de développement du porc du Québec inc.
Dépôt légal 2025
Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada
ISBN 978-2-925175-29-2

Équipe de réalisation

Répondant	Patrick Gagnon, Ph. D., responsable analyse et valorisation des données, Centre de développement du porc du Québec (CDPQ)
Direction scientifique	Patrick Gagnon, Ph. D., responsable analyse et valorisation des données, CDPQ
Collaborateurs	Laurence Maignel, M. Sc., généticienne, Centre canadien pour l'amélioration des porcs (CCAP) Luis Andrés Gonzalez-Gracia, candidat au Ph. D., Université Laval Anne Marthe Sophie Ngo-Bibinbe, candidate au Ph. D., Université Laval Jamie Ahloy-Dallaire, Ph. D., professeur agrégé, Faculté des sciences de l'agriculture et de l'alimentation, Université Laval Éric Paquet, Ph. D., professeur agrégé, Faculté des sciences de l'agriculture et de l'alimentation, Université Laval

Remerciements

Ces travaux ont été réalisés grâce à une aide financière du Programme Innov'Action agroalimentaire, un programme issu de l'Accord Canada-Québec de mise en œuvre du Partenariat canadien pour l'agriculture conclu entre le ministère de l'Agriculture, des Pêcheries et de l'Alimentation et Agriculture et Agroalimentaire Canada. Le projet a également été réalisé grâce à la contribution financière d'Olymel, de PIC Canada, du Porc Notre-Dame, du Centre canadien pour l'amélioration des porcs et du Centre de développement du porc du Québec.



On nourrit le monde



Résumé

Les stations d'alimentation avec identification individuelle sont traditionnellement utilisées pour mesurer la quantité d'aliment consommée par jour à l'échelle individuelle, mais recueillent aussi l'information à chaque visite. Il est possible de calculer des indicateurs du comportement alimentaire individuel, mais aussi d'inférer le comportement social. Des séquences vidéo ont permis de déterminer que la grande majorité des déplacements détectés sur la base de ces données, soit un animal qui commence une visite immédiatement après celle d'un autre animal, était de nature agressive. Ainsi, des indicateurs sur la hiérarchie et plus généralement sur les interactions sociales ont pu être calculés à partir des séquences des visites aux stations. Le nombre et la fréquence de déplacements provoqués, comme agresseur, ou subis, comme agressé, étaient à la base des indicateurs sociaux calculés. Par exemple, le classement Elo, communément utilisé pour classer les joueurs d'échecs, est une des méthodes qui ont été utilisées pour classer les animaux d'un groupe (parc, enclos) selon leurs interactions présumées gagnées ou perdues. À l'échelle du groupe, l'analyse des réseaux sociaux a permis de calculer des indicateurs comme la densité du réseau (*network density*) qui indique le degré d'interactions entre les animaux. Tous ces indicateurs peuvent varier dans le temps et ces variations temporelles sont elles-mêmes des indicateurs précieux pour évaluer la vitesse d'établissement de la hiérarchie au sein d'un groupe.

Les stations d'alimentation avec identification individuelle sont surtout utilisées dans les fermes de sélection. Les candidats à la sélection sont principalement choisis en fonction de leur potentiel génétique individuel pour la vitesse de croissance et l'efficacité alimentaire. Comme les animaux sont élevés en groupe et que les interactions sociales peuvent avoir un impact sur les performances du groupe, des indicateurs sociaux développés dans le cadre de ce projet ont été analysés pour connaître leurs liens avec les critères classiques de comportement alimentaire, d'efficacité énergétique et de croissance. La plupart des critères de comportement alimentaire étudiés (ex. : nombre de visites quotidiennes, durée moyenne quotidienne passée à l'automate, vitesse d'ingestion moyenne) ont une héritabilité moyenne à forte (26 à 59 %), tandis que les critères de comportement social (ex. : nombre de déplacements provoqués ou subis à l'automate, proportion de jours en test au cours desquels l'animal provoque plus de déplacements qu'il n'en subit) ont une héritabilité faible à moyenne (7 à 37 %), ce qui indique qu'ils pourraient être améliorés génétiquement. Des indices de potentiel génétique (IPG) ont été calculés sur les caractères héréditaires pour les parents des animaux testés en 2023 dans un des troupeaux de sélection partenaires, afin de tester leur valeur prédictive mais aussi pour mesurer leur impact sur les performances de production classiques de la descendance.

Les enregistrements vidéo ont aussi été utilisés pour le développement d'une méthode de suivi individuel des porcs. Une approche novatrice, basée sur la théorie des chaînes de Markov cachées, combinant la vision numérique et l'information fournie par une station d'alimentation ou d'abreuvement, a été développée. Cette méthode permet de suivre les animaux et d'obtenir les identifiants individuels sur une plus longue période en comparaison avec les méthodes utilisant uniquement la vision numérique.

Table des matières

Liste des tableaux.....	iv
Liste des figures	v
1 Introduction	1
2 Objectif.....	2
3 Méthodologie.....	3
3.1 Données	3
3.2 Analyse du comportement alimentaire et social	4
3.2.1 Suivi des animaux via la vision numérique	4
3.2.2 Comportement alimentaire	5
3.2.3 Comportement social via les données des automates.....	5
3.3 Analyse de l'impact du comportement alimentaire et social sur les performances.....	7
3.4 Calcul des paramètres génétiques.....	8
3.5 Validation des indices sociaux	9
4 Résultats.....	11
4.1 Performance de l'outil de suivi des porcs.....	11
4.2 Analyses statistiques de l'impact des variables comportementales sur les critères de performances classiques.....	13
4.3 Paramètres génétiques des variables comportementales.....	15
4.4 Analyses statistiques de l'impact des indices génétiques sur les critères de performances classiques.....	16
5 Conclusion	21
Références	22

Liste des tableaux

Tableau 1. Données retenues pour chaque partenaire.....	3
Tableau 2. Liste des critères individuels de comportement alimentaire calculés à partir des données des automates d'alimentation.....	5
Tableau 3. Liste des variables de comportement social estimées à partir des données des automates d'alimentation.	7
Tableau 4. Liste des différentes méthodes évaluées et leurs performances respectives.....	11
Tableau 5. Impact des variables comportementales sur les performances individuelles des animaux du partenaire A.	13
Tableau 6. Impact des variables de comportement social sur les performances par parc des animaux du partenaire A.	14
Tableau 7. Paramètres génétiques des critères individuels de comportement alimentaire.	15
Tableau 8. Paramètres génétiques des critères individuels de comportement social.....	15
Tableau 9. Corrélations entre les IPGs estimés pour les parents et les variables de comportement observées chez leurs descendants durant la période de validation.	17
Tableau 10. Impact des IPGs comportementaux du père et de la mère sur les performances individuelles de production durant la période de validation.	17
Tableau 11. Impact des IPGs liés au comportement social du père et de la mère sur les performances par parc durant la période de validation.	18
Tableau 12. Corrélations génétiques et phénotypiques entre les caractères étudiés – Race Duroc ($n = 4480$ animaux)	25

Liste des figures

Figure 1. Exemple d'analyse de réseaux sociaux.....	6
Figure 2. Performances des différentes approches évaluées en fonction du temps dans la vidéo de 10 minutes.	12
Figure 3. L'évolution du score F1 en fonction du nombre d'identifications synthétiques à la station d'abreuvement..	13
Figure 4. Étendue de l'IPG de la mère pour la variable D+MQ pour les 16 parcs faibles et forts pour la CA 80-120 kg.....	19
Figure 5. Étendue de l'IPG du père pour la variable ELO pour les 16 parcs faibles et forts pour la CA 80-120 kg.	20

1 Introduction

La sélection des porcs pour une meilleure croissance et contre l'adiposité a permis d'améliorer l'efficacité alimentaire. L'utilisation d'automates qui distribuent et mesurent la consommation alimentaire individuelle s'est répandue dans les élevages de sélection au cours des dernières années. L'efficacité alimentaire globale ou par phase est utilisée comme critère de sélection, cependant il existe une marge de progrès importante à exploiter. La sélection des porcs est généralement réalisée sur les performances et les valeurs génétiques individuelles, alors que les animaux sont élevés en groupe et que les interactions sociales ont un impact sur les performances de l'ensemble du groupe. La sélection pour des animaux à fort potentiel de croissance, en particulier dans un contexte d'animaux de plus en plus lourds, est connue pour favoriser l'émergence d'animaux dominants qui doivent leurs bonnes performances à leur statut social dans le groupe. Les animaux dominants sont souvent parmi les moins efficaces car ils dépensent beaucoup d'énergie à déclencher des batailles pour maintenir leur statut. La compétition entre les porcs est source de stress, de blessures et de mortalités qui impactent le bien-être des animaux et se traduisent par des pertes économiques pour le producteur.

L'exploitation de données collectées par les automates d'alimentation peut fournir des informations sur le comportement alimentaire, mais aussi sur les interactions sociales des animaux élevés en groupe. L'analyse du comportement alimentaire et de ses composantes sociales permet de combiner les effets individuels et sociaux sur les performances de croissance, d'efficacité énergétique et de composition de carcasse. L'information fournie par les stations d'alimentation va bien au-delà des critères classiques utilisés par l'industrie et peut maintenant être exploitée. De même, la vision numérique connaît un essor important et peut permettre de valider ou même de bonifier l'analyse du comportement à partir de données d'alimentation. Les améliorations récentes des capacités d'analyses numériques ont permis une démocratisation de l'utilisation des outils d'analyse et de visualisation de données. Ces nouvelles approches sont prometteuses pour fournir de nouveaux critères de sélection permettant une amélioration conjointe des coûts de production et du bien-être animal.

Le projet réalisé visait à appliquer ces nouvelles approches pour analyser le comportement alimentaire et social des porcs en croissance et évaluer son impact sur les performances des animaux du même parc.

2 Objectif

L'objectif général du projet était d'explorer les relations entre croissance, efficacité alimentaire, comportement alimentaire et composition de la carcasse chez des porcs en croissance afin de développer de nouveaux critères de sélection s'appuyant sur le comportement alimentaire et social.

Dans cette optique, des données de stations d'alimentation provenant de différents élevages ont été utilisées pour développer des modèles de réseaux sociaux et définir des critères basés sur le comportement social. Des enregistrements vidéo étaient disponibles pour un sous-groupe d'animaux pour valider les interactions analysées aux stations d'alimentation et explorer des méthodes de suivi comportemental par vision numérique.

Des valeurs génétiques ont été calculées pour des critères héréditaires de comportement alimentaire et de comportement social. Des valeurs génétiques pour des animaux non utilisés pour le développement des modèles ont permis d'évaluer la capacité de prédiction de ces nouveaux critères sur les performances.

3 Méthodologie

3.1 Données

Les données provenaient de deux partenaires privés (A et B) et de la Station d'évaluation des porcs de Deschambault (CDPQ). Les animaux des partenaires A et B étaient des femelles et mâles entiers de race pure alors que les animaux au CDPQ étaient des castrats croisés (Tableau 1). Au cours de la phase d'engraissement, les animaux étaient placés dans des parcs de 12 à 14 animaux équipés d'automates d'alimentation pour une durée de 10 à 11 semaines (partenaire A et CDPQ) ou 4 à 8 semaines (partenaire B) afin de mesurer leur efficacité alimentaire sur toute leur phase de croissance (partenaire A) ou sur la fin de celle-ci (partenaire B). Les parcs étaient allotés par race et par sexe (partenaires A et B) et équipés d'un automate d'alimentation individuelle électronique de type FIRE® (Osborne Industries Inc.) ou IVOG® (Hokofarm group) selon les sites. Chaque animal portait une puce RFID à l'oreille, qui permettait sa détection dans les automates d'alimentation et la collecte de l'heure de début/fin et de la quantité d'aliment consommée à chaque visite dans l'automate. Les animaux étaient alimentés à volonté pendant toute la phase dans les automates.

Les animaux utilisés dans le projet avaient au moins deux pesées individuelles (une au début et une en fin de la période dans les automates d'alimentation) et les données de chaque visite à la station d'alimentation étaient disponibles pour chaque animal sur toute la période de test. Pour les animaux du CDPQ, un seul parc a été utilisé (15 animaux) pour le développement des preuves de concept pour l'analyse du comportement social autant par les données d'alimentation que par l'analyse des vidéos, mais aussi pour valider les interactions sociales détectées à l'automate avec des observations visuelles. Le logiciel BORIS (Friard et Gamba, 2016) a été utilisé pour annoter le début et la fin des visites observées sur les enregistrements vidéo et les comparer avec les données du système d'alimentation.

Tableau 1. Données retenues pour chaque partenaire.

Source	Nb sites	Nb animaux	Races	Sexes	Pesées individuelles	Alimentation individuelle	Abreuvement individuel	Enregistrements vidéo
Partenaire A	5	7772	Duroc Yorkshire Landrace	Mâles entiers Femelles	●	●		
Partenaire B	1	3566	Duroc	Mâles entiers Femelles	●	●		
CDPQ	1	15	Commerciaux	Castrats	●	●	●	●

Pour les animaux des partenaires A et B, le jeu de données disponible incluait plusieurs générations successives d'animaux contrôlés en ferme, ce qui rendait possible certaines analyses génétiques pour les critères calculés. Un contrôle de qualité a été effectué pour retirer les parcs avec trop de données d'alimentation manquantes. Toutes les données retenues du partenaire A et les données antérieures à 2023 pour le partenaire B ont été utilisées pour développer tous les modèles statistiques et génétiques. Les données du partenaire B collectées en 2023 ont été utilisées afin de valider les modèles.

3.2 Analyse du comportement alimentaire et social

3.2.1 Suivi des animaux via la vision numérique

La vision numérique offre la possibilité de suivre les animaux, et leurs interactions en temps réel, partout dans le parc, contrairement aux stations d'alimentation. Avant de s'attaquer à la détection d'interactions, comme des agressions, en tenant compte des identifiants individuels, il faut d'abord être en mesure de suivre individuellement des porcs sans perdre leur trace. Or, malgré les avancées récentes en vision numérique, le suivi individuel des animaux en continu est une tâche ardue. En fait, si la seule source d'information est une caméra surplombant le parc, l'identité d'un animal sera inévitablement perdue au bout d'un certain temps, même si toutes les identités sont connues au départ. Les entassements des animaux, les occlusions et les déplacements erratiques des animaux sont parmi les causes des pertes d'identité.

Pour assurer un suivi individuel à long terme, il est nécessaire d'utiliser une source d'information indépendante qui confirme ou corrige ponctuellement les identités proposées par le système de vision. Pour ce projet, une approche novatrice intégrant le suivi par vision numérique et les identifications au système d'abreuvement individuel a été développée. L'approche permet de corriger les séquences vidéo précédant et suivant une visite tout en tenant compte de l'incertitude des identifications au système d'abreuvement (e.g., lorsque plusieurs animaux se trouvent à proximité de l'antenne du système d'abreuvement). La probabilité qu'un animal détecté au point (x,y) au temps t soit l'animal i dépend notamment de la position des animaux détectés au temps $t-1$ et des visites antérieures et postérieures à la station d'abreuvement. Il en résulte une série d'équations impliquant des probabilités conditionnelles impossible à résoudre analytiquement. C'est pourquoi l'approche proposée utilise la théorie des chaînes de Markov cachées (HMM). Le détail des équations utilisées et de l'implantation numérique de l'approche est présenté dans Ngo-Bibinbe *et al.* (2024).

La preuve de concept a été développée sur une séquence vidéo de 10 minutes à la station de Deschambault à 25 images par seconde dont 780 avec annotations sur l'identité et l'emplacement de chacun des animaux. Trois approches ont été comparées, toutes utilisant le modèle de suivi par détection représentant l'état de l'art ByteTrack (Zhang *et al.*, 2021) : (i) le modèle ByteTrack seul, (ii) ByteTrack avec une approche de réidentification plus classique (ByteTrack+Re-ID) et (iii) ByteTrack avec l'approche HMM proposée (ByteTrack+HMM). Brièvement, ByteTrack est une méthode de suivi multi-objets (*multi-object tracking*) qui identifie

les objets à suivre dans un premier temps et utilise l'identité de ces objets dans le temps pour faire le suivi. L'approche HMM que nous proposons pourrait en théorie être appliquée à d'autres modèles de suivi par détection que ByteTrack. Nous avons choisi ByteTrack parce que c'est un des modèles les plus robustes actuellement.

3.2.2 Comportement alimentaire

Plusieurs indicateurs du comportement ont été calculés uniquement à partir des données brutes des systèmes d'alimentation (Tableau 2). La plupart de ces indicateurs ont été calculés à l'échelle individuelle, indépendamment pour chaque animal, mais peuvent tout de même être révélateurs du comportement social, tels que, par exemple, la proportion des visites hors périodes de pointe.

Tableau 2. Liste des critères individuels de comportement alimentaire calculés à partir des données des automates d'alimentation.

Variable	Définition	Unité
CMQ	Consommation moyenne quotidienne	kg/j
VMQ	Nombre moyen quotidien de visites à la station d'alimentation	visites/j
DMQ	Temps moyen quotidien passé à la station d'alimentation	min/j
DMV	Durée moyenne par visite = DMQ/VMQ	min
CMV	Consommation moyenne par visite = CMQ/VMQ	kg
VIM	Vitesse d'ingestion moyenne = CMQ/DMQ	kg/min
VMQP	VMQ durant la période de pointe ¹	visites/j
VMQO	VMQ en dehors de la période de pointe ¹ = VMQ - VMQP	visites/j
CMQP	CMQ durant la période de pointe ¹	kg/j
CMQO	CMQ en dehors de la période de pointe ¹ = CMQ - CMQP	kg/j
PVMP	Proportion des visites durant la période de pointe ¹ = VMQP/VMQ	%
PVMO	Proportion des visites en dehors de la période de pointe ¹ = VMQO/VMQ	%
PCMP	Proportion de la consommation durant la période de pointe ¹ = CMQP/CMQ	%
PCMO	Proportion de la consommation en dehors de la période de pointe ¹ = CMQO/CMQ	%

¹La période de pointe est définie de 8h à 20h.

3.2.3 Comportement social via les données des automates

D'autres indicateurs, également dérivés des données d'automates, sont basés sur les interactions entre les animaux lors des repas (Tableau 3). En effet, la séquence des visites à l'automate et leurs caractéristiques temporelles (durée des visites, écart entre les visites de deux animaux différents, etc.) permettent de prédire certaines interactions sociales entre les animaux utilisant les automates d'alimentation au même moment de la journée. Dans chaque parc, l'automate

d'alimentation est l'un des points clefs où les animaux interagissent pour accéder à une ressource essentielle (l'aliment) et c'est également là que s'établit la hiérarchie à l'intérieur du groupe. Plus spécifiquement, il est possible d'identifier les événements de « déplacements » à l'automate, lorsqu'un animal en déloge un autre pour pouvoir visiter l'automate à son tour. Il est ainsi possible, à partir des données de visites élémentaires, de quantifier toutes les interactions de déplacement entre les porcs d'un même parc et d'identifier pour chaque événement l'animal « déplaceur » et l'animal « déplacé ». Les interactions de ce type sont plus fréquentes en début d'engraissement, lorsque les animaux viennent d'être regroupés et que la hiérarchie n'est pas encore établie au sein du groupe. Elles sont plus ou moins intenses et peuvent mener à des comportements agressifs parfois problématiques. Une fois la hiérarchie établie, le passage des animaux à l'automate donne moins lieu à des déplacements.

L'analyse des réseaux sociaux (*social network analysis*) permet de quantifier les interactions au sein d'un groupe au cours du temps et de calculer un certain nombre de critères individuels et pour le groupe (parc) (Figure 1). La densité du réseau (*network density*) a été calculée à chaque jour et est basée sur le nombre de connexions. Deux animaux d'un parc sont dits *connectés* une journée donnée s'ils ont eu au moins une interaction de type déplacement à la station d'alimentation. La densité du réseau pour une journée est le nombre de paires d'animaux connectés divisé par le nombre total de paires d'animaux, soit $n \times (n-1)/2$ où n est le nombre d'animaux dans le parc (Gonzalez-Gracia *et al.*, 2023a,b).

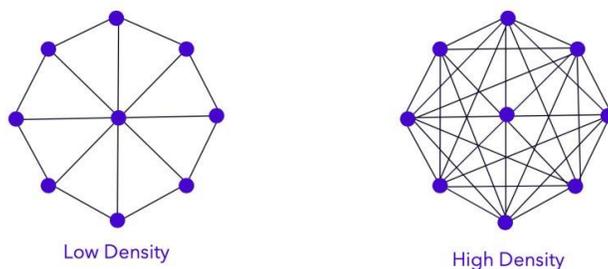


Figure 1. Exemple d'analyse de réseaux sociaux. Chaque ligne représente deux animaux (points) connectés. Pour le schéma de gauche, un total de 16 connexions relie les 9 animaux (densité du réseau = 16 connexions/36 connexions possibles = 0,44). Le schéma de droite représente un groupe pleinement connecté (densité du réseau = 1).

Les données utilisées dans l'analyse des réseaux sociaux provenaient de 649 groupes dans des parcs équipés de stations d'alimentation automatique provenant de cinq fermes du partenaire A, totalisant 7772 porcs, engraisés dans des groupes constitués de porcs mâles ou femelles Duroc, ou de mâles Landrace ou Yorkshire. Les événements de déplacements agressifs ont été détectés dans l'automate (*feder-detected aggressive displacements* - FDAD) en fonction de la proximité temporelle entre la sortie d'un animal et le suivant entrant dans l'automate, la

justification étant que des intervalles de temps très courts entre les visites (<2 secondes) reflètent des déplacements agressifs probables. En parallèle, un sous-ensemble de parcs utilisant des alimenteurs automatiques a été équipé de caméras de surveillance pour permettre une validation ultérieure de l'algorithme de détection d'agression en comparant ses résultats à l'observation visuelle (Gonzalez-Gracia *et al.*, 2024).

En plus de la densité du réseau, le classement Elo a aussi été utilisé (Elo, 1967). Brièvement, le classement Elo, donne un score à chaque animal basé sur ses interactions gagnées et perdues selon la force relative de ses « adversaires ». Cet indicateur est utilisé notamment pour le classement des joueurs d'échecs et a été adopté dans les études du comportement social des animaux.

Tableau 3. Liste des variables de comportement social estimées à partir des données des automates d'alimentation.

Variable	Définition	Unité	Niveau
D+MQ	Nombre moyen quotidien de déplacements provoqués à la station d'alimentation	Déplacements/jour	Animal
D-MQ	Nombre moyen quotidien de déplacements subis à la station d'alimentation	Déplacements/jour	Animal
PD+	Proportion des déplacements à la station d'alimentation qui sont provoqués = $D+MQ/(D+MQ + D-MQ)$	%	Animal
PJD+	Proportion des jours où l'animal a provoqué plus de déplacements qu'il en a subis	%	Animal
ELO	Classement Elo à la fin de l'essai	Pts Elo	Animal
NDM	Densité du réseau (<i>network density</i>) moyenne durant le lot	Proportion (entre 0 et 1)	Parc
NDP	Pente linéaire ajustée pour l'évolution quotidienne de la densité du réseau	Δ proportion/jour	Parc
NDR2	Coefficient de détermination (R ²) pour le modèle linéaire ajusté sur la densité quotidienne du réseau (indication de la stabilité des interactions sociales en cours de lot)	Proportion (entre 0 et 1)	Parc

3.3 Analyse de l'impact du comportement alimentaire et social sur les performances

Les principaux critères de performance pour les porcs en croissance sont le gain moyen quotidien (GMQ) et la conversion alimentaire (CA), qui est la quantité d'aliment (kg) nécessaire pour obtenir 1 kg de gain de poids vif. Des analyses de la variance (ANOVA) ont été réalisées pour

quantifier l'impact des variables de comportement alimentaire (Tableau 2) et social (Tableau 3) sur les performances individuelles et par parc. Seules les données du partenaire A ont été utilisées à cette étape.

Le modèle de base pour les données individuelles est donné par :

$$Y_{i,j,a,b,c,k}^{30-120} = \mu + (\alpha \times \beta)_{ij} + IPG_k + A_a + B_{b(a)} + C_{c(b)} + \varepsilon_{i,j,a,b,c,k} \quad \text{Équation 1}$$

où $Y_{i,j,a,b,c,k}^{30-120}$ est l'indicateur de performance (GMQ ou CA) ajusté pour l'intervalle 30-120 kg, μ est la moyenne globale, α_i est l'effet de la race i (Duroc, Landrace ou Yorkshire), β_j est l'effet du sexe j (mâle entier pour les trois races ou femelle uniquement pour le Duroc), IPG_k est l'indice de potentiel génétique (IPG) de l'animal k pour l'indicateur de performance (GMQ naissance-120 kg ou CA 30-120 kg), A_a est l'effet aléatoire du bâtiment a , B_b est l'effet aléatoire du lot b (emboîté dans le bâtiment), C_c est l'effet aléatoire du parc c (emboîté dans le lot) et $\varepsilon_{i,j,a,b,c,k}$ est le terme aléatoire pour la performance de l'individu k . Chaque variable de comportement alimentaire et social à l'échelle de l'animal a été ajoutée à ce modèle.

Le modèle de base pour les données par parc est donné par :

$$\bar{Y}_{i,j,a,b,c}^{30-120} = \mu + (\alpha \times \beta)_{ij} + \overline{IPG}_c + A_a + B_{b(a)} + \varepsilon_{i,j,a,b,c} \quad \text{Équation 2}$$

où $\bar{Y}_{i,j,a,b,c}^{30-120}$ est la moyenne pour le parc c de l'indicateur de performance (GMQ ou CA) ajusté pour l'intervalle 30-120 kg, \overline{IPG}_c est la moyenne de l'IPG pour l'indicateur de performance des animaux de parc c et $\varepsilon_{i,j,a,b,c}$ est le terme aléatoire pour la performance du parc c . Les analyses à l'échelle du parc se sont concentrées sur les variables de comportement social (Tableau 3). Pour chaque variable analysée, la moyenne par parc ainsi que d'autres statistiques telles que l'écart-type, l'écart interquartile (EIQ) et l'étendue ont été ajoutées au modèle de base par parc.

Les analyses ont été réalisées avec les bibliothèques lme4 version 1.1-35.1 (Bates *et al.*, 2015) et lmerTest version 3.1-3 (Kuznetsova *et al.*, 2017) du logiciel R version 4.2.2 (R Core Team, 2022). Pour l'inférence statistique, il est supposé que les termes ε sont indépendants et identiquement distribués selon la distribution Normale centrée à 0. Une transformation monotone (*e.g.*, racine carrée, logarithmique) est appliquée sur la variable Y si l'hypothèse n'est pas respectée.

3.4 Calcul des paramètres génétiques

Les données disponibles dans ce projet se prêtent bien à des calculs de paramètres génétiques, en particulier celles du partenaire A, car elles ont été collectées sur plusieurs générations successives d'animaux de race pure, pour lesquels les généalogies sont bien connues. L'estimation de paramètres tels que l'héritabilité permet de savoir si un caractère pourrait être facilement sélectionné, et le cas échéant, quel serait l'impact sur les autres caractères d'intérêt (via les corrélations génétiques). L'héritabilité d'un caractère donné est la proportion de la variance de ce caractère qui est d'origine génétique dans une population donnée. Dans ce projet, il a été possible de calculer les paramètres génétiques pour plusieurs critères de comportement

alimentaire et de les comparer avec les résultats publiés sur d'autres populations, et pour de nouveaux critères individuels de comportement social développés dans le cadre de ce projet.

Les héritabilités ont été estimées à l'aide du logiciel VCE 5.1.2 (Kovac *et al.*, 2002) en utilisant un modèle animal monocaractère incluant les effets fixes du groupe contemporain (combinaison troupeau-année-saison) et du sexe et les effets aléatoires du parc (intra-groupe contemporain), de la portée de naissance et de la valeur génétique additive de l'animal. Les effets du parc (p), de la portée de naissance (q) et de l'animal (a) et la résiduelle (e) sont supposés être indépendants et suivre une distribution normale avec les paramètres respectifs suivants: $p \sim N(0, I\sigma_p^2)$, $q \sim N(0, I\sigma_q^2)$, $a \sim N(0, A\sigma_a^2)$ et $e \sim N(0, I\sigma_e^2)$, où σ_p^2 , σ_q^2 , σ_a^2 et σ_e^2 sont les variances d'effets du parc, de la portée de naissance, de la valeur génétique additive et de l'effet résiduel, respectivement, I est la matrice identité et A est la matrice de parenté entre tous les animaux phénotypés.

3.5 Validation des indices sociaux

Le potentiel d'amélioration génétique d'un caractère dépend en grande partie de son niveau d'héritabilité. Cependant, l'amélioration du caractère et son impact sur les autres caractéristiques des animaux est plus difficile à prédire. Pour valider certains des nouveaux critères développés dans ce projet, des indices de potentiel génétique (IPG) ont été estimés pour les animaux du partenaire B testés jusqu'en 2022, en utilisant une approche BLUP-modèle animal incluant les mêmes effets de milieu et les composantes de variance décrites à la section 3.4. Les IPGs sont des indicateurs de la supériorité ou de l'infériorité génétique d'un animal par rapport à la population évaluée. Parmi ces animaux se trouvaient les parents des animaux testés en 2023. Pour savoir si les IPGs des parents sont de bons prédicteurs des performances de leurs descendants pour les critères de comportement alimentaire et social, une première analyse statistique des performances des animaux du partenaire B testés en 2023 (donc non inclus dans le calcul des valeurs génétiques) a été réalisée, à l'aide du modèle suivant :

$$Y_{j,b,c,k}^{80-120} = \mu + \beta_j + B_{b(a)} + IPG_k + IPGp_k + IPGm_k + \varepsilon_{j,b,c,k} \quad \text{Équation 3}$$

où $IPGp_k$ et $IPGm_k$ sont l'IPG du père et de la mère respectivement pour un critère de comportement alimentaire ou social donné. Les autres variables ont été définies à l'équation 1; les termes α et A ont été retirés puisqu'il n'y a qu'une seule race et un seul bâtiment dans le jeu de données du partenaire B. L'intervalle de standardisation de performance est maintenant 80-120 kg pour mieux refléter la période de croissance suivie chez le partenaire B.

On peut supposer que dans le cas du comportement social, des phénomènes non additifs sont en jeu, par exemple la dynamique d'un groupe peut varier grandement selon la composition de ce groupe, menant par exemple à l'exacerbation de comportements agressifs si plusieurs animaux dominants sont regroupés dans le même parc. L'analyse des performances par parc est dans ce contexte très pertinente pour valider l'approche. Le modèle utilisé à l'échelle du parc est le suivant :

$$\bar{Y}_{j,b,c}^{80-120} = \mu + \beta_j + B_{b(a)} + \overline{IPG}_c + \widetilde{IPGp}_c + \widetilde{IPGm}_c + \varepsilon_{j,b,c} \quad \text{Équation 4}$$

où \widetilde{IPGp}_c et \widetilde{IPGm}_c sont une statistique (la moyenne, l'écart-type, l'EIQ ou l'étendue), calculée sur les animaux du parc c , des IPGs des pères et des mères respectivement pour un critère de comportement social donné. Les autres variables ont été définies à l'équation 2.

4 Résultats

4.1 Performance de l'outil de suivi des porcs

Nous avons développé une méthode permettant de combiner les informations sur l'identité des animaux provenant des stations d'abreuvement avec les vidéos provenant des caméras pour améliorer les suivis des porcs à long terme avec la vision numérique. Grâce à ces informations combinées, nous sommes en mesure, en utilisant les chaînes de Markov cachées, d'assigner et de propager l'information pour obtenir l'identification des animaux sur de plus longues périodes.

Les performances des différentes approches de suivi des porcs (ByteTrack, ByteTrack+Re-ID, ByteTrack+HMM) ont été évaluées sur une vidéo de 10 minutes que nous avons annotée pour obtenir la localisation des porcs ainsi que leurs associations entre 2 images consécutives dans la séquence vidéo. Brièvement, l'approche ByteTrack seule est l'évaluation de référence puisqu'elle consiste en l'utilisation de l'approche ByteTrack directement sans modification (Zhang *et al.* 2021). À noter, pour l'utilisation de cette méthode, nous avons dû développer un modèle d'identification des porcs dans une image. Pour ce faire, nous avons entraîné un modèle YOLOX sur 11 700 annotations de porc. L'approche ByteTrack+ReID est l'approche ByteTrack qu'on combine à une approche de réassignation des identités des porcs en tenant compte de l'identité et finalement l'approche ByteTrack+HMM qui combine ByteTrack avec notre méthode HMM utilisant les informations des stations d'abreuvement.

Pour l'évaluation nous utilisons différentes métriques comme la précision sur le suivi multi-objets (MOTA), le score IDF1, la précision (Pr), le rappel (Ra) et le score F1 (Psota *et al.*, 2020 ; Bernardin et Stiefelhagen, 2008). Pour toutes ces métriques qui varient entre 0 et 1, une valeur de 0 est associée à de mauvaises performances et une valeur de 1 à des performances parfaites. Il est important de préciser que ces métriques évaluent deux éléments. Dans un premier temps la méthode doit identifier les porcs dans l'image et dans un second temps elle doit suivre et associer les porcs entre deux images consécutives. Pour toutes les métriques utilisées, la méthode ByteTrack+HMM obtient des performances supérieures aux deux autres approches (Tableau 4).

Tableau 4. Liste des différentes méthodes évaluées et leurs performances respectives. Précision sur le suivi multi-objets (MOTA), le score IDF1, précision (Pr), rappel (Ra) et le F1 score.

Méthode	MOTA	IDF1	Pr	Ra	F1
ByteTrack	0,21	0,59	0,52	0,50	0,51
ByteTrack-ReID	0,24	0,59	0,51	0,51	0,51
ByteTrack+HMM	0,42	0,61	0,64	0,56	0,60

L'objectif de la nouvelle méthode proposée est l'amélioration des performances de suivi à long-terme et pour évaluer ces performances, nous avons étudié l'évolution des différentes métriques en fonction du temps dans la vidéo de 10 minutes (Figure 2). La Figure 2 montre que pour toutes les métriques les performances de notre approche sont supérieures aux deux autres après 7 minutes (environ 10 000 images).

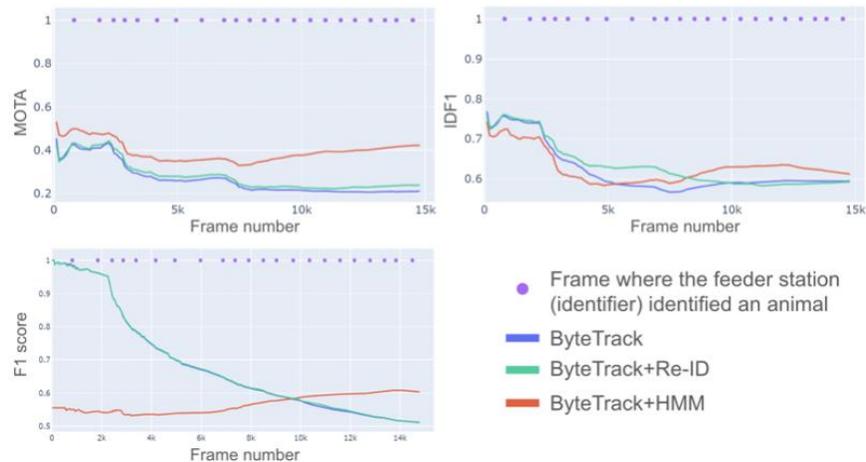


Figure 2. Performances des différentes approches évaluées en fonction du temps dans la vidéo de 10 minutes. Les différentes méthodes sont identifiées avec différentes couleurs. ByteTrack en bleu, ByteTrack+ReID en vert et ByteTrack+HMM, notre méthode, en rouge. Les points de couleur magenta identifient les moments des visites à la station d’abreuvement utilisés pour l’identification des porcs. Métriques utilisées : précision sur le suivi multi-objets (MOTA), le score IDF1 et le F1 score.

Finalement, nous nous sommes aussi intéressés à l’évolution des performances de notre approche en fonction du nombre d’identifications à la station d’abreuvement (Figure 3). Pour ce faire nous avons généré des séries de données synthétiques dans lesquelles nous avons ajouté un nombre croissant de visites à la station d’abreuvement. Chacune de ces « visites » artificielles étaient générées en choisissant aléatoirement une des annotations d’un porc à un endroit donné à un temps donné. Nous avons de plus attribué la mauvaise identité au porc sur 25 % de ces visites artificielles, afin de simuler l’incertitude des identifications à la station d’alimentation. On peut constater que plus il y a de visites à la station, plus notre approche est performante, ce qui n’est pas le cas de l’approche ByteTrack+ReID.

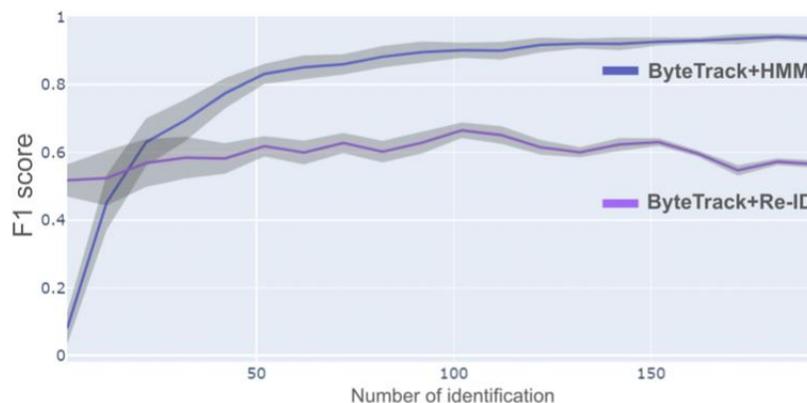


Figure 3. L'évolution du score F1 en fonction du nombre d'identifications synthétiques à la station d'abreuvement. Les performances sont présentées pour notre méthode ByteTrack+HMM en bleu et ByteTrack+ReID en magenta. La moyenne et le minimum et maximum sont présentés.

En résumé, l'approche ByteTrack+HMM que nous proposons possède des performances intéressantes pour le suivi à long-terme des porcs et pourra être utilisée pour le suivi des interactions des porcs dans de futurs projets pour permettre l'identification de comportements problématiques.

4.2 Analyses statistiques de l'impact des variables comportementales sur les critères de performances classiques

Les performances individuelles, soit le GMQ et la CA, dépendent du comportement alimentaire (Tableau 5). Comme la consommation d'un animal a un impact direct sur sa croissance, l'impact majeur de la CMQ sur les variables de performance était attendu. La plupart des autres variables de comportement alimentaire ont un impact statistiquement significatif (*i.e.* valeur $p < 0,05$). Par exemple, les animaux avec une DMV, CMV ou VIM supérieure ont en moyenne un meilleur GMQ et une meilleure efficacité alimentaire (*i.e.* une plus faible CA).

Tableau 5. Impact des variables comportementales sur les performances individuelles des animaux du partenaire A. Le fond en gris fait ressortir les variables de comportement social.

Variable ¹	GMQ 30-120 kg			CA 30-120 kg ³		
	Valeur <i>t</i>	Valeur <i>p</i>	Δ (g/j) ²	Valeur <i>t</i>	Valeur <i>p</i>	Δ (-) ²
CMQ 30-120 kg	77,87	< 0,001	0,180	9,04	< 0,001	0,075
VMQ	-3,48	0,001	-0,012	9,56	< 0,001	0,070
DMQ	6,78	< 0,001	0,020	3,60	< 0,001	0,047
DMV	7,13	< 0,001	0,027	-5,67	< 0,001	-0,047
CMV	24,12	< 0,001	0,077	-6,80	< 0,001	-0,063
VIM	23,63	< 0,001	0,081	-2,45	0,014	-0,030
PVMP	-2,11	0,035	-0,009	5,87	< 0,001	0,052
PCMP	-0,50	0,619	-0,003	0,75	0,454	0,009
D+MQ	-2,67	0,008	-0,009	7,31	< 0,001	0,070
D-MQ	-6,06	< 0,001	-0,020	7,04	< 0,001	0,077
PD+	4,42	< 0,001	0,018	-1,09	0,276	-0,031
PJD+	1,91	0,056	0,002	2,27	0,023	0,030
ELO	2,63	0,009	0,009	-0,90	0,370	-0,004

¹Variable de comportement (Tableaux 2 et 3) ajoutée au modèle de base (Équation 1).

²Différence entre la moyenne des groupes 25 % supérieur et inférieur pour la variable analysée. Les groupes ont été définis par race et par sexe.

³Une transformation logarithmique a été appliquée sur la CA pour les tests statistiques. Les différences entre groupes (Δ) ont été calculées sur la variable non transformée.

Les variables de comportement social D+MQ et D-MQ ont un impact statistiquement significatif sur les deux variables de performance ($p < 0,008$), suggérant que les animaux avec beaucoup d'interactions, qu'elles soient provoquées ou subies, ont un moins bon GMQ et surtout une moins bonne CA. Les autres variables de comportement social (PD+, PJD+, ELO) ont moins d'impact sur les performances individuelles (Tableau 5). Seule la variable PJD+ a à la fois un impact statistiquement significatif et une différence entre les groupes numériquement importante pour la CA. Les résultats suggèrent qu'un animal qui provoque plus de déplacements qu'il en subit sur plusieurs jours aurait en moyenne une moins bonne efficacité alimentaire.

À l'échelle du parc, les seules variables de comportement social ayant un impact statistiquement significatif sur le GMQ ajusté sont l'écart-type du classement Elo ($p = 0,029$), l'étendue du classement Elo ($p = 0,031$) et le coefficient de détermination de la densité du réseau ($p < 0,001$; Tableau 6). Les différences entre les groupes, en g/j, sont cependant relativement faibles, mais les résultats suggèrent qu'une stabilité dans la hiérarchie d'un parc est favorable pour le GMQ. Pour la CA ajustée, les seuls résultats statistiquement significatifs sont avec l'écart-type ($p = 0,045$) et l'étendue ($p = 0,026$) de la variable D-MQ.

Tableau 6. Impact des variables de comportement social sur les performances par parc des animaux du partenaire A.

Variable ¹	GMQ 30-120 kg ³			CA 30-120 kg		
	Valeur <i>t</i>	Valeur <i>p</i>	Δ (g/j) ²	Valeur <i>t</i>	Valeur <i>p</i>	Δ (-) ²
moyenne(D+MQ)	-1,08	0,283	-0,017	0,43	0,671	0,072
écart-type(D+MQ)	-0,63	0,529	-0,008	1,02	0,308	0,039
EIQ(D+MQ)	-1,78	0,075	-0,013	0,46	0,647	0,041
étendue(D+MQ)	-0,29	0,771	-0,007	1,12	0,263	0,048
moyenne(D-MQ)	-1,11	0,268	-0,018	0,48	0,629	0,073
écart-type(D-MQ)	-1,12	0,263	-0,010	2,01	0,045	0,072
EIQ(D-MQ)	-1,71	0,087	-0,015	0,97	0,330	0,059
étendue(D-MQ)	-0,97	0,331	-0,008	2,24	0,026	0,073
moyenne(PD+)	-0,07	0,947	0,016	0,48	0,634	-0,051
écart-type(PD+)	1,05	0,296	0,010	1,09	0,276	-0,042
EIQ(PD+)	0,14	0,886	0,007	1,02	0,310	-0,026
étendue(PD+)	1,42	0,156	0,018	1,47	0,143	-0,044
moyenne(PJD+)	-1,10	0,271	-0,020	1,43	0,152	0,082
écart-type(PJD+)	-1,15	0,253	-0,003	-0,46	0,644	-0,005
EIQ(PJD+)	-0,26	0,795	-0,003	0,95	0,340	-0,002
étendue(PJD+)	-0,58	0,565	0,002	0,01	0,995	-0,004
écart-type(ELO)	2,19	0,029	0,013	-0,37	0,712	-0,052
EIQ(ELO)	1,31	0,190	0,010	-0,66	0,507	-0,049
étendue(ELO)	2,16	0,031	0,011	0,00	0,999	-0,045
NDP	-1,79	0,074	0,000	-0,02	0,981	-0,031
NDM	-1,33	0,185	-0,018	0,16	0,871	0,064
NDR2	3,60	< 0,001	0,014	-0,21	0,837	0,014

¹Variable de comportement social (Tableau 3) ajoutée au modèle de base (Équation 2).

²Différence entre la moyenne des groupes 25 % supérieur et inférieur pour la variable analysée. Les groupes ont été définis par race et par sexe.

³Une transformation racine carrée a été appliquée sur le GMQ pour les tests statistiques. Les différences entre groupes (Δ) ont été calculées sur la variable non transformée.

4.3 Paramètres génétiques des variables comportementales

Pour chacun des caractères étudiés, les composantes de la variance phénotypique, incluant la variance génétique additive et la variance de l'effet de la portée de naissance, ont été estimées. Les rapports de variances correspondants sont listés aux tableaux 7 (comportement alimentaire) et 8 (comportement social) pour les trois races étudiées.

Tableau 7. Paramètres génétiques des critères individuels de comportement alimentaire.

Variable	Yorkshire (n=1636)		Landrace (n=1639)		Duroc (n=4480)	
	h ²	c ²	h ²	c ²	h ²	c ²
CMQ	0,52	0,04	0,44	0,08	0,21	0,10
VMQ	0,33	0,15	0,39	0,11	0,49	0,01
DMQ	0,57	0,06	0,59	0,05	0,56	0,05
DMV	0,26	0,23	0,53	0,08	0,47	0,07
CMV	0,41	0,17	0,40	0,10	0,34	0,11
VIM	0,56	0,04	0,52	0,07	0,48	0,10
PVMP	0,20	0,09	0,21	0,06	0,31	0,09
PCMP	0,26	0,08	0,28	0,06	0,37	0,09

h² : hérabilité, proportion de la variance phénotypique d'origine génétique additive;
c² : proportion de la variance phénotypique liée à l'environnement commun de la portée de naissance
Les erreurs standards des hérabilités estimées varient de 0.02 à 0.06.

L'hérabilité estimée pour les critères de comportement alimentaire est moyenne à élevée et varie d'un caractère à l'autre et d'une race à l'autre. Ceci est en accord avec différentes études réalisées sur d'autres populations (Chen *et al.*, 2000; Rohrer *et al.*, 2013). Pour la plupart des caractères, les races maternelles (Yorkshire et Landrace) ont des résultats proches, en comparaison de la race Duroc. Les caractères les plus hérables toutes races confondues sont la DMQ et la VIM, avec des hérabilités proches de 50 %. La plupart des caractères sont influencés par un effet significatif de la portée de naissance.

Tableau 8. Paramètres génétiques des critères individuels de comportement social.

Variable	Yorkshire (n=1636)		Landrace (n=1639)		Duroc (n=4480)	
	h ²	c ²	h ²	c ²	h ²	c ²
D+MQ	0,23	0,12	0,16	0,12	0,26	0,02
D-MQ	0,28	0,09	0,07	0,09	0,20	0,03
PD+	0,21	0,10	0,36	0,03	0,38	0,02

PJD+	0,28	0,10	0,19	0,07	0,37	0,02
ELO	0,18	0,10	0,22	0,08	0,31	0,03

h^2 : hérabilité, proportion de la variance phénotypique d'origine génétique additive;
 c^2 : proportion de la variance phénotypique liée à l'environnement commun de la portée de naissance;
 Les erreurs standards des hérabilités estimées varient de 0.02 à 0.06.

Les critères de comportement social étudiés ont des niveaux d'hérabilité plus faibles que les critères de comportement alimentaire (7 à 38 %) et sont pour la plupart affectés par des effets d'environnement commun de la portée de naissance, surtout chez les races maternelles. Ceci indique d'une part qu'il existe un déterminisme génétique transmissible pour ces nouveaux critères et qu'il est théoriquement possible de sélectionner les géniteurs pour améliorer ces critères dans les races porcines étudiées.

Une analyse des corrélations génétiques et phénotypiques (présentée à l'annexe 1 pour la race Duroc) montre peu de liens entre les performances de production, les critères de comportement alimentaire et les critères de comportement social. On note cependant quelques tendances intéressantes au niveau des corrélations génétiques estimées :

- Un lien faible mais positif (défavorable) entre la conversion alimentaire d'une part, et certains critères de comportement social (PD+, PJD+, ELO), suggérant qu'une partie de l'énergie dépensée par les animaux dominants pourrait l'être dans les interactions sociales aux dépens de la conversion alimentaire. Ceci est en accord avec les résultats rapportés à la section 4.2.
- Un lien positif significatif entre la vitesse d'ingestion et les variables PD+, PJD+, ELO, indiquant que les animaux les plus dominants sont aussi ceux qui mangent le plus vite.
- Un lien positif entre la proportion de visites et d'aliment consommé durant la période de pointe et les variables PD+, PJD+ et ELO.
- Les caractères PD+, PJD+ et ELO sont hautement corrélés au niveau génétique. Pour une utilisation en sélection, le critère PD+ est le plus héritable des trois.

4.4 Analyses statistiques de l'impact des indices génétiques sur les critères de performances classiques

Les IPGs calculés pour les parents de chaque animal inclus dans les groupes de validation (2023) ont été comparés avec les performances observées pour ces mêmes critères dans ces groupes (Tableau 9). Les corrélations sont généralement plus grandes avec les IPGs de la mère, surtout pour les variables de comportement social PD+, PJD+ et ELO. Ceci indique que même si ces caractères sont relativement hérables, les IPGs parentaux pour les critères de comportement alimentaire et social ne sont pas de très bons prédicteurs des performances de leurs descendants, et que beaucoup d'effets de milieu (ex: lot) et d'environnement non contrôlé (composition des parcs) influencent l'expression de ces caractères.

Ces IPGs ont peu d'impact sur les performances de production individuelles (Tableau 10). Les variables qui ont le plus d'impact sont, pour le GMQ, l'IPG paternel du classement Elo ($p < 0,001$)

et pour la CA, l'IPG paternel pour le nombre moyen quotidien de visites pour la CA ($p = 0,015$). Ces deux IPG ont pourtant des corrélations faibles avec la variable correspondante observée durant la période de validation (Tableau 9). L'expression du comportement d'un animal ne dépend pas seulement de sa génétique, elle dépend aussi des autres individus dans le parc.

Tableau 9. Corrélation entre les IPGs estimés pour les parents et les variables de comportement observées chez leurs descendants durant la période de validation. Le fond en gris fait ressortir les variables de comportement social.

Variable observée	IPG père	IPG mère	IPG parents
VMQ	3,7 %	17,4 %	17,0 %
CMV	10,1 %	23,7 %	25,9 %
DMQ	21,8 %	18,1 %	29,6 %
VIM	26,2 %	14,0 %	30,4 %
D+MQ	-5,5 %	12,1 %	6,1 %
D-MQ	-6,2 %	9,4 %	4,1 %
PD+	-0,9 %	19,6 %	15,0 %
PJD+	-5,9 %	21,3 %	13,6 %
ELO	-0,2 %	16,8 %	12,0 %

Tableau 10. Impact des IPGs comportementaux du père et de la mère sur les performances individuelles de production durant la période de validation. Le fond en gris fait ressortir les variables de comportement social.

Variable ¹	GMQ 80-120 kg				CA 80-120 kg ³			
	IPG père ²		IPG mère ²		IPG père ²		IPG mère ²	
	Valeur <i>t</i>	Valeur <i>p</i>	Valeur <i>t</i>	Valeur <i>p</i>	Valeur <i>t</i>	Valeur <i>p</i>	Valeur <i>t</i>	Valeur <i>p</i>
VMQ	1,86	0,063	-0,95	0,345	-2,44	0,015	-0,72	0,474
CMV	-0,91	0,366	1,64	0,102	1,64	0,101	0,20	0,841
DMQ	1,44	0,151	0,20	0,845	-1,19	0,234	0,49	0,624
VIM	0,40	0,693	0,75	0,454	0,25	0,802	-1,21	0,228
D+MQ	1,97	0,049	-1,77	0,078	-1,39	0,165	0,05	0,964
D-MQ	2,22	0,027	-1,74	0,083	-1,16	0,246	0,18	0,859
PD+	1,45	0,148	2,07	0,039	0,91	0,363	-0,35	0,730
PJD+	0,79	0,428	2,16	0,031	1,18	0,238	-0,32	0,748
ELO	3,71	< 0,001	1,10	0,271	0,29	0,770	-0,15	0,883

¹Variable de comportement définie aux tableaux 2 et 3.

²IPG_p et IPG_m de l'Équation 3.

³Une transformation logarithmique a été appliquée sur la CA pour les tests statistiques.

L'analyse de l'impact des IPGs liés au comportement social sur les performances par parc a montré certains résultats statistiquement significatifs (Tableau 11), malgré le nombre relativement faible de parcs retenus pour la période de validation ($n = 62$).

Il est difficile de voir l'impact direct des IPGs sur les performances à partir de statistiques descriptives simples étant donné l'effet du lot qui a été important durant la période de validation

et aussi étant donné la méthode d'allotement, qui favorisait l'homogénéité des parcs, ce qui nuit à la détection de différence entre les parcs. Pour faire ressortir l'effet des IPGs, le parc le plus performant et le parc le moins performant dans chacun des 8 lots et pour chacun des 2 sexes ont été retenus. Un total de 16 couples de parcs contrastés, un « fort » et l'autre « faible » ont ainsi été créés pour chaque variable de performance (GMQ et CA). Les valeurs des IPGs entre les parcs forts et faibles pouvaient ainsi être directement comparés. Un exemple avec l'étendue de l'IPG des mères pour le nombre moyen quotidien de déplacements provoqués, variable ayant l'impact le plus significatif sur la CA ($p = 0,003$; Tableau 11), est présenté à la Figure 4. Le fait que 4 points sur 16 se retrouvent sur ou très près de la fonction identité (ligne hachurée; Figure 4) illustre l'homogénéité entre les parcs d'un même lot, rendant difficile la détection d'effets génétiques. Le fait que 10 des 12 points restants se retrouvent au-dessus de la fonction identité illustre qu'un parc formé d'animaux avec une grande étendue de l'IPG de la mère pour la variable D+MQ mènerait à une meilleure CA.

Tableau 11. Impact des IPGs liés au comportement social du père et de la mère sur les performances par parc durant la période de validation.

Variable ¹	GMQ 80-120 kg				CA 80-120 kg ³			
	IPG père ²		IPG mère ²		IPG père ²		IPG mère ²	
	Valeur <i>t</i>	Valeur <i>p</i>	Valeur <i>t</i>	Valeur <i>p</i>	Valeur <i>t</i>	Valeur <i>p</i>	Valeur <i>t</i>	Valeur <i>p</i>
moyenne(D+MQ)	-1,27	0,252	0,35	0,728	1,35	0,212	-0,36	0,722
écart-type(D+MQ)	1,64	0,107	2,20	0,033	0,25	0,806	-1,74	0,088
EIQ(D+MQ)	1,00	0,326	1,03	0,307	0,55	0,588	0,50	0,621
étendue(D+MQ)	1,00	0,322	2,62	0,012	0,90	0,372	-3,12	0,003
moyenne(D-MQ)	-1,04	0,341	0,27	0,785	1,28	0,236	-0,25	0,802
écart-type(D-MQ)	0,81	0,424	1,80	0,078	1,12	0,268	-1,50	0,140
EIQ(D-MQ)	-0,71	0,482	0,39	0,700	2,04	0,048	1,29	0,203
étendue(D-MQ)	0,50	0,621	2,48	0,016	1,25	0,217	-3,07	0,003
moyenne(PD+)	1,03	0,321	0,51	0,611	-0,57	0,581	0,32	0,751
écart-type(PD+)	-1,99	0,079	0,85	0,400	1,64	0,113	0,53	0,595
EIQ(PD+)	-2,51	0,017	-0,10	0,920	-0,74	0,464	-1,80	0,078
étendue(PD+)	-0,98	0,346	1,23	0,224	-0,31	0,764	0,01	0,989
moyenne(PJD+)	0,87	0,396	0,67	0,507	-0,33	0,742	-0,97	0,339
écart-type(PJD+)	-1,07	0,309	-0,08	0,936	1,28	0,208	0,03	0,976
EIQ(PJD+)	-2,12	0,039	-0,69	0,496	-0,91	0,368	-0,73	0,468
étendue(PJD+)	-0,03	0,973	-0,15	0,883	-0,46	0,651	0,90	0,371
moyenne(ELO)	0,93	0,367	-0,52	0,608	-1,65	0,109	-1,03	0,306
écart-type(ELO)	1,29	0,205	1,13	0,264	0,52	0,607	-0,83	0,412
EIQ(ELO)	-0,30	0,768	1,10	0,278	0,63	0,562	-1,21	0,230
étendue(ELO)	1,14	0,260	0,55	0,586	-1,89	0,067	-0,36	0,722

¹Variable de comportement social définie au Tableau 3.

² \overline{IPGp} et \overline{IPGm} de l'Équation 4.

³Une transformation logarithmique a été appliquée sur la CA pour les tests statistiques.

Les constats sont similaires avec l'étendue des IPGs de la mère pour la variable D-MQ, autant pour la CA que pour le GMQ. Pour les autres IPGs statistiquement significatifs ($p < 0,05$), les points sont répartis plus également de part et d'autre de la fonction identité. Il arrive aussi qu'il y ait trop de points identiques pour tirer des conclusions (Figure 5).

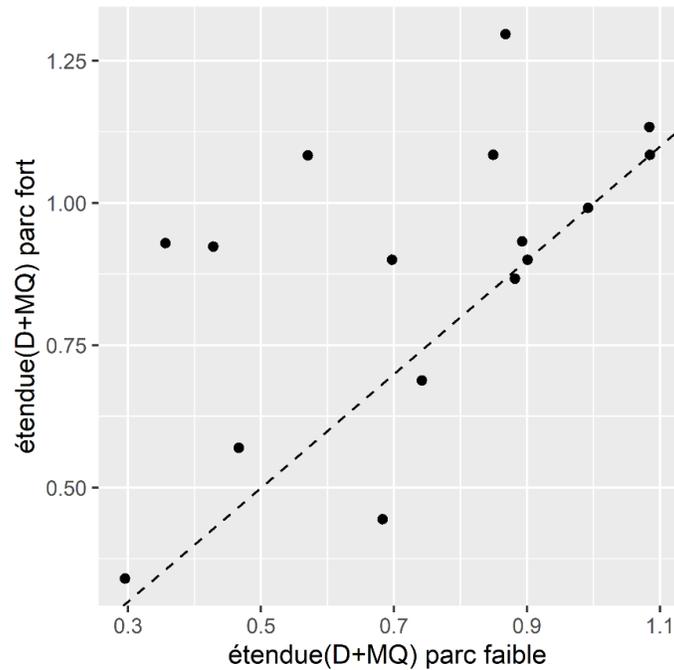


Figure 4. Étendue de l'IPG de la mère pour la variable D+MQ pour les 16 parcs faibles et forts pour la CA 80-120 kg. Un point (x,y) représente x : la valeur pour le parc le moins performant et y : la valeur pour le parc le plus performant dans un lot donné pour un sexe donné.

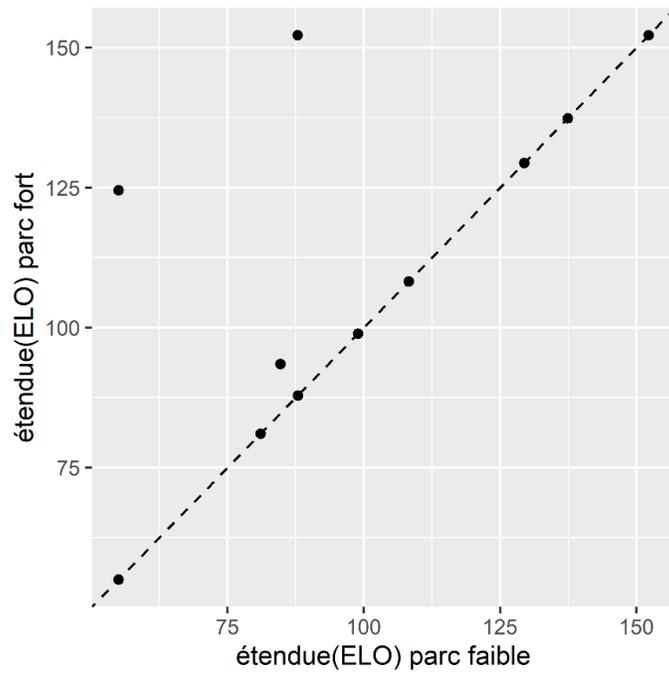


Figure 5. Étendue de l'IPG du père pour la variable ELO pour les 16 parcs faibles et forts pour la CA 80-120 kg. Un point (x,y) représente x : la valeur pour le parc le moins performant et y : la valeur pour le parc le plus performant dans un lot donné pour un sexe donné.

5 Conclusion

L'objectif général du projet était d'explorer les relations entre croissance, efficacité alimentaire, comportement alimentaire et composition de la carcasse chez des porcs en croissance afin de développer de nouveaux critères de sélection s'appuyant sur le comportement alimentaire et social.

Ce projet a montré la valeur des données fournies par les automates d'alimentation. Au-delà des variables classiques de performance comme la consommation quotidienne individuelle, l'analyse approfondie de chaque visite à l'automate a permis deux avancées prometteuses.

D'une part, les données d'identification individuelles ont permis d'améliorer la performance d'un algorithme de vision numérique pour le suivi individuel des animaux. La preuve de concept a été développée à partir de données d'abreuvement, mais l'approche peut être appliquée à partir de données d'alimentation. Des chaînes de Markov cachées permettent de pondérer la valeur des différentes sources d'information en combinaison avec le suivi des animaux, atténuant le risque de perdre la trace d'un animal ou de mélanger les identités. Il y a plusieurs défis techniques à résoudre pour passer de la preuve de concept à une application dans un contexte de production, mais le potentiel de l'approche est intéressant.

D'autre part, les données d'automates ont permis de caractériser le comportement alimentaire et social individuel et par parc à partir de l'analyse approfondie des séquences des visites. Les données d'animaux de races pures de deux différentes sources ont été utilisées pour développer les différents modèles et calculer de nouveaux indices de potentiel génétique basés sur le comportement social observé aux stations d'alimentation. Certains de ces indices ont montré une assez forte héritabilité pour envisager leur utilisation en sélection. D'autres études devraient être réalisées pour préciser la relation entre les IPGs de comportement social estimés sur les parents et les performances de production individuelles et par parc.

L'analyse approfondie des données d'automates pourrait aussi être utilisée pour caractériser le comportement social des truies en gestation élevées en groupes. Plus généralement, des méthodes d'apprentissage automatique pourraient être développées pour détecter différents événements à partir des données d'alimentation, des problèmes de boiterie à la détection de mise bas par exemple.

Références

- Bates, D., M. Maechler, B. Bolker et S. Walker, 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. DOI: 10.18637/jss.v067.i01.
- Bernardin, K. et R. Stiefelhagen, 2008. Evaluating Multiple Object Tracking Performance: The CLEAR MOT Metrics. *EURASIP Journal on Image and Video Processing*, Volume 2008, Article ID 246309, 10 pages. DOI: 10.1155/2008/246309
- Chen C.Y., I. Misztal, S. Tsuruta, W.O. Herring, J. Holl et M. Culbertson, 2000. Influence of heritable social status on daily gain and feeding pattern in pigs. *Animal Breeding and Genetics*, 127, 107-112.
- Elo, A.E., 1967. The Proposed USCF Rating System, Its Development, Theory, and Applications. *Chess Life*, XXII (8): 242-247.
- Friard, O. et M. Gamba, 2016. BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution* 2016, 7, 1325-1330.
- Gonzalez-Gracia, L.A., P. Gagnon, L. Maignel, É. Paquet et J. Dallaire, 2023a. Understanding social dynamics with social networks of pigs based on automatic feeder data: preliminary results. *Third International Student Symposium on Animal Behaviour and Cognition (ISSABC)*. Conjunto Amoxcalli, Facultad de Ciencias, UNAM, Mexico City, Mexico, 21 au 23 juin 2023.
- Gonzalez-Gracia, L.A., P. Gagnon, L. Maignel, É. Paquet et J. Dallaire, 2023b. First approaches to modelling social dynamics of intensively housed growing pigs based on automatic feeder data. *14e symposium du Centre de recherche en infectiologie porcine et avicole (CRIPA)*. Drummondville, QC, 8 décembre 2023.
- Gonzalez-Gracia, L.A., P. Gagnon, L. Maignel, É. Paquet et J. Dallaire, 2024. Social dynamics and aggression in growing pigs inferred from automatic feeder data. *57th Congress of the International Society of Applied Ethology. Animals and Society: Striking the right balance through Ethology*. Curitiba, Brésil, 22 au 26 juillet 2024.
- Kovac, M., E. Groenveld et L.A. Garcia-Cortez, 2002. VCE-5, a package for the estimation of dispersion parameters. *Proc. 7th WCGALP*, Montpellier (2002).
- Kuznetsova, A., P.B. Brockhoff et R.H.B. Christensen, 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13), 1-26. DOI: 10.18637/jss.v082.i13
- Ngo-Bibinbe, A.M.S, P. Gagnon, J. Dallaire et É. Paquet, 2024. An HMM-based framework for identity-aware long-term MOT from sparse and uncertain identification: use case on long-term

tracking in livestock. *4th Workshop on CV4Animals: Computer Vision for Animal Behavior Tracking and Modeling, In conjunction with CVPR 2024*. Seattle, États-Unis, 17 juin 2024.

Psota E.T., T. Schmidt, B. Mote, et L.C. Pérez, 2020. Long-term tracking of group-housed livestock using keypoint detection and map estimation for individual animal identification. *Sensors*, 20(13), 3670. DOI: 10.3390/s20133670

R Core Team, 2022. R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. Vienne, Autriche. URL <https://www.R-project.org/>.

Rohrer G.A., T. Brown-Brandl, L.A. Rempel, J.F. Schneider et J. Holl, 2013. Genetic analysis of behavior traits in swine production. *Livestock Science*, 157, 28-37.

Zhang, Y., P. Sun, Y. Jiang, D. Yu, F. Weng, Z. Yuan, P. Luo, W. Liu et Xi. Wang, 2021. Bytetrack: Multi-object tracking by associating every detection box. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 13682. LNCS: 1-21.

Annexe 1

Tableau 12. Corrélations génétiques et phénotypiques entre les caractères étudiés – Race Duroc (n = 4480 animaux)

Variable	GMQ	CA	EGD	CMQ	VMQ	DMQ	DMV	CMV	VIM	PVMP	PCMP	D+MQ	D-MQ	PD+	PJD+	ELO
GMQ	0.17	-0.24	+0.27	+0.79	+0.04	+0.03	+0.10	+0.21	+0.18	+0.16	+0.15	+0.11	+0.09	+0.07	+0.09	+0.04
CA	-0.47	0.16	+0.43	+0.20	+0.00	-0.01	-0.06	+0.06	+0.03	+0.05	+0.02	+0.08	-0.10	+0.26	+0.22	+0.24
EGD	+0.27	+0.18	0.45	+0.46	-0.01	-0.08	-0.02	+0.15	+0.21	+0.02	+0.11	-0.04	+0.08	-0.06	-0.04	+0.05
CMQ	+0.65	+0.22	+0.42	0.21	+0.06	+0.04	+0.01	+0.19	+0.27	+0.15	+0.13	+0.18	+0.08	+0.24	+0.22	+0.23
VMQ	-0.04	+0.11	-0.06	+0.00	0.49	+0.38	-0.50	-0.91	-0.38	+0.15	+0.12	+0.69	+0.83	+0.26	+0.04	+0.08
DMQ	+0.03	+0.03	+0.01	+0.02	+0.21	0.56	+0.48	-0.17	-0.96	+0.16	+0.10	+0.22	+0.60	-0.23	-0.40	-0.34
DMV	+0.08	-0.08	+0.04	+0.04	-0.57	+0.40	0.47	+0.76	-0.46	+0.08	+0.03	-0.43	-0.27	-0.42	-0.29	-0.30
CMV	+0.29	-0.07	+0.17	+0.33	-0.69	-0.16	+0.75	0.34	+0.24	-0.02	-0.05	-0.56	-0.72	-0.18	+0.08	+0.02
VIM	+0.30	+0.00	+0.17	+0.41	-0.17	-0.10	-0.33	+0.32	0.48	-0.10	-0.01	-0.16	-0.58	+0.31	+0.45	+0.38
PVMP	0.00	+0.09	-0.02	+0.08	+0.06	+0.08	+0.00	-0.03	-0.05	0.31	+0.94	+0.40	+0.06	+0.51	+0.45	+0.52
PCMP	+0.01	+0.02	+0.01	+0.04	+0.07	+0.09	-0.04	-0.06	-0.01	+0.76	0.37	+0.38	+0.05	+0.47	+0.46	+0.48
D+MQ	-0.03	+0.00	-0.05	+0.03	+0.58	+0.19	-0.31	-0.37	+0.16	+0.20	+0.13	0.26	+0.71	+0.77	+0.56	+0.78
D-MQ	-0.04	+0.00	-0.05	-0.01	+0.58	+0.28	-0.25	-0.38	-0.26	+0.14	+0.03	+0.80	0.20	+0.11	-0.16	-0.07
PD+	-0.01	+0.08	-0.03	+0.07	+0.24	-0.01	-0.24	-0.21	+0.04	+0.19	+0.19	+0.63	+0.14	0.38	+0.95	+0.99
PJD+	+0.03	+0.05	+0.00	+0.08	+0.04	-0.14	-0.07	+0.05	+0.17	+0.11	+0.17	+0.39	-0.15	+0.81	0.37	+0.99
ELO	+0.01	+0.16	+0.01	+0.05	+0.07	-0.07	-0.10	-0.03	+0.09	+0.15	+0.16	+0.34	-0.03	+0.67	+0.66	0.31

Sur la diagonale : hérabilités; au-dessus de la diagonale : corrélations génétiques; au-dessous de la diagonale : corrélations phénotypiques; GMQ = gain moyen quotidien; CA = conversion alimentaire; EGD = épaisseur de gras dorsal; Pour la définition des autres caractères, voir tableaux 2 et 3.



Centre de développement du porc du Québec inc.
815 Rte Marie-Victorin
Lévis (secteur Saint-Nicolas)
Québec G7A 3S6
www.cdpg.ca