

Identification de biomarqueurs de l'efficacité alimentaire chez le porc en croissance

Florence GONDRET (1), Mette SKOU HEDEMAN (2), Farouk MESSAD (1), Lianne VERSHUREN (3,7), Junjun WANG (4), Gabriel de la FUENTE (5), Hélène GILBERT (6), Alfons JANSMAN (7)

(1) PEGASE, INRAE, AGROCAMPUS Ouest, 35590 Saint-Gilles, France

(2) Animal Sciences Department, Aarhus University, 8830 Tjele, Denmark

(3) Topigs Norsvin Research Center B.V., Beuningen, Pays-Bas

(4) College of Animal Science and Technology, China Agricultural University (CAU), Beijing, Chine

(5) Dpt. Ciencia Animal, Campus ETSEA, Universitat de Lleida, Espagne

(6) GenPhyse, INRAE, INPT, INPT-ENV, Université de Toulouse, 31320 Castanet-Tolosan, France

(7) Wageningen Livestock Research, 6700 AH Wageningen, Pays-Bas

florence.gondret@inrae.fr

Avec la collaboration du consortium et le soutien financier du programme européen Feed-a-Gene

Identification de biomarqueurs de l'efficacité alimentaire chez le porc en croissance

L'efficacité alimentaire inclut deux composantes, l'une relative à l'efficacité digestive des aliments et la seconde à l'utilisation post-absorptive des nutriments. Pour améliorer la prédiction de l'efficacité alimentaire et en identifier de nouveaux proxys, il est important de disposer de marqueurs biologiques pertinents pour ces deux composantes. De nombreuses molécules peuvent être mesurées dans les tissus, les fluides ou les fèces de l'animal. Différents travaux au sein du projet européen Feed-a-Gene ont étudié la possibilité d'identifier des biomarqueurs de l'efficacité alimentaire globale ou de l'une de ses composantes. Pour cela, des porcs présentant des capacités différentes à utiliser l'azote induites par des leviers génétiques ou des facteurs nutritionnels, ou bien issus d'une sélection divergente sur l'ingéré résiduel ou encore choisis pour présenter des valeurs extrêmes d'efficacité alimentaire, ont été étudiés. Les signatures géniques ou métaboliques dans le muscle, le sang ou l'urine de ces animaux ont été obtenues par des technologies le plus souvent sans *a priori* et à haut-débit. Les résultats valident globalement l'hypothèse selon laquelle il est possible d'identifier des molécules discriminantes des différences phénotypiques d'efficacité alimentaire, et même prédictives des valeurs individuelles. La variété des fonctions biologiques représentées par les gènes ou les métabolites inclus dans les modèles de discrimination ou de prédiction confirme la nature intégrative et complexe de l'efficacité alimentaire chez le porc en croissance. Les conditions d'une utilisation future de ces biomarqueurs pour raffiner les mesures d'efficacité alimentaire et permettre le tri ou la sélection d'animaux sur ce caractère sont évoquées en perspective.

Identifying biomarkers as new traits for feed efficiency in growing pigs

Feed efficiency has two components, one related to digestive efficiency and the second to post-absorptive nutrient efficiency. Therefore, new data and proxies referring to digestive or post-absorptive efficiencies are needed to improve feed efficiency in growing pigs. A large number of small molecules involved in or generated by metabolic processes in tissues, body fluids or feces may be considered as molecular signatures to assess inter-individual differences in feed efficiency. Various studies within the European Feed-a-Gene project have investigated the possibility of identifying biomarkers of overall food efficiency or one of its components. Pigs with genetically- or nutritionally-induced differences in nitrogen utilization, divergent for residual feed intake, or extremes of feed-efficiency traits were compared. Data were generated for the pigs using untargeted high-throughput metabolomics or transcriptomics applied to muscle, blood or urine. The results support the hypothesis that it is possible to identify subsets of gene transcripts or metabolites in tissues and body fluids that can discriminate pigs according to feed efficiency or predict individual values of feed efficiency. The variety of biological functions represented by the genes or metabolites included in the predictive models confirms the integrative and complex nature of feed efficiency in growing pigs. Perspectives to include these biomarkers as new traits in subsequent selection procedures or to help in sort of animals during particular periods are discussed.

INTRODUCTION

L'efficacité alimentaire est un caractère important pour la filière porcine, dans la mesure où l'aliment représente près des deux tiers du coût de production, et où les rejets consécutifs à une mauvaise valorisation de l'aliment par l'animal sont sources de pollution environnementale. Des mesures de l'efficacité alimentaire, accessibles et adoptables dans les élevages de sélection ou commerciaux, ont été proposées depuis plusieurs années. Il s'agit de l'indice de conversion alimentaire ou IC, défini comme le rapport entre la quantité d'aliment ingérée et le gain de poids pendant une période de référence, et de l'ingéré résiduel ou CMJR (pour consommation moyenne journalière résiduelle), défini comme la différence entre la quantité d'aliment ingéré observée et celle prédite relativement aux besoins d'entretien et aux besoins de production de l'individu. De nombreuses études ont été conduites pour comprendre les sources de variabilité de ces mesures, qu'elles soient d'origine environnementales ou internes à l'animal (Gilbert *et al.*, 2017). L'ensemble de ces études s'accorde sur le fait que l'efficacité alimentaire est un caractère complexe. En outre, sa valeur n'est connue qu'à l'issue de la période de test, généralement fixée entre 30 et 115 kg de poids vif chez le porc. Aussi, disposer de proxys de ce caractère pourrait permettre d'investiguer plus facilement l'origine de la variabilité de l'efficacité alimentaire, de donner des éléments objectifs de tri des animaux à des périodes clés, et d'améliorer les schémas de sélection par la mesure d'un plus grand nombre d'animaux ou l'utilisation de mesures plus précises.

L'efficacité alimentaire inclut deux composantes, l'une relative à l'efficacité digestive des aliments et la seconde à l'utilisation post-absorptive des nutriments. Il est important d'identifier des biomarqueurs pour l'une ou l'autre de ces deux composantes et pour l'efficacité alimentaire globale. Pour cela, utiliser des leviers de variation déjà connus de l'efficacité alimentaire permet de disposer d'une gamme relativement large de valeurs individuelles. De nombreuses molécules peuvent aussi être mesurées dans les tissus et les fluides (sang, salive, urine) collectés chez l'animal. Elles concernent différents niveaux d'organisation du vivant, comme des séquences géniques, des transcrits de gènes (ARN messagers), des ARN non codants, des protéines ou des métabolites. Ces molécules sont actrices de processus biologiques ou sont les produits de ces processus. Des séquences de micro-organismes comme celles du microbiote intestinal hébergé par l'animal peuvent également être analysées dans les fèces. Différents travaux au sein du projet européen Feed-a-Gene ont donc étudié la possibilité d'identifier des biomarqueurs de l'efficacité alimentaire parmi ces molécules, en utilisant différentes sources de variation du phénotype d'intérêt. Cet article expose de façon synthétique les principaux résultats. Ils valident globalement la possibilité d'identifier des molécules discriminantes des différences d'efficacité alimentaire, et même prédictives des valeurs individuelles d'efficacité alimentaire chez le porc en croissance.

1. MATERIEL ET METHODES

1.1. Principaux leviers de variation de l'efficacité alimentaire

Les dispositifs expérimentaux incluaient :

- des porcs en croissance présentant des différences dans leurs capacités à valoriser l'azote. Celles-ci étaient induites par des leviers génétiques tels que la race (Duroc, Large White, Large

- White x Landrace ou lignées composites) ou s'appuyaient sur les valeurs génétiques individuelles contrastées pour le dépôt de protéines existantes au sein d'une même race. L'application de facteurs nutritionnels, et notamment la comparaison entre des régimes équilibrés et des régimes à basse teneur en protéines (70% des besoins), ont également permis de générer des différences entre porcs dans leur efficacité d'utilisation de l'azote. Ces études ont été respectivement menées au sein de l'Université de Wageningen (Pays-Bas) par Jansman et van der Peet-Schwering (2018), et au sein de l'Université de Lleida (Espagne) par Sarri *et al.* (2019).

- des groupes de porcs constitués à l'issue d'une période de test pour présenter des différences importantes d'ingéré alimentaire pour un même gain de poids, des différences de gain de poids pour un même ingéré alimentaire, ou bien encore des valeurs extrêmes d'IC. Ces études ont été menées, respectivement, par l'Université de Pékin (Chine) par He *et al.* (2018), et par l'Université de Wageningen et l'entreprise Topigs Norsvin (Pays-Bas) par Vershuren *et al.* (2019).

- des porcs issus d'une sélection génétique divergente intra-race sur l'ingéré résiduel entre 30 et 95 kg de poids vif (lignées à CMJR faible ou élevée), et soumis à des régimes de composition nutritionnelle différente, notamment quant à la teneur en fibres, ou à des niveaux alimentaires contrastés (notamment, à volonté ou restreint). Ces études ont été menées par l'INRA et les modalités expérimentales sont détaillées par Gilbert *et al.* (2019) et par Messad *et al.* (2019).

1.2. Analyses des molécules et compartiments cibles

Classiquement, la recherche de biomarqueurs procède soit par la quantification de molécules cibles sur la base d'hypothèses biologiques quant à leur rôle potentiel dans le phénotype, soit par la mesure sans *a priori* de plusieurs dizaines à plusieurs milliers de molécules dans un objectif de découverte de nouveaux biomarqueurs. Dans le premier cas, les groupes expérimentaux sont comparés quant à leur différence d'expression des candidats. C'est sur *IGF1* (un gène codant pour le facteur de croissance *insulin-like growth factor-1*) qu'a notamment porté l'étude de He *et al.* (2018).

Cependant, la majeure partie des études menées dans le cadre du projet Feed-a-Gene concernait le second objectif visant à la découverte de nouveaux biomarqueurs. Pour cela, des analyses ont été conduites dans les tissus ou fluides animaux, soit à l'échelle des transcrits des gènes par analyses des transcriptomes par puce à ADN (études conduites à l'INRA), soit à l'échelle des métabolites par spectrométrie de masse couplée à la chromatographie liquide (études conduites à l'Université d'Aarhus Danemark) selon la méthodologie décrite par Curtasu *et al.* (2019).

1.3. Analyses statistiques

Pour déterminer si une molécule cible peut être un candidat biomarqueur, les groupes de porcs expérimentaux ont été comparés quant à leur différence d'expression de la molécule cible par des analyses statistiques simples de type ANOVA.

Dans le cas de la recherche de nouveaux biomarqueurs, deux grands types de méthodologies ont été utilisées pour prendre en compte le grand nombre de variables mesurées (c'est-à-dire les centaines à milliers de données acquises sans *a priori*) relativement au petit nombre d'observations (les porcs expérimentaux). Pour proposer les molécules les plus aptes à discriminer des groupes d'individus constitués selon leur race,

leur lignée, leur type génétique ou leur régime, des analyses par régression multivariée des moindres carrés partiels (PLS-DA) ont été mises en œuvre. Ces analyses maximisent la variance des variables explicatives, ainsi que la corrélation entre chacune des variables explicatives et la variable à expliquer, tout en supervisant l'analyse grâce à la connaissance des groupes expérimentaux. Elles permettent donc de répondre à deux questions : 1) les groupes expérimentaux sont-ils réellement différents ? et 2) quelles sont les variables qui décrivent le mieux les différences entre groupes ? Ces analyses peuvent être complétées par une analyse discriminante des composantes principales (DAPC) pour identifier des groupes de métabolites et leurs relations.

Pour proposer des modèles prédictifs pour des phénotypes discrets (comme les classes de CMJR) ou bien des phénotypes continus (valeur individuelle d'IC), des algorithmes d'apprentissage automatique (dits de « *machine learning* ») procédant par de multiples combinaisons d'arbres de régression décisionnelle, ont été utilisés. Ils permettent d'identifier un petit nombre de variables dites les plus importantes dans la prédiction (VIP pour « *very important variable in prediction* ») à partir d'un jeu de données permettant à l'algorithme d'apprendre la structure des données et de corriger pas à pas ses erreurs. Un jeu de données de validation de la qualité du modèle de prédiction ainsi construit est ensuite utilisé. Cette méthodologie d'apprentissage automatique vise donc à proposer des modèles décisionnels permettant de prédire le caractère d'intérêt, mais cette fois sans *a priori* sur les groupes expérimentaux d'origine.

2. RESULTATS

2.1. Identification de marqueurs géniques musculaires pour classer les porcs ou prédire leur niveau d'efficacité alimentaire

Le muscle long dorsal a été choisi comme tissu cible pour identifier des marqueurs moléculaires pouvant servir de proxys aux mesures d'efficacité alimentaire. En effet, des études antérieures (Gondret *et al.*, 2017) ont montré que ce muscle est particulièrement affecté par les divergences d'efficacité alimentaire, à la fois quant à son développement relatif (les porcs les plus efficaces étant plus musclés) et à sa fonctionnalité (de nombreux gènes et voies métaboliques sont modifiés en relation avec les différences d'efficacité alimentaire).

Une analyse des variations de quelques gènes cibles décrits dans la littérature comme jouant un rôle important dans la croissance musculaire a été conduite pour un petit nombre de porcs présentant des différences d'indice de consommation.

Tableau 1 – Expression du gène *IGF1* dans le muscle pour des groupes de porcs à efficacité alimentaire contrastée¹

Variables	Groupes			
	LFI	HFI	HADG	LADG
IC, kg/kg	1,97a	2,28b	2,01a	2,46b
IGF1, UA	1,00a	0,48b	1,00a	0,80a

¹IC : indice de conversion alimentaire ; IGF1 : insulin like growth factor ; LFI : ingéré alimentaire faible et HFI : ingéré alimentaire élevé, pour un même gain moyen quotidien ; HADG : croissance journalière élevée et LADG : croissance journalière faible pour un même ingéré alimentaire ; UA : unités arbitraires (quantité relative de transcrit). Les valeurs moyennes affectées d'une même lettre ne diffèrent pas ($P > 0,05$) entre groupes. $n = 16$ porcs.

Cette analyse a révélé des différences dans les niveaux d'expression du facteur de croissance *IGF1* (Tableau 1) entre des porcs choisis pour présenter des différences d'ingéré alimentaire pour une même vitesse de croissance mesurée sur une période de 28 jours en début d'engraissement (He *et al.*, 2018).

Une seconde approche a été d'utiliser des porcs Large White issus d'une sélection divergente sur l'ingéré résiduel (CMJR) au cours de deux expériences successives pour maximiser les sources de variations (génération, sexe, nutrition, saison d'élevage, etc.) et donc les amplitudes des différences interindividuelles d'efficacité alimentaire. Des mesures à haut-débit des ARN messagers ont été réalisées dans le muscle long dorsal (environ 7 000 gènes exprimés) chez les porcs autour du stade commercial d'abattage (entre 80 et 115 kg de poids vif). Il s'agissait alors de déterminer dans quelle mesure il était possible de trouver une combinaison d'un petit nombre de transcrits de gènes permettant de prédire, respectivement, la valeur génétique de CMJR et la valeur mesurée d'IC (Messad *et al.*, 2019).

Le modèle bâti à partir d'un algorithme d'apprentissage automatique procédant par de multiples combinaisons d'arbres de régression décisionnelle sur un jeu d'apprentissage ($n = 50$ porcs) puis validé sur un second jeu de données ($n = 21$ porcs), retient les valeurs de transcrits d'une trentaine de gènes uniques. Ce modèle explique 65% de la variabilité des valeurs génétiques d'ingéré résiduel, et 67% de la variabilité des mesures d'IC (Tableau 2). Pour l'IC, l'erreur de prédiction (RMSE) est donc seulement de 8% pour une valeur moyenne d'IC de 2,75 pour l'ensemble des porcs considérés dans cette étude.

Tableau 2 – Modèles de prédiction de la valeur génétique d'ingéré résiduel et de la valeur phénotypique individuelle d'IC à partir de l'expression de gènes musculaires¹

	Valeur du caractère		Évaluation du modèle de prédiction		
	Min	Max	Nb VIP	R ²	RMSE
CMJR, g/j	-108	92	27	0,65	39,3
IC, kg/kg	2,25	3,28	33	0,67	0,22

¹CMJR : ingéré résiduel (valeur génétique) ; IC : indice de conversion alimentaire ; Min : valeur minimale, et Max : valeur maximale, mesurées (pour l'IC) ou prédites (pour la CMJR) dans la population de porcs considérés ; Nb VIP : nombre de variables importantes dans la prédiction ; R² : coefficient de régression ; RMSE : erreur de prédiction. $n = 71$ porcs.

Les gènes identifiés comme de bons prédicteurs (VIP) dans le modèle de prédiction de l'IC sont décrits dans la littérature comme participant à des processus relatifs au métabolisme énergétique, au métabolisme protéique, à la régulation du développement musculaire, à l'adhésion cellulaire et à la fonction immunitaire. Ces différentes fonctions avaient été précédemment décrites comme affectées par les différences d'ingéré résiduel entre les lignées divergentes d'origine (Gondret *et al.*, 2017).

L'utilisation d'une puce à ADN pour obtenir les mesures des transcrits de gènes n'étant pas concrètement envisageable en routine, l'expression de 24 de ces gènes VIP a été ensuite mesurée par qPCR (une technique d'analyse ciblée). Des modèles statistiques simples procédant par régression linéaire pas à pas ont alors permis de mettre en lumière un petit sous-ensemble de gènes les plus significatifs.

La prédiction reste plutôt bonne avec un sous-ensemble de seulement 10 gènes, ceci notamment pour l'IC ($R^2 = 0,77$; Tableau 3).

Tableau 3 – Gènes musculaires identifiés comme les plus importants dans la prédiction de l'efficacité alimentaire¹

CMJR		IC	
$R^2 = 0,58$		$R^2 = 0,77$	
Gènes	P	Gènes	P
<i>FKBP5</i>	<0,001	<i>FKBP5</i>	<0,001
<i>SERINC3</i>	0,02	<i>MUM1</i>	0,03
<i>IGF2</i>	0,03	<i>AKAP12</i>	0,03
<i>CSRNP3</i>	0,03	<i>FYN</i>	0,03
<i>EZR</i>	0,09	<i>TMED3</i>	0,08
<i>RPL16</i>	0,08	<i>PHKB</i>	0,08
		<i>TFG</i>	0,02
		<i>SOCS6</i>	0,07
		<i>ILR4</i>	0,10
		<i>FRAS1</i>	0,12

¹CMJR : ingéré résiduel (valeur génétique); IC : indice de conversion alimentaire; P : probabilité d'entrée ($< 0,15$) dans un modèle de régression linéaire pas à pas. Les gènes sont nommés par leur symbole international. $n = 71$ porcs.

Cette démarche de recherche sans *a priori* et à haut débit des transcrits dans le muscle des porcs permet donc bien d'identifier des proxys de phénotypes complexes comme l'efficacité alimentaire. Elle suggère que, sur la base d'un échantillon de muscle prélevé à l'abattoir, il est possible d'obtenir une bonne prédiction du niveau d'efficacité alimentaire de tous les animaux de la population, que ce soit en terme de valeur génétique estimée pour la CMJR, ou de valeur phénotypique mesurée pour l'IC. Il reste à déterminer si ces biomarqueurs candidats seront validés pour un plus grand nombre d'individus que celui considéré dans notre démarche de découverte.

2.2. Identification de biomarqueurs circulants discriminants de l'efficacité alimentaire

Au vu des résultats encourageants obtenus sur le muscle, il a été proposé d'investir massivement dans la recherche de biomarqueurs dans des compartiments biologiques plus accessibles, comme les fluides circulants.

Les études ont notamment porté sur l'identification de métabolites dans l'urine, susceptibles de discriminer l'efficacité d'utilisation de l'azote. Pour cela, des groupes de porcs (croisements entre verrats d'une lignée synthétique et femelles Large White x Landrace), présentant des différences génétiques dans leur capacité à déposer des protéines (Jansman *et al.*, 2018), ont été comparés. Les analyses métabolomiques révèlent qu'il est possible de discriminer ces deux groupes de porcs comme présentant une valeur génétique basse ou une valeur génétique élevée pour le dépôt protéique, sur la base d'une composante latente (PC2) qui combine les concentrations de plusieurs métabolites mesurés dans l'urine (Figure 1). Les métabolites discriminants identifiés sur cette composante sont notamment les concentrations en cinnamoylglycine (un produit de bactéries intestinales métabolisant la phénylalanine) pour expliquer les faibles valeurs génétiques de dépôt de protéines, et les concentrations en pipéridine et créatinine pour expliquer les valeurs génétiques élevées de dépôt de protéines.

La pipéridine résulte de la fermentation bactérienne des protéines et des peptides non digestibles, tandis que la créatinine est directement associée au métabolisme énergétique. Ces résultats suggèrent que les porcs à capacité élevée de dépôt protéique ont des différences d'activité fermentaire digestive, et un métabolisme énergétique plus élevé que les porcs à moindre capacité de dépôt protéique. Ce dernier élément est sans doute lié au fait que la synthèse protéique est considérée comme ayant un coût énergétique élevé pour l'animal. Il faut noter que, dans cette expérience, les métabolites discriminants n'expliquaient cependant que 7,2% de la variabilité totale.

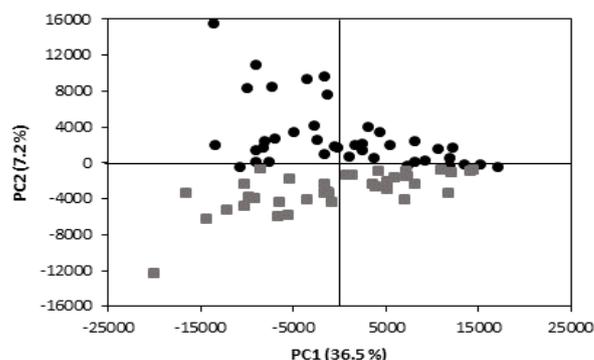


Figure 1 – Analyse discriminante (PLS-DA) des métabolites urinaires entre groupes de porcs à valeur génétique contrastée par leur capacité génétique de dépôt de protéines

Les porcs ($n = 40$) présentant une capacité faible (●) ou élevée (■) à déposer des protéines sont projetés sur le plan principal formé par les deux premières composantes latentes (PC)

Lorsque l'on s'intéresse aux métabolites circulants dans le plasma, il est également possible de discriminer ces mêmes groupes de porcs à partir d'une analyse métabolomique, mais là aussi, avec un faible pourcentage d'explication (10%) de la variance totale du caractère (données non présentées). De même, il est possible de discriminer deux lignées de porcs (race Large White) issus d'une sélection divergente sur l'ingéré résiduel conduite à l'INRA (porcs à CMJR faible ou élevée; Gilbert *et al.*, 2019) avec une variance expliquée autour de 20% (Figure 2) par analyse métabolomique du plasma collecté à 15 et 23 semaines d'âge.

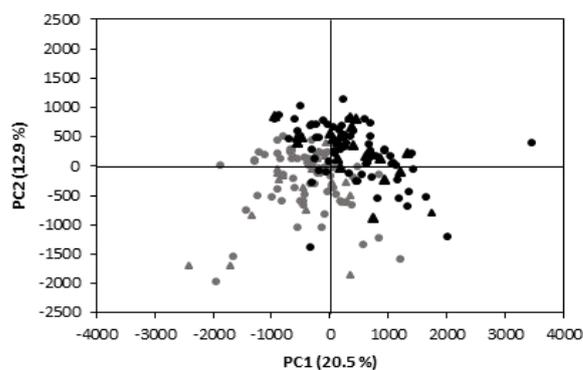


Figure 2 – Analyse discriminante (PLS-DA) des métabolites plasmatiques collectés à 15 et 23 semaines d'âge dans deux lignées de porcs sélectionnées sur l'ingéré résiduel.

Les porcs ($n = 90$) présentant un ingéré résiduel faible (mesuré : ●; prédit : ◐) ou élevé (mesuré : ◑; prédit : ◓) sont projetés sur le plan principal formé par les deux premières composantes latentes (PC)

Les métabolites les plus discriminants sont des acides aminés (phénylalanine, glutamate, tryptophane) et des dérivés de leurs métabolismes, comme la cétoleucine qui est le résultat du catabolisme incomplet des acides aminés ramifiés, ou l'acide hippurique, un dérivé du métabolisme hépatique de la glycine.

Toujours sur le plasma, deux groupes de porcs (issus d'un croisement entre verrats d'une lignée synthétique et femelles de type Large White x Landrace), constitués sur la base de valeurs extrêmes d'IC (faible ou élevée ; Vershuren *et al.*, 2018) ont été également comparés. Ils peuvent être discriminés par les concentrations de métabolites mesurés dans le plasma collecté le jour de l'abattage par analyse métabolomique (données non présentées). Dans ce cas, les métabolites participant à la discrimination des valeurs d'IC sont là aussi des acides aminés (leucine, tryptophane, valine), mais aussi la pipéridinone, et le p-crésol, un produit de la fermentation bactérienne des protéines.

L'ensemble de ces résultats suggère que les différences d'efficacité protéique comme d'efficacité alimentaire globale chez des porcs en croissance résultent, au moins en partie, de différences dans le métabolisme protéique, des acides aminés et de leurs dérivés. Du fait de l'identification de métabolites connus pour être issus du catabolisme de ces acides aminés, ils suggèrent aussi que les besoins protéiques des porcs à efficacité alimentaire élevée sont insuffisamment couverts par les régimes standards. D'ailleurs, Saintilan *et al.* (2015) soulignaient que les besoins des porcs les plus efficaces, que ce soit sous l'angle de l'ingéré résiduel ou de l'IC, étaient légèrement plus élevés que les recommandations actuelles formulées pour les porcs en croissance. Dans toutes ces études de métabolomique, la discrimination des porcs sur la base de leur efficacité alimentaire est cependant bien moins bonne (entre 7 et 20%) que les discriminations obtenues quand on cherche à différencier les animaux sur la base de la race ou des régimes qu'ils ont reçus (plus de 50% de variabilité expliquée ; données non présentées). Cette observation confirme les études précédentes (Jégou *et al.*, 2016) montrant qu'il existe peu de différences dans les concentrations en métabolites plasmatiques entre lignées de porcs issus d'une sélection divergente sur la CMJR comparativement aux différences induites par un régime riche en fibres et en lipides.

2.3. Identification de biomarqueurs sanguins précoces d'efficacité alimentaire

L'efficacité alimentaire est un phénotype qui se mesure durant une période de test relativement longue, avec un coût important d'évaluation en raison de la mesure individuelle de l'ingéré alimentaire. Aussi, disposer de biomarqueurs précoces de ce phénotype connu uniquement à l'issue de la période de test représenterait une avancée significative en élevage : tri précoce des animaux à phénotyper selon leur potentiel d'efficacité alimentaire, pilotage individuel de la composition de leur ration en fonction de ce potentiel, etc. Des études précédentes montraient des différences importantes entre des lignées de porcs divergentes pour l'ingéré résiduel lorsque l'on s'intéressait aux transcrits de gènes présents dans les cellules sanguines circulantes dans le sang (Jégou *et al.*, 2016). L'avantage de ce compartiment est aussi qu'il peut être prélevé à n'importe quel stade de la croissance. C'est pourquoi des études ont porté sur l'identification de transcrits de gènes circulants comme biomarqueurs candidats à différents stades de la croissance pour expliquer les différences d'IC lors des périodes de test. Plusieurs jeux de données transcriptomiques

sanguines précédemment ou nouvellement acquises chez des porcs âgés de 40 jours (Gondret, données non publiées), 87 jours (Reis Furtado Campos *et al.*, 2014) ou 132 jours (Jégou *et al.*, 2016) ont été associés. Tous ces jeux de données concernaient les lignées divergentes INRA (race Large White) sélectionnées sur l'ingéré résiduel (CMJR) pendant 6 à 9 générations. Les périodes de mesures de l'IC étaient, respectivement, de 90 à 160 jours d'âge, de 87 à 100 jours d'âge, et de 76 à 132 jours d'âge. La valeur moyenne de l'IC était de 2,85 pour un total de 148 porcs ayant pu être analysés quant à leur transcriptome sanguin. Considérer différents jeux de données acquises à différentes périodes de la croissance et différentes périodes de mesures de l'IC doit en théorie garantir une certaine généralité des biomarqueurs. Cela devrait éviter qu'un biomarqueur identifié à un stade précis ne soit plus valide quelques semaines avant ou après. La moitié des individus disponibles (69 porcs) a été utilisée comme un dispositif d'apprentissage décisionnel pour construire des modèles capables de prédire, d'une part la classe d'ingéré résiduel (caractère dichotomique : faible ou élevée), et d'autre part l'IC (valeurs individuelles continues). Les modèles sont basés sur des arbres de classification dans le premier cas, et des arbres de régression dans le second cas. Puis, les modèles ont été validés quant à leur capacité prédictive sur la seconde moitié des individus (79 porcs). Les résultats quant à la prédiction de la classe d'ingéré résiduel sont synthétisés dans le tableau 4. Les résultats indiquent qu'il est donc possible de classer les porcs sur un caractère dichotomique (faible ou élevé) d'ingéré résiduel avec une très bonne précision (97,5% de succès) à partir de la combinaison des niveaux d'expression de 25 gènes uniques mesurés dans le sang total durant le post-sevrage et la croissance. Autrement dit, le modèle se trompe moins de quatre fois sur 100 pour classer un porc inconnu en individu à CMJR faible ou à CMJR élevée.

Tableau 4 – Taux de succès d'un modèle de classification des porcs sur leur ingéré résiduel¹

Classes réelles	N porcs	Classes prédites		
		CMJR élevée	CMJR faible	% succès
CMJR élevée	42	41	1	97,6
CMJR faible	37	1	36	93,5
Taux de succès moyen (% de bons classements) : 97,5%				

¹Le modèle est construit à partir de la mesure de l'expression de gènes (quantité de transcrits) dans le sang total prélevé chez des porcs en post-sevrage ou en croissance. CMJR: ingéré résiduel.

Les 25 gènes retenus (données non détaillées) dans le modèle de classification sont décrits dans la littérature comme impliqués dans les processus de réponse aux nutriments, le stress oxydant, la phosphorylation oxydative, la différenciation cellulaire et la régulation des processus développementaux.

Pour la prédiction de la valeur mesurée d'IC, le modèle de régression décisionnelle permet là aussi de prédire les valeurs individuelles avec une bonne précision (Tableau 5). L'erreur de prédiction est estimée à 9% par rapport à la valeur moyenne d'IC mesurée, et la part de la variance totale expliquée est estimée à 70% du caractère. Parmi les 27 gènes retenus comme candidats biomarqueurs dans ce dernier cas (données non détaillées), on peut notamment souligner la présence de gènes liés au système IGF comme *IGF2*, mais aussi plusieurs gènes liés au catabolisme protéique (par exemple, *PSMB9*, *SHPRH*, *FEMC1*).

Tableau 5 – Modèle de prédiction de la valeur individuelle d'IC à partir de l'expression de gènes sanguins¹

	Valeur du caractère		Modèle de prédiction		
	Min	Max	Nb VIP	R ²	RMSE
IC, kg/kg	1,92	5,00	27	0,70	0,26

¹IC : indice de conversion alimentaire; Min : valeur minimale et Max : valeur maximale, mesurées dans la population considérée; Nb VIP : nombre de variables importantes dans la prédiction; R² : coefficient de régression; RMSE : erreur de prédiction. n = 79 porcs.

CONCLUSION

Les travaux réalisés au sein du programme européen Feed-a-Gene, et notamment ceux décrits dans cette synthèse sur la recherche de biomarqueurs de l'efficacité alimentaire du porc en croissance, confirment que l'efficacité alimentaire est un caractère complexe car mettant en œuvre la participation d'une grande variété de molécules impliquées dans différentes fonctions biologiques. Les mesures de l'efficacité alimentaire, comme l'IC et la CMJR correspondraient à des caractères au moins en partie différents, puisque les biomarqueurs identifiés comme de possibles candidats ne sont pas identiques pour ces deux caractères. Les résultats obtenus suggèrent aussi que la mesure à haut débit de l'expression des gènes est sans doute plus efficace que la mesure des concentrations des métabolites

circulants si l'on souhaite proposer des molécules discriminantes ou prédictives de l'efficacité alimentaire. Cela s'explique par le fait que les variations de concentrations des métabolites sont sans doute plus sensibles aux conditions externes à l'animal, et notamment aux différences de régimes alimentaires. Il reste encore beaucoup de travail pour convertir ces preuves de concepts quant à la possibilité d'identifier des biomarqueurs de caractères complexes comme l'efficacité alimentaire, en outils et proxys réellement utilisables. Il s'agit en particulier de vérifier le pouvoir discriminant ou prédictif des molécules identifiées en étudiant des animaux moins extrêmes dans leurs différences d'efficacité alimentaire, et notamment des croisements commerciaux. Il s'agit aussi d'étudier si ces marqueurs pourraient être utilisables pour du tri en élevage basé sur le phénotype mais aussi pour un usage en sélection basée sur une valeur génétique. Il s'agit enfin d'aller vers des conditions de terrain pour intégrer les variations environnementales qui sont souvent des sources de non validation des biomarqueurs découverts en conditions contrôlées.

REMERCIEMENTS

Les auteurs remercient l'ensemble des participants aux différentes études rapportées dans cet article. Le programme Feed-a-Gene a reçu le soutien de la Communauté Européenne au sein de l'action H2020 729 (agrément no 633531). Farouk Messad a reçu une allocation postdoctorale de la Région Bretagne.

REFERENCES BIBLIOGRAPHIQUES

- Bunter KL., Cai W., Johnston DJ., Dekkers JC., 2010. Selection to reduce residual feed intake in pigs produces a correlated response in juvenile insulin-like growth factor-I concentration. *J. Anim. Sci.*, 88, 1973-1981.
- Campos PH., Noblet J., Jaguelin-Peyraud Y., Gilbert H., Mormède P., Donzele RF., Donzele JL., Renaudeau D., 2014. Thermoregulatory responses during thermal acclimation in pigs divergently selected for residual feed intake. *Int. J. Biometeorol.*, 58, 1545-1557.
- Curtasu MV., Bach Knudsen KE., Callesen H., Purup S., Stagsted J., Hedemann MS., 2019. Obesity development in a miniature Yucatan pig model – a multi-compartmental metabolomics study on cloned and normal pigs fed restricted or ad libitum high-energy diets. *J. Proteome Res.*, 18, 30-47.
- Gilbert H., Billon Y., Brossard L., Faure J., Gatellier P., Gondret F., Labussière E., Lebreton B., Lefaucheur L., Le Floc'h N., Louveau I., Merlot E., Meunier-Salaün MC., Montagne L., Mormède P., Renaudeau D., Riquet J., Rogel-Gaillard C., van Milgen J., Vincent A., Noblet J., 2017. Review: divergent selection for residual feed intake in the growing pig. *Animal*, 11, 1427-1439.
- Gilbert H., Terenina E., Ruesche J., Gress L., Billon Y., Mormède P., Larzul C., 2019. Efficacité alimentaire et activité de l'axe corticotrope : comparaison des caractères de production dans des lignées divergentes soumises à un aliment fibreux. *Journées Rech. Porcine*, 51, 321-326.
- Gondret F., Vincent A., Houée-Bigot M., Siegel A., Lagarrigue S., Causeur D., Gilbert H., Louveau I., 2017. A transcriptome multi-tissue analysis identifies biological pathways and genes associated with variations in feed efficiency of growing pigs. *BMC Genomics*, 18:244.
- He B. Li T., Wang W., Gao H., Bai Y., Zhang S., Zang J., Li D., Wang, J., 2018. Metabolic characteristics and nutrient utilization in high feed-efficiency pigs selected using different feed conversion ratio models. *Sci. China Life Sci.*, 61.
- Jansman AJM., van der Peet-Schwering CMC., 2018. Effects of birth weight on nitrogen digestion and utilization in grower pigs. 14th International symposium on digestive physiology in pigs, Brisbane, Australia.
- Jégou M., Gondret F., Lalande-Martin J., Tea I., Baéza E., Louveau I., 2016. NMR-based metabolomics highlights differences in plasma metabolites in pigs exhibiting diet-induced differences in adiposity. *Eur. J. Nutr.*, 55, 1189-1199.
- Jégou M., Gondret F., Vincent A., Tréfeu C., Gilbert H., Louveau I., 2016. Whole blood transcriptomics is relevant to identify molecular changes in response to genetic selection for feed efficiency and nutritional status in the pig. *PLoS One*, 11:e0146550.
- Messad F., Louveau I., Koffi B., Gilbert H., Gondret F., 2019. Investigation of muscle transcriptomes using gradient boosting machine learning identifies molecular predictors of feed efficiency in growing pigs. *BMC Genomics*, 20:659.
- Saintilan R., Brossard L., Vautier B., Sellier P., Bidanel J., van Milgen J., Gilbert H., 2015. Phenotypic and genetic relationships between growth and feed intake curves and feed efficiency and amino acid requirements in the growing pig. *Animal*, 9, 18-27.
- Sarri L., Seradj AR., Tor M., De La Fuente G., Balcells, J., 2019. Effect of productive variety and dietary protein on digestive efficiency and fractional synthesis rate. 70th Annual meeting of the European federation of animal science, Ghent, Belgium.
- Verschuren LMG., Jansman AJM., Calus MPL., Bergsma R., Hedemann MS., 2018. Plasma metabolites related to nitrogen efficiency in low and high birth weight pigs. 14th International symposium on digestive physiology in pigs, Brisbane, Australia.