

UNIVERSIDADE ESTADUAL DE MARINGÁ  
CENTRO DE CIÊNCIAS AGRÁRIAS

ESTUDO DE ASSOCIAÇÃO GENÔMICA AMPLA PARA A  
CARACTERÍSTICA DE EFICIÊNCIA ALIMENTAR EM  
SUÍNOS

Autor: Éverton Ferreira da Silva  
Orientadora: Prof.<sup>a</sup> Dr.<sup>a</sup> Eliane Gasparino

MARINGÁ  
Estado do Paraná  
dezembro - 2017

UNIVERSIDADE ESTADUAL DE MARINGÁ  
CENTRO DE CIÊNCIAS AGRÁRIAS

ESTUDO DE ASSOCIAÇÃO GENÔMICA AMPLA PARA A  
CARACTERÍSTICA DE EFICIÊNCIA ALIMENTAR EM  
SUÍNOS

Autor: Éverton Ferreira da Silva  
Orientadora: Prof.<sup>a</sup> Dr.<sup>a</sup> Eliane Gasparino

Dissertação apresentada, como parte das exigências para obtenção do título de MESTRE EM ZOOTECNIA, no programa de Pós-Graduação em Zootecnia da Universidade Estadual de Maringá – Área de Concentração Produção Animal.

MARINGÁ  
Estado do Paraná  
dezembro - 2017

Dados Internacionais de Catalogação na Publicação (CIP)  
(Biblioteca Central - UEM, Maringá, PR, Brasil)

3586e Silva, Éverton Ferreira da, 1971-  
Estudo de associação genômica ampla para a característica de eficiência alimentar em suínos / Éverton Ferreira da Silva. -- Maringá, 2017.  
xii, 38 f. : figs., tabs.

OrientadorA: Prof.ª Dr.ª Eliane Gasparino.  
Dissertação (mestrado) - Universidade Estadual de Maringá, Centro de Ciências Agrárias, Programa de Pós-Graduação em Zootecnia, 2017.

1. Estudos de associação genômica ampla (GWAS). 2. Polimorfismos de base única (SNPs). 3. Conversão alimentar. 4. Suíno - Ganho de peso diário. 5. Suíno - Consumo de ração. I. Gasparino, Eliane, orient. II. Universidade Estadual de Maringá. Centro de Ciências Agrárias. Programa de Pós-Graduação em Zootecnia III. Título.

CDD 23.ed. 636.4

Gláucia Volponi de Sousa - CRB-9/948



UNIVERSIDADE ESTADUAL DE MARINGÁ  
CENTRO DE CIÊNCIAS AGRÁRIAS

ESTUDO DE ASSOCIAÇÃO GENÔMICA AMPLA  
PARA A CARACTERÍSTICA DE EFICIÊNCIA  
ALIMENTAR EM SUÍNOS

Autor: Éverton Ferreira da Silva  
Orientadora: Profª Drª Eliane Gasparino

TITULAÇÃO: Mestre em Zootecnia - Área de Concentração Produção  
Animal

APROVADA em 19 de dezembro de 2017.

Prof. Dr. Carlos Antonio Lopes de  
Oliveira

Dr. Marcos Soares Lopes

Profª Drª Eliane Gasparino  
(Orientadora)

À memória de meu pai, Adroaldo Martins da Silva, que colocou minha formação acadêmica como um de seus objetivos de vida;

À minha mãe, Tânia Maria Ferreira da Silva, que com seu esforço e exemplos me transmitiu valores que fortaleceram minha formação humana;

À minha avó Terezinha Ferreira, quem, em minha infância, sugeriu-me ser médico veterinário (certamente ela não lembra disso), e que com sua tranquilidade, bom humor e pensamento positivo, cativa a todos que a conhecem;

Aos meus filhos Juliana, Lucas e Laura, os maiores presentes que recebi de Deus e que sempre são um estímulo ao crescimento pessoal e profissional, embora, às vezes, isso represente aumentar a distância física entre nós;

À minha esposa Sanny Duarte, que nem mesmo a distância impediu de nos encontrarmos e compartilhar nossas vidas e que me incentiva na busca de meus objetivos com muito amor e carinho, suportando os períodos de ausência.

DEDICO

## AGRADECIMENTOS

À minha família, pelo apoio incondicional e incentivo constantes.

À Topigs Norsvin, pela oportunidade, pelo apoio financeiro durante as semanas dedicadas à análise dos dados na Holanda e pelas dispensas para dedicar-me ao mestrado.

À Universidade Estadual de Maringá e ao Programa de Pós-Graduação em Zootecnia, pela oportunidade de cumprir nessa instituição uma importante fase de minha formação.

À André Costa, pela oportunidade e todo o apoio durante a realização do curso.

À Profa. Dr.<sup>a</sup> Eliane Gasparino, pela orientação, pelo conhecimento transmitido, pela paciência, pelo incentivo e pela grande amizade.

Ao colega e grande amigo Dr. Marcos Soares Lopes, pelo grande auxílio prestado em várias etapas deste trabalho.

Aos professores do Programa de Pós-Graduação em Zootecnia da UEM, por compartilharem seus conhecimentos e contribuir com esta etapa em minha formação profissional.

Aos professores e amigos Dr. Paulo Sávio Lopes, Dr.<sup>a</sup> Simone Guimarães e Dr. Fabiano Fonseca e Silva, pela amizade, pelo auxílio na discussão do projeto e pelo conhecimento transmitido durante a disciplina cursada na Universidade Federal de Viçosa.

Aos colegas do curso de pós-graduação, agradeço as trocas de experiências, discussões e companheirismo, principalmente àqueles com os quais tive a oportunidade de desenvolver algum trabalho em conjunto.

À CAPES, pela bolsa concedida.

## BIOGRAFIA

ÉVERTON FERREIRA DA SILVA, filho de Adroaldo Martins da Silva e Tânia Maria Ferreira da Silva, nasceu em Rio Grande, Rio Grande do Sul, no dia 28 de janeiro de 1971.

Em janeiro de 1995, concluiu o curso de Medicina Veterinária pela Universidade Federal de Pelotas – UFPEL.

Em março de 1995, foi contratado pela Cooperativa Agrícola Mista Rondon Ltda - COPAGRIL, de Marechal Cândido Rondon, Paraná, atuando como extensionista do fomento de suínos até dezembro de 1998.

Em dezembro de 1998, foi contratado pela Cooperativa Central Agropecuária Sudoeste Ltda - FRIMESA, também em Marechal Cândido Rondon, Paraná, onde desenvolveu atividades como suporte técnico e sanitário do fomento de suínos até fevereiro de 2000.

Em fevereiro de 2000, foi contratado pela Dalland do Brasil Agropecuária Ltda (atual Topigs Norsvin do Brasil Ltda) exercendo atualmente a função de Gerente de Genética Latam, sendo responsável pelo programa de melhoramento genético da empresa no Brasil e gerenciando o programa de melhoramento genético nos países da América Central e América do Sul.

Em março de 2015, iniciou no Programa de Pós-Graduação em Zootecnia, em nível de Mestrado, área de concentração Produção Animal, na Universidade Estadual de Maringá - UEM, realizando estudos na área de melhoramento genético de suínos. No dia 19 de dezembro de 2017, submeteu-se à banca para defesa da Dissertação.



## ÍNDICE

	Página
LISTA DE TABELAS .....	vii
LISTA DE FIGURAS .....	viii
RESUMO .....	ix
ABSTRACT .....	xi
I. INTRODUÇÃO .....	1
1. A eficiência alimentar (EA) em suínos .....	2
2. Seleção para eficiência alimentar em suínos .....	3
3. Polimorfismos de base única (SNPs) .....	4
4. Estudos de associação genômica ampla (GWAS) .....	5
5. Referências .....	7
III. Associação genômica ampla para características relacionadas à eficiência alimentar em uma população de suínos cruzados .....	12
Resumo .....	12
Implicações .....	13
Introdução .....	13
Material e métodos .....	15
Resultados .....	20
Discussão .....	22
Conclusões .....	27
Agradecimentos .....	28
Referências .....	28

## LISTA DE TABELAS

	Página
<b>Tabela 1</b> Estatística descritiva dos grupos de informações fenotípicas e genotípicas para as características de ganho de peso médio diário (ADG), consumo de ração médio diário (ADFI) e conversão alimentar (FCR) .....	33
<b>Tabela 2</b> Componentes de variância e herdabilidades estimadas para as características de ganho de peso médio diário (ADG), consumo de ração médio diário (ADFI) e conversão alimentar (FCR) utilizando fenótipos pré-corrigidos .....	34
<b>Tabela 3</b> Caracterização dos SNPs com efeito significativo sobre as características de ganho de peso médio diário (ADG) e consumo de ração médio diário (ADFI) .....	35
<b>Tabela 4</b> SNPs com efeito significativo sobre as características de ganho de peso médio diário (ADG) e consumo de ração médio diário (ADFI), regiões QTL associadas a estas características e genes candidatos .....	36

## LISTA DE FIGURAS

	Página
<b>Figura 1</b> Gráficos quantil-quantil dos P-valores dos SNPs para as características de ganho de peso médio diário (ADG), consumo de ração médio diário (ADFI) e conversão alimentar (FCR). .....	37
<b>Figura 2</b> Estudo de associação genômica de características relacionadas à eficiência alimentar na fase de recria/terminação, representado pelo $-\log_{10}$ (P-valor) da associação do SNP com as características de ganho de peso médio diário (ADG), consumo de ração médio diário (ADFI) e conversão alimentar (FCR) (eixo y), sobre a posição física dos SNPs nos 18 cromossomos autossômicos (eixo x). .....	38

## RESUMO

A eficiência alimentar (EA), característica que relaciona o ganho de peso do animal à quantidade de ração consumida, é considerada uma das mais importantes características na produção de suínos, pois está relacionada ao custo de produção da carne suína. A avaliação dessa característica em grande número de indivíduos é difícil de ser realizada e apresenta custo elevado, contudo, utilizar estratégias e tecnologias que possam auxiliar em seu progresso genético é de extrema importância e interesse dos programas de melhoramento genético dos suínos. Nesse contexto, as estações eletrônicas de alimentação têm sido utilizadas em programas de melhoramento genético de suínos e têm contribuído com a coleta de informações de consumo de ração em maior número de animais. Em geral, informações de EA são coletadas em animais puros e utilizadas com o objetivo de melhorar essa característica em animais cruzados. Entretanto, espera-se que a utilização de informações de animais cruzados comerciais para compor a população de referência melhore a acurácia das informações, pois geralmente os rebanhos puros estão alojados sob condições de sanidade e manejo superiores as que estão submetidos os animais comerciais. A utilização de densos painéis comerciais de SNPs (polimorfismos de base única) em GWAS (estudos de associação genômica ampla) tem demonstrado ser uma estratégia eficiente para identificar genes ou regiões genômicas que explicam variação em características de interesse. Esses GWAS são utilizados para detectar associações significativas entre marcadores SNPs ou regiões genômicas com as características investigadas graças ao desequilíbrio de ligação (LD) entre SNPs e as mutações causais para estas características. Dessa forma, este estudo foi realizado com o objetivo de identificar marcadores do tipo SNP, com associação significativa com características relacionadas à EA, além de identificar potenciais genes candidatos para as características de ganho de peso médio diário (ADG), consumo de ração médio diário

(ADFI) e conversão alimentar (FCR). Foram coletadas informações de uma população de 2.386 suínos cruzados, genotipados para 51.468 marcadores SNPs. Associações significativas ( $FDR \leq 0,005$ ) entre SNPs e ADG foram identificadas entre 177,01 e 185,47 Mb no cromossomo suíno 1 (SSC1), e entre 2,87 e 3,22 Mb no cromossomo suíno 12 (SSC12). Associações significativas entre SNPs e ADFI também foram identificadas entre 173,26 e 185,47 Mb no SSC1. Embora associações significativas entre SNPs e as características ADG e ADFI foram identificadas na mesma região do SSC1 (177,01 - 185,47 Mb), não foi identificada associação significativa dessa região com a característica FCR, podendo esse fato ser explicado pela complexidade dessa característica que é obtida pela divisão do ADFI pelo ADG. Os genes candidatos identificados nas regiões com associações significativas com as características investigadas foram: *CDH19*, *CDH7*, *RNF152*, *MC4R*, *PMAIP1*, *FEM1B* e *GAA*. Dentre esses genes, destaca-se o gene *MC4R*, conhecidamente associado às características ADG, ADFI e espessura de toucinho em suínos. Neste trabalho, foram identificadas regiões QTL fortemente associadas com ADG e ADFI no SSC1 e com ADG no SSC12 em uma população de suínos cruzados. Essas regiões haviam sido previamente descritas em populações de linhagens puras de suínos, porém, este é o primeiro estudo a confirmar a relevância dessas regiões em uma população de suínos cruzados. Os SNPs e genes candidatos identificados neste estudo poderão ser avaliados quanto ao seu potencial para utilização em modelos de predição aliando seleção genômica ampla com seleção assistida por marcadores/genes.

**Palavras-chave:** GWAS, SNP, conversão alimentar, ganho de peso diário, consumo de ração

## ABSTRACT

Feed efficiency (FE) is a trait that relates the weight gain of an animal to the amount of consumed feed to achieve this gain. The FE is considered one of the most important traits in pork production because it is directly related to the production cost. The evaluation of FE on a large number of animals is difficult to be carried out and it is expensive. Nevertheless, applying strategies and technologies that may aid the genetic progress of FE is of utmost importance and interest of pig breeding programs. In this context, electronic feeding stations has been applied in pig breeding and it has contributed to the recording of feed consumption information on a large number of animals. In general, FE information is recorded on purebred animals and used to improve FE in crossbred animals. However, using commercial crossbred information to compose the reference population is expected to improve the information accuracy because purebred populations are usually kept under superior health and management conditions when compared to crossbred commercial populations. The use of single nucleotide polymorphisms (SNPs) in genome-wide association studies (GWAS) has been pointed out as an efficient strategy to identify genes and genomic regions that explain the phenotypic variation of the traits of interest. The GWAS are used to identify association between SNPs and the evaluated traits due to the linkage disequilibrium (LD) between the SNPs and the causal mutations that influence these traits. Therefore, the aim of this study was to identify SNPs significantly associated with FE-related traits and to identify potential candidate genes controlling the traits of average daily gain (ADG), average daily feed intake (ADFI) and feed conversion rate (FCR). Data was available on a crossbred pig population with 2,386 animals genotyped for 51,468 SNPs. Significant associations (FDR  $\leq 0.005$ ) between SNPs and ADG were identified between 177.01 and 185.47 Mb on *Sus scrofa* chromosome 1 (SSC1) and between 2.87 and 3.22 Mb on SSC12. Significant associations

between SNPs and ADFI were identified between 173.26 and 185.47 Mb on SSC1. Although the same region on SSC1 (177.01-185.47 Mb) was found to be significant for both ADG and ADFI traits, no SNPs in this region showed significant association with FCR, that could be explained by the complexity of this trait which is a ratio between ADFI and ADG. The candidate genes identified in the regions with significant associations with the evaluated traits were: *CDH19*, *CDH7*, *RNF152*, *MC4R*, *PMAIP1*, *FEM1B* and *GAA*. Among these genes, we can highlight *MC4R* which is known to affect the traits ADG, ADFI and backfat thickness in pigs. In this study, SNPs with significant association with ADG and ADFI on SSC1 and with ADG on SSC12 were identified in a crossbred pig population. The same regions have been described to be associated with these traits in previous studies using purebred pig populations. However, this is the first study to confirm the relevance of these genomic regions for these traits in a crossbred pig population. The potential use of the SNPs and candidate genes identified in this study should be further evaluated in prediction models in a genomic selection scheme associated with marker/gene-assisted selection.

**Key words:** GWAS, SNP, feed conversion, average daily gain, feed intake

## I. INTRODUÇÃO

O suíno começou a ser domesticado há aproximadamente 9.000 anos e as linhagens comerciais atuais são originárias de cruzamentos entre dois grupos independentes formados por raças de suínos selvagens procedentes da Europa e da Ásia, conforme demonstrado por estudos com DNA mitocondrial (Giuffra *et al.*, 2000; Megens *et al.*, 2008; Amills *et al.*, 2010). Os genótipos asiáticos foram introduzidos na Europa a partir do século XVII (White, 2011) e informações históricas indicam sua utilização para melhoria de características dos genótipos europeus durante o século XVIII e início do século XIX (Giuffra *et al.*, 2000; White, 2011).

O período da introdução do suíno asiático na Europa coincidiu com a emergência da revolução industrial, época em que ocorreram importantes transformações na atividade agropecuária, motivada pelo aumento da atividade comercial entre os continentes asiático e europeu. Esse fato demonstra o importante papel que a introdução das linhagens asiáticas na Europa exerceu na mudança do perfil produtivo de subsistência para o perfil de produção comercial (White, 2011).

No Brasil, os primeiros suínos foram introduzidos por Martin Afonso de Souza no ano de 1532, que trouxe de Portugal as raças Alentejana, Transtagana, Galega, Bizarra, Beiroa e Macau (Fávero e Pereira de Figueiredo, 2009). Nos séculos seguintes, essas raças deram origem às raças naturalizadas brasileiras, como a Piau. A partir do final da década de 1950 e durante a década de 1960, começaram a ser introduzidas no Brasil raças de maior potencial produtivo originárias de vários países com o objetivo de melhorar a produção da suinocultura nacional, uma vez que a banha havia perdido espaço para os óleos vegetais (Fávero e Pereira de Figueiredo, 2009). Na década de 1970, por iniciativa da Associação Brasileira de Criadores de Suínos (ABCS), começaram a ser realizadas avaliações de reprodutores nas Estações de Testes de Reprodutores de Suínos (ETRS),



passando a ser medidas as características de ganho de peso médio diário, conversão alimentar e espessura de toucinho dos animais. Dessa forma, a grande ênfase dada à seleção para essas características levou a ganhos genéticos consideráveis (Lopes, 2010).

A partir de meados da década de 1970 até o final da década de 1990, instalaram-se no Brasil, várias empresas especializadas em melhoramento genético de suínos, oriundas da Europa, Estados Unidos e Canadá (Lopes, 2010), que naquela época, trouxeram de seus países de origem um material genético moderno e competitivo (Fávero e Pereira de Figueiredo, 2009). Nessa mesma época, também foram formados alguns programas nacionais de melhoramento genético em suínos, sendo que alguns persistem ativos até os dias atuais.

A evolução no melhoramento genético obtida nos últimos anos possibilitou ao Brasil, tornar-se um dos maiores produtores e exportadores mundiais de carne suína. A alta competitividade da cadeia de produção dessa proteína, exige a continuidade da aplicação de práticas que possibilitem a melhoria na eficiência dos suínos, seja através da melhoria de sua eficiência reprodutiva, buscando melhorar características relacionadas, por exemplo, com prolificidade e robustez, assim como características que estejam relacionadas com a eficiência produtiva, como o ganho de peso, a eficiência alimentar e a qualidade de carcaças (carne magra e espessura de toucinho).

## 1. A eficiência alimentar (EA) em suínos

A Eficiência Alimentar (EA) é considerada uma das mais importantes características na produção de suínos, relacionando o ganho de peso do animal com a quantidade de alimento ingerido. Sua importância é justificada principalmente pelo fato dos custos com alimentação representarem entre 60 e 70% do custo total de produção da carne suína (Patience *et al.*, 2015).

As ações relacionadas com a redução no custo de produção têm contribuído na redução do custo final de produtos como carne, leite e ovos, possibilitando à população mundial maior acesso à proteína de alta qualidade (Hayes *et al.*, 2013). Além disso, o desperdício de rações utilizadas para alimentação animal constitui importante impacto ambiental, e também tem motivado produtores a buscar alternativas para redução desse desperdício e aumento da EA (Rothschild, 2004).

A EA é um fator crítico em vários níveis dentro da cadeia de produção da carne suína, implicando na lucratividade das granjas, na competitividade da indústria em relação às demais fontes de proteína, na demanda global de alimentos, além de aspectos ambientais e sociais (Patience *et al.*, 2015).

Tipicamente a EA é medida como conversão alimentar (CA) que é definida como a quantidade de ração consumida para produção de uma unidade de ganho de peso do animal.

Existem variações substanciais no consumo individual de ração em suínos e aproximadamente 30% dessa variação é explicada pela variação genética (Cai *et al.*, 2008).

A medição do consumo de ração de grande número de indivíduos é difícil de ser realizada e apresenta alto custo (Jiao *et al.*, 2014) e esse é um dos motivos pelo qual em muitos casos a CA é melhorada através de dados fenotípicos indiretos, como o ganho de peso e características relacionadas à composição corporal (por exemplo, espessura de toucinho e profundidade de lombo).

## 2. Seleção para eficiência alimentar em suínos

De acordo com Merks (2000), durante o período entre os anos de 1900 e 1960, os programas de melhoramento genético regionais da Europa focaram somente nas características de carne magra e eficiência alimentar, embora alguns tenham enfatizado características como aptidão, qualidade de aprumos e características raciais. Os animais de linhagens puras começaram a ser testados com o objetivo principal de comparar seus valores genéticos para ganho de peso, conversão alimentar e espessura de toucinho sob condições padronizadas.

A partir do início da década de 1960, iniciou-se a especialização em linhagens maternas e paternas, e levou às linhagens paternas ao aumento no progresso genético das características produtivas, como ganho de peso, eficiência alimentar e espessura de toucinho; e as linhagens maternas ao aumento no progresso genético das características reprodutivas, como tamanho de leitegada (Merks, 2000). Com essa especialização, as empresas de melhoramento deram origem ao cruzamento das linhagens paternas e maternas para originar o animal comercial moderno (um animal cruzado).

As alterações genéticas ocorridas nos suínos, entre os anos 1900 e 2000, resultaram principalmente no progresso genético das características de ganho de peso (+100%) e na espessura de toucinho (-75%) (Merks, 2000).

Tornar o suíno mais magro é uma das formas de melhorar a eficiência alimentar. Essa estratégia indireta de seleção foi utilizada no melhoramento genético de suínos durante décadas, entretanto Knap e Wang (2012) reforçaram que aproximadamente um terço da variação no consumo de ração é devido a processos que não dependem do crescimento e da composição corporal, principalmente aqueles ligados à manutenção corporal, como o metabolismo basal, *turnover* proteico, termorregulação, atividade física, sistema imune, além da eficiência de absorção e digestão de nutrientes.

A estimativa de valores genéticos para consumo de ração, ao invés de CA, possibilita maior controle sobre outros efeitos. Ademais, a coleta de dados de consumo de ração em animais alojados em condições comerciais tende a aumentar a acurácia da estimativa dos valores genéticos, em virtude de que o consumo de ração é mais sensível às variações ambientais (Knap e Wang, 2012), que geralmente são menores nas granjas núcleos quando comparadas às granjas comerciais.

O uso das estações de alimentação automatizadas tem permitido aumentar o registro de informações de consumo de alimento. Desde que elas foram introduzidas, as empresas de melhoramento genético de suínos tem conseguido considerável progresso genético na conversão alimentar, principalmente pelo progresso genético sobre as características de crescimento e de composição corporal (Knap e Wang, 2012).

### 3. Polimorfismos de base única (SNPs)

Os polimorfismos de base única (SNPs) são o mais abundante tipo de variação no DNA do genoma vertebrado (Matsumoto *et al.*, 2012), representando uma das abordagens mais interessantes na genotipagem, pois além de abundantes, são geneticamente estáveis e acessíveis em análises automatizadas de alto rendimento (Yadav *et al.*, 2017).

Para ser considerado como um SNP, o alelo menos frequente do marcador de DNA deve apresentar abundância de pelo menos 1% na população (Brookes, 1999; Vignal *et al.*, 2002). Ainda que, em princípio, em cada posição de uma sequência de DNA alguma das quatro possíveis bases nucleotídicas possam estar presentes, os SNPs são na prática bialélicos (Vignal *et al.*, 2002).

Por volta do ano de 2009, os marcadores SNPs tornaram-se disponíveis comercialmente para genotipagem de suínos (Ramos *et al.*, 2009; Knol *et al.*, 2016). A criação de densos painéis comerciais de SNPs, combinada com a utilização de métodos estatísticos para estimativa dos valores genéticos, resultou na aplicação extensiva da seleção genômica ampla (GWAS) em bovinos e posteriormente em outras espécies (Dekkers, 2012).

A utilização de SNPs de alta densidade em GWAS tem demonstrado ser uma alternativa eficiente na identificação de genes ou regiões genômicas que explicam variação em características pecuárias (Onteru *et al.*, 2013).

#### 4. Estudos de associação genômica ampla (GWAS)

Os estudos de associação genômica ampla (GWAS) baseiam-se na possibilidade de detectar associações significativas entre marcadores SNPs ou regiões genômicas com determinadas características. Isso é possível caso haja desequilíbrio de ligação (LD) entre SNPs e mutações causais para características de interesse (Dekkers, 2012; Zhang *et al.*, 2012).

De acordo com Pearson e Manolio (2008), o GWAS típico possui 4 partes: (1) seleção de grande número de indivíduos com a característica de interesse e um grupo de comparação; (2) isolamento do DNA, genotipagem e revisão dos dados para assegurar a alta qualidade da genotipagem; (3) teste estatístico para associações entre os SNPs que passaram pelos limiares de qualidade e a característica; e (4) replicação das associações identificadas em uma amostra da população independente ou exame das implicações funcionais experimentalmente.

O procedimento de controle de qualidade das informações deve ser realizado antes do GWAS, excluindo-se animais com falhas de genotipagem  $>0,05$ , além dos SNPs localizados nos cromossomas sexuais, com posição desconhecida no genoma suíno (Groenen *et al.*, 2012), com falhas de genotipagem  $>0,05$ , ou com forte desvio no equilíbrio de Hardy-Weinberg ( $\chi^2 >600$ ) (Hayes *et al.*, 2009; Resende *et al.*, 2012; Lopes *et al.*, 2013; Do *et al.*, 2015).

Vários métodos estatísticos têm sido utilizados para GWAS, entretanto muitos estudos têm utilizado o modelo *single*-SNP no qual cada SNP é ajustado a cada vez no

modelo, em geral como um efeito fixo (Dekkers, 2012; Lopes, 2016). O P-valor obtido para cada SNP é utilizado como medida da força da associação.

O uso do P-valor tem sido considerado uma importante limitação dessa metodologia de GWAS (Sham e Purcell, 2014). Sua interpretação é difícil sem considerar o poder estatístico da análise (por exemplo, o tamanho da amostra e a menor frequência alélica) (Sham e Purcell, 2014). Apesar dessas limitações, o *single*-SNP GWAS tem sido amplamente utilizado com sucesso na identificação de regiões QTL (Li *et al.*, 2011).

Diferenças artificiais na frequência alélica devido à estratificação da população e erros de genotipagem podem afetar todos os SNPs, e dessa forma, o teste estatístico pode ser inflado em todo o genoma (Yang *et al.*, 2011). Uma forma de evitar falsos positivos devido à estratificação populacional é a aplicação de modelos mistos e a utilização da matriz de parentesco (Price *et al.*, 2010).

O controle genômico (estimativa do fator de inflação  $\lambda$ ) é outra estratégia que pode ser aplicada para avaliar se os procedimentos utilizados durante a GWAS foram suficientes para evitar resultados falsos positivos (Hinrichs *et al.*, 2009; Price *et al.*, 2010).

Após o GWAS, a apresentação dos P-valores obtidos e esperados em gráficos quantil-quantil, ordenados de menor a maior, permite a avaliação visual da normalidade dos dados e identificação de possível inflação dos P-valores (Diniz, 2013).

A FDR (*False Discovery Rate*), definida como a proporção da rejeição da hipótese nula quando ela é verdadeira (Benjamini e Hochberg, 1995), é estimada utilizando os P-valores para evitar falsos positivos devido à múltipla testagem. Os resultados da GWAS normalmente são apresentados na forma de gráficos (*Manhattan plot*), onde os SNPs são dispostos de acordo com sua posição em cada cromossomo, relacionados ao seu P-valor para a característica alvo do GWAS, destacando àqueles com FDR acima do limiar de significância.

Alguns GWAS com suínos de diferentes linhagens têm identificado regiões de QTL e SNPs relacionados com EA (Houston *et al.*, 2005; Zhang *et al.* 2012; Sahana *et al.*, 2013), contudo são necessários mais estudos que auxiliem na identificação de outros marcadores ou que possibilitem reforçar a associação dos marcadores já identificados com a expressão dessa característica ou das características que constituem a EA.

## 5. Referências

- Amills M, Clop A, Ramírez O and Pérez-Encizo M 2010. Origin and genetic diversity of pig breeds. In *Encyclopedia of Life Sciences (ELS)* (ed. John Wiley & Sons, Ltd), 1-10. Chichester. doi: 10.1002/9780470015902.a0022884.
- Benjamini Y and Hochberg Y 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 57, 289-300.
- Brookes AJ 1999. The essence of SNPs. *Gene* 234, 177-186.
- Cai W, Casey DS and Dekkers JCM 2008. Selection response and genetic parameters for residual feed intake in Yorkshire swine. *Journal of Animal Science* 86, 287–298.
- Dekkers JCM 2012. Application of genomics tools to animal breeding. *Current Genomics* 13, 207-212.
- Diniz DB 2013. Genome-wide association study for sperm motility in pigs. MSc thesis, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brasil.
- Do DN, Janss LLG, Jensen J and Kadarmideen HN 2015. SNP annotation-based whole genomic prediction and selection: An application to feed efficiency and its component traits in pigs. *Journal of Animal Science* 93, 2056–2063.
- Fávero JA and Pereira de Figueiredo EA 2009. Evolução do melhoramento genético de suínos no Brasil. *Revista Ceres* 56, 420-427.
- Giuffra E, Kijas JMH, Amarger V, Carlborg Ö, Jean J-T and Andersson L 2000. The origin of the domestic pig: independent domestication and subsequent introgression. *Genetics* 154, 1785-1791.
- Groenen MAM, Archibald AL, Uenishi H, Tuggle CK, Takeuchi Y, Rothschild MF, Rogel-Gaillard C, Park C, Milan D, Megens H-J, Li S, Larkin DM, Kim H, Frantz LAF, Caccamo M, Ahn H, Aken BL, Anselmo A, Anthon C, Auvil L, Badaoui B, Beattie CW, Bendixen C, Berman D, Blecha F, Blomberg J, Bolund L, Bosse M, Botti S, Bujie Z, Bystrom M, Capitanu B, Carvalho-Silva D, Chardon P, Chen C, Cheng R, Choi S-H, Chow W, Clark RC, Clee C, Crooijmans RPMA, Dawson HD, Dehais P, De Sapia F, Dibbits B, Drou N, Du Z-Q, Eversole K, Fadista J, Fairley S, Faraut T, Faulkner GJ, Fowler KE, Fredholm M, Fritz E, Gilbert JGR, Giuffra E, Gorodkin J, Griffin DK, Harrow JL, Hayward A, Howe K, Hu Z-L, Humphray SJ, Hunt T, Hornshøj H, Jeon J-T, Jern P, Jones M, Jurka J, Kanamori H, Kapetanovic R, Kim J, Kim J-H, Kim KW, Kim T-H, Larson G, Lee K, Lee K-T, Leggett R, Lewin HA, Li Y, Liu W, Loveland JE, Lu Y, Lunney JK, Ma J, Madsen O, Mann K, Matthews L,

- McLaren S, Morozumi T, Murtaugh MP, Narayan J, Nguyen DT, Ni P, Oh S-J, Onteru S, Panitz F, Park E-W, Park H-S, Pascal G, Paudel Y, Perez-Enciso M, Ramirez-Gonzalez R, Reecy JM, Rodriguez-Zas S, Rohrer GA, Rund L, Sang Y, Schachtschneider K, Schraiber JG, Schwartz J, Scobie L, Scott C, Searle S, Servin B, Southey BR, Sperber G, Stadler P, Sweedler JV, Tafer H, Thomsen B, Wali R, Wang J, Wang J, White S, Xu X, Yerle M, Zhang G, Zhang J, Zhang J, Zhao S, Rogers J, Churcher C and Schook LB 2012. Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* 491, 393-398.
- Hayes BJ, Bowman PJ, Chamberlain AJ and Goddard ME 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. *Journal of Dairy Science* 92, 433-443.
- Hayes BJ, Harris AL and Goddard ME 2013. The future of livestock breeding: genomic selection for efficiency, reduced emissions intensity, and adaptation. *Trends in genetics* 29, 206-214.
- Hinrichs AL, Larkin EK and Suarez BK 2009. Population stratification and patterns of linkage disequilibrium. *Genetic Epidemiology* 33, S88-S92. doi: 10.1002/gepi.20478.
- Houston RD, Haley CS, Archibald AL and Rance KA 2005. A QTL affecting daily feed intake maps to chromosome 2 in pigs. *Mammalian Genome* 16, 464-470.
- Jiao S, Maltecca C, Gray KA and Cassady JP 2014. Feed intake, average daily gain, feed efficiency, and real-time ultrasound traits in Duroc pigs: II. Genomewide association. *Journal of Animal Science* 92, 2846-2860.
- Knap PW and Wang L 2012. Pig breeding for improved feed efficiency. In *Feed efficiency in swine* (ed. J. F. Patience), pp. 167-179. Wageningen Academic Publishers, Wageningen, the Netherlands.
- Knol EF, Nielsen B and Knap PW 2016. Genomic selection in commercial pig breeding. *Animal Frontiers* 6, 15-22.
- Li J, Das K, Fu G, Li R and Wu R 2011. The Bayesian lasso for genome-wide association studies. *Bioinformatics* 27, 516-523.
- Lopes PS 2010. Melhoramento genético de suínos. Retrieved on 15 october 2017, from <http://sbmaonline.org.br/anais/viii/palestras/pdfs/3.pdf>.
- Lopes MS, Silva FF, Harlizius B, Duijvesteijn N, Lopes PS, Guimarães SEF and Knol EF 2013. Improved estimation of inbreeding and kinship in pigs using optimized SNP panels. *BMC Genetics* 14, 92. doi: 10.1186/1471-2156-14-92.
- Lopes MS 2016. Genomic selection for improved crossbred performance. Phd Thesis, Wageningen University, Wageningen, the Netherlands.

- Matsumoto T, Okumura N, Uenishi H, Hayashi T, Hamasima N and Awata T 2012. Porcine single nucleotide polymorphism (SNP) development and population structure of pigs assessed by validated SNPs. *Biochemical Genetics* 50, 428–439.
- Megens H-J, Crooijmans RPMA, San Cristobal M, Hui X, Li N and Groenen MAM 2008. Biodiversity of pig breeds from China and Europe estimated from pooled DNA samples: differences in microsatellite variation between two areas of domestication. *Genetic Selection Evolution* 40, 103-128.
- Merks JWM 2000. One century of genetic changes in pigs and the future needs. British Society of Animal Science, occasional publication, 8-19.
- Onteru SK, Gorbach DM, Young JM, Garrick DJ, Dekkers JCM and Rothschild MF 2013. Whole genome association studies of residual feed intake and related traits in the pig. *Plos One* 8, e61756.
- Patience JF, Rossoni-Serão MC and Gutiérrez NA 2015. A review of feed efficiency in swine: biology and application. *Journal of Animal Science and Biotechnology* 6, 33. doi: 10.1186/s40104-015-0031-2.
- Pearson TA and Manolio TA 2008. How to interpret a genome-wide association study. *JAMA: Journal of American Medical Association* 299, 1335-1344.
- Price AL, Zaitlen NA, Reich D and Patterson N 2010. New approaches to population stratification in genome-wide association studies. *Nature Review Genetics* 11, 459-463.
- Ramos AM, Crooijmans RPMA, Affara NA, Amaral AJ, Archibald AL, Beever JE, Bendixen C, Churcher C, Clark R, Dehais P, Hansen MS, Hedegaard J, Hu Z-L, Kerstens HH, Law AS, Megens H-J, Milan D, Nonneman DJ, Rohrer GA, Rothschild MF, Smith TPL, Schnabel RD, Tassell CPV, Taylor JF, Wiedmann RT, Schook LB and Groenen MAM 2009. Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. *Plos One* 4, e6524.
- Resende MDV, Silva FF, Lopes PS and Azevedo CF 2012. Seleção genômica ampla (GWS) via modelos mistos (REML/BLUP), inferência bayesiana (MCMC), regressão aleatória multivariada e estatística espacial. Universidade Federal de Viçosa/Departamento de Estatística, Viçosa, Minas Gerais, Brasil.
- Rothschild MF 2004. Porcine genomics delivers new tools and results: This little piggy did more than just go to market. *Genetics Research* 83, 1-6. doi: 10.1017/S0016672303006621.
- Sahana G, Kadlecová V, Hornshøj H, Nielsen B and Christensen OF 2013. A genome-wide association scan in pig identifies novel regions associated with feed efficiency trait. *Journal of Animal Science* 91, 1041-1050.



- Sham PC and Purcell SM 2014. Statistical power and significance testing in large-scale genetic studies. *Nature Reviews Genetics* 15, 335-346.
- Vignal A, Milan D, SanCristobal M and Eggen A 2002. A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution* 34, 275-305.
- White S 2011. From globalized pig breeds to capitalist pigs: a study in animal cultures and evolutionary history. *Environmental History* 16, 94–120.
- Yadav AK, Tomar SS, Jha AK and Singh J 2017. Importance of molecular markers in livestock improvement: a review. *International Journal of Agriculture Innovations and Research* 5, 614-621.
- Yang J, Weedon MN, Purcell S, Lettre G, Estrada K, Willer CJ, Smith AV, Ingelsson E, O'Connell JR, Mangino M, Mägi R, Madden PA, Heath AC, Nyholt DR, Martin NG, Montgomery GW, Frayling TM, Hirschhorn JN, McCarthy MI, Goddard ME, Visscher PM and the GIANT Consortium 2011. Genomic inflation factors under polygenic inheritance. *European Journal of Human Genetics* 19, 807–812.
- Zhang H, Wang Z, Wang S and Li H 2012. Progress of genome wide association study in domestic animals. *Journal of Animal Science and Biotechnology* 3, 26. doi: 10.1186/2049-1891-3-26.

## II. OBJETIVOS GERAIS

Identificar marcadores do tipo SNP com associação significativa com características relacionadas à eficiência alimentar, utilizando GWAS (estudo de associação genômica ampla) em suínos cruzados.

Detectar potenciais genes candidatos para as características de ganho de peso médio diário, consumo de ração médio diário e conversão alimentar.

### III. Associação genômica ampla para características relacionadas à eficiência alimentar em uma população de suínos cruzados

(Revista – Animal – An International Journal of Animal Bioscience)

Título curto: GWAS para eficiência alimentar em suínos cruzados

#### Resumo

A eficiência alimentar (**EA**) é considerada uma das mais importantes características na produção de suínos. Porém, a mensuração desse fenótipo é difícil e de custo elevado, limitando a coleta de grandes volumes de dados para uma acurada seleção para melhor EA. Portanto, a identificação de marcadores moleculares relacionados à EA que possam ser utilizados na avaliação genética é de grande interesse dos programas de melhoramento de suínos, a fim de aumentar a acurácia de seleção e o progresso genético dessa característica. Neste estudo, objetivou-se identificar marcadores SNPs com associação significativa com características relacionadas à EA, além de identificar potenciais genes candidatos para as características avaliadas: ganho de peso médio diário (**ADG**), consumo de ração médio diário (**ADFI**) e conversão alimentar (**FCR**). Essas informações foram coletadas em uma população de 2.386 suínos cruzados que foram genotipados para 51.468 marcadores **SNPs** (polimorfismos de base única). Foram identificadas regiões de **QTL** (*locus* de características quantitativas) com forte associação ao ADG, localizadas entre 177,01 e 185,47 Mb no cromossomo suíno 1 (SSC1), com  $FDR \leq 0,005$ , e entre 2,87 e 3,22 Mb no cromossomo suíno 12 (SSC12), com  $FDR \leq 0,005$ . Também evidenciou-se associações entre ADFI com a região QTL entre 173,26 e 185,47 Mb no SSC1 ( $FDR \leq 0,005$ ). Não se evidenciou associação significativa ( $FDR \leq 0,005$ ) com a característica FCR, embora tenha-se evidenciado forte associação da região QTL entre 177,01 e 185,47 Mb do SSC1 com as características ADG e ADFI, que compõem a característica FCR. A não identificação de associação de região genômica com a FCR pode ser explicada pela complexidade dessa característica. Além disso, identificou-se os genes *CDH19*, *CDH7*, *RNF152*, *MC4R*, *PMAIP1*, *FEM1B* e *GAA*, localizados nas janelas de 1Mb das regiões

QTLs com significativa associação identificadas neste estudo. Dentre eles, destaca-se o gene *MC4R*, conhecidamente associado às características ADG, ADFI e espessura de toucinho. Foram identificadas regiões de QTL com forte associação com as características ADG e ADFI no SSC1 e com o ADG no SSC12 em uma população de suínos cruzados. Essas regiões foram previamente descritas em populações de linhagens puras de suínos, porém, este é o primeiro estudo a confirmar a relevância dessas regiões de QTL em uma população de suínos cruzados. Os genes candidatos e SNPs identificados neste estudo poderão ser avaliados quanto ao seu potencial para utilização em modelos de predição, aliando seleção genômica ampla com seleção assistida por marcadores.

**Palavras-chave:** GWAS, SNP, conversão alimentar, ganho de peso diário, consumo de ração

### **Implicações**

A EA é uma das mais importantes características na produção de suínos. Animais com melhor EA produzirão mais carne por unidade de ração consumida. Melhorar a EA significa reduzir o custo de produção, possibilitando o acesso de uma maior parte da população mundial à proteína de alta qualidade. Os GWAS possibilitam identificar marcadores SNPs para características zootécnicas. O uso de marcadores para melhor EA permite aumentar o progresso genético dessa característica. Informações fenotípicas de populações de suínos cruzados criados em ambientes similares aos que os animais normalmente são produzidos, podem ser usadas como referência para seleção dos animais puros nas populações que lhes deram origem.

### **Introdução**

A eficiência alimentar (EA) é considerada uma das mais importantes características na produção de suínos, relacionando o ganho de peso do animal com a quantidade de alimento ingerido. Sua importância é justificada principalmente pelo fato dos custos relacionados com alimentação representarem de 60 a 70% do custo total de produção (Patience *et al.*, 2015).

Porém, além de seu aspecto econômico, uma melhor EA também possui um aspecto social, já que as ações relacionadas com a redução no custo de produção contribuem para redução do custo final de produtos como carne, leite e ovos, possibilitando à população mundial maior acesso à proteína de alta qualidade (Hayes *et al.*, 2013), além de reduzir o impacto ambiental relacionado aos resíduos gerados.

Apesar do grande interesse da indústria suinícola em EA, a mensuração desse fenótipo é difícil e de custo elevado, limitando a coleta de grandes volumes de dados para uma acurada seleção para melhor EA. Dessa forma, a utilização de ferramentas que auxiliem no aumento da acurácia de seleção para melhor EA tem despertado grande interesse dos programas de melhoramento genético de suínos. Uma dessas ferramentas é a utilização de informações genômicas como auxílio à seleção.

Nesse contexto, alguns estudos (Onteru *et al.*, 2013; Sahana *et al.*, 2013; Do *et al.*, 2014a e 2014b; Jiao *et al.*, 2014, Howard *et al.*, 2015; Qiao *et al.*, 2015; Wang *et al.*, 2015; Sato *et al.*, 2016; Guo *et al.*, 2017) têm utilizado painéis densos de polimorfismos de base única (SNPs) para realizar estudos de associação genômica ampla (**GWAS**), a fim de identificar SNPs associados a características ligadas à EA que possam ser incorporados à seleção. Porém, devido à dificuldade e os altos custos para medir esse fenótipo, GWAS para características relacionadas à EA ainda são limitados. Portanto, novos GWAS para essas características são necessários para identificação de SNPs que possam, além de auxiliar no processo de seleção, promover maior entendimento dos mecanismos genéticos envolvidos no controle dessas características.

É importante salientar que os GWAS realizados previamente têm focado principalmente em fenótipos coletados em populações de animais de raças puras criadas em ambientes de alto nível sanitário, enquanto o produto final do melhoramento de suínos são animais cruzados criados em ambiente de menor nível sanitário (Hidalgo *et al.*, 2015), utilizando geralmente dietas com níveis nutricionais distintos, além de, na grande maioria das vezes, serem submetidos a condições de manejo bastante diferenciadas. De acordo com Veroneze *et al.* (2014), o uso de populações cruzadas como populações de referência é de grande interesse, uma vez que os animais normalmente estão alojados em

ambientes similares aos que os animais comerciais estarão submetidos, e, em virtude de que uma população cruzada pode ser utilizada como referência para seleção em todas as populações puras que lhe deram origem. Portanto, SNPs relacionados à EA identificados em populações cruzadas poderiam ser utilizados em estratégias de seleção assistida por marcadores para selecionar animais de linhas puras, visando melhorar o desempenho de sua progênie cruzada. Todavia, GWAS utilizando dados de populações cruzadas ainda são escassos, devido à baixa disponibilidade de dados fenotípicos relacionados à EA de populações cruzadas também genotipadas.

Nesse estudo, objetivou-se identificar SNPs com associação significativa com características relacionadas à EA: 1) ganho de peso médio diário (ADG), 2) consumo de ração médio diário (ADFI), e 3) conversão alimentar (FCR); coletadas em uma população de suínos cruzados. Além disso, objetivou-se identificar potenciais genes candidatos para estas características.

## **Material e métodos**

### *Aspectos éticos*

A condição de alojamento dos animais, a coleta de informações e a coleta de amostras para extração do DNA cumpriram os protocolos e seguiram as rotinas regularmente utilizadas pela granja experimental, sendo praticadas em conformidade com as normas de proteção e bem estar animal dos Países Baixos.

### *Origem das informações*

As informações fenotípicas utilizadas neste estudo fazem parte do banco de dados de um programa comercial de melhoramento genético de suínos. Os dados foram coletados na granja experimental Schothorst Feed Research (Lelystad, Países Baixos), de acordo com protocolos de coleta de informações regularmente utilizados nessa unidade. O nível sanitário dessa granja se assemelha ao nível sanitário das demais granjas comerciais dos Países Baixos.

### *Animais e coleta de informações fenotípicas*

O banco de dados fenotípicos utilizado continha inicialmente informações de um total de 4.981 suínos terminados (2.553 machos castrados e 2.428 fêmeas),

originados do cruzamento entre machos de uma linhagem sintética e fêmeas híbridas F1 (Landrace x Large White). As informações foram coletadas entre agosto de 2012 e novembro de 2016.

Os animais foram pesados no início e no final do período de teste nas fases de crescimento/terminação e apresentaram pesos médios de  $23,3 \pm 4,8$  kg e  $123,6 \pm 8,4$  kg, respectivamente. As médias de idade dos animais no início e no final do período de teste foram de  $62,8 \pm 3,2$  e  $168,3 \pm 9,8$  dias, respectivamente. As informações de consumo individual de ração foram registradas durante todo o período de teste utilizando estações de alimentação computadorizadas IVOG (<http://insentec.eu/en/pig-management/ivog>), com obtenção do consumo individual a cada visita do animal à estação de alimentação.

O ADG de cada animal foi calculado pela diferença entre o peso final e peso inicial, dividido pelo número de dias do período de teste. O ADFI de cada animal durante o período de teste foi obtido pela soma da ração total consumida durante o teste, dividido pelo número de dias do período de teste. A FCR de cada animal foi obtida pela divisão de seus valores de ADFI pelo ADG.

Durante os procedimentos de controle de qualidade, foram excluídos dados de animais que não apresentassem informações para as características em estudo, sendo removidas informações de 2.242 animais por não apresentarem informação de ADFI e FCR. Também foram removidos dados de animais que apresentaram valores superiores ou inferiores a  $3,5 * DP$  (Desvio Padrão) para as características. Além disso, foram removidos dados de animais que compuseram grupos contemporâneos de rebanho-ano-semana (**HYW**) formados por menos de 9 animais. Assim, informações de 4.964 animais foram utilizadas para a pré-correção dos fenótipos.

Para a realização do GWAS, após a pré-correção dos fenótipos, foram excluídas as informações de animais que não foram genotipados. Ao final, resultaram informações dos totais de animais apresentados na Tabela 1.

#### *Pré-correção dos fenótipos*

Para a estimativa dos efeitos não genéticos sobre cada característica estudada e posterior pré-correção dos fenótipos, foi utilizado o software ASReml v. 3.0

(Gilmour *et al.*, 2009), sendo incluídas informações de parentesco de no mínimo quatro gerações e utilizando o seguinte modelo:

$$y_{ijklmno} = \mu + sex_i + HYWSt_j + WSt_k + T_l + L_m + P_n + animal_o + e_{ijklmno}, \quad (1)$$

em que  $y_{ijklmno}$  é a medida fenotípica do animal  $o$ ;  $\mu$  é o valor médio da característica na população;  $sex_i$  é o efeito fixo do sexo  $i$ ;  $HYWSt_j$  é o efeito fixo de grupo de contemporâneos  $j$  (animais do mesmo rebanho ( $H$ ), ano ( $Y$ ) e semana ( $W$ )  $j$  em que foi realizada a pesagem no início do teste);  $WSt_k$  é o efeito fixo do peso  $k$  no início do período de teste;  $T_l$  é o efeito fixo do experimento  $l$  a que foi submetido o animal;  $L_m$  é o efeito da leitegada de nascimento  $m$ , assumido como aleatório;  $P_n$  é o efeito aleatório do grupo de animais que ocuparam a mesma baia  $n$ ;  $animal_o$  é o efeito aleatório genético aditivo do próprio animal, assumido ser distribuído como  $\sim N(0, \mathbf{A}\sigma_a^2)$ , sendo  $\mathbf{A}$  a matriz dos numeradores dos coeficientes de parentesco e  $\sigma_a^2$  a variância genética aditiva; e  $e_{ijklmno}$  é o efeito residual aleatório, assumido ser distribuído como  $\sim N(0, \mathbf{I}\sigma_e^2)$ , sendo  $\mathbf{I}$  a matriz identidade e  $\sigma_e^2$  a variância residual.

Os fenótipos foram, então, corrigidos para os efeitos não genéticos estimados pelo modelo (1) para serem posteriormente utilizados na GWAS. Os fenótipos pré-corrigidos foram utilizados para estimar os componentes de variâncias e herdabilidades utilizando o software ASReml v. 3.0 (Gilmour *et al.*, 2009) (Tabela 2).

#### *Genotipagem e controle de qualidade*

Um total de 2.386 animais da população de suínos cruzados avaliados neste estudo foram genotipados utilizando o *chip* GeneSeek Custom 10K SNP (*chip* 10K) (GeneSeek Inc., Lincoln, NE). Porém, para maior resolução do GWAS, os animais cruzados genotipados com o *chip* 10K tiveram seus genótipos imputados para o *chip* GeneSeek Custom 80K SNP (*chip* 80K) (GeneSeek Inc., Lincoln, NE). Essa imputação pode ser realizada com alta acurácia porque ambos os pais dos animais cruzados foram também genotipados utilizando o *chip* 80K ou o Illumina Porcine SNP60 BeadChip (*chip* 60K) (Illumina, Inc., San Diego, CA). Além dos pais dos animais cruzados (174 reprodutores e 313 matrizes F1), outros animais genotipados de cada uma das linhas paternas foram



utilizados para aumentar a acurácia do processo de imputação (totalizando 8.110 animais da população sintética paterna e 3.012 animais da população F1).

Animais com falhas de genotipagem  $>0,05$  dentro do *chip* para os quais foram genotipados (10K, 60K ou 80K), seriam excluídos do banco de dados. Porém, isso não foi necessário porque todos os animais de todas as populações apresentaram falhas de genotipagem  $<0,05$ . O controle de qualidade dos SNPs, realizado logo em seguida, consistiu na exclusão de SNPs localizados em cromossomos sexuais, com posição desconhecida no Pig Genome build10.2 (Groenen *et al.*, 2012), com falhas de genotipagem  $>0,05$ , ou com forte desvio no equilíbrio de Hardy-Weinberg, apresentando  $\chi^2 >600$  (Hayes *et al.*, 2009; Resende *et al.*, 2012; Lopes *et al.*, 2013; Do *et al.*, 2015).

Antes de realizar o processo de imputação dos genótipos dos animais cruzados, a imputação dentro de cada população parental foi realizada separadamente utilizando o software AlphaImpute (Hickey *et al.*, 2011). Nesse processo, os animais genotipados tiveram seus genótipos imputados para o painel de SNPs do *chip* 80K.

Em um segundo passo, a população sintética paterna e a população F1 foram combinadas como uma única população de referência para a imputação dos genótipos dos animais cruzados utilizando todo o painel de SNPs que passaram pelo controle de qualidade em ambas as populações parentais. O software AlphaImpute (Hickey *et al.*, 2011) também foi utilizado nesse passo. Após a imputação, foram mantidos somente os SNPs que estavam segregando na população cruzada (menor frequência alélica  $>0,01$ ), totalizando 51.468 SNPs. Nas populações parentais, o DNA foi extraído a partir de bulbos capilares ou de fragmentos de tecidos do pavilhão auricular (coletados por cânula localizada na agulha do aplicador do brinco de identificação aplicado nos leitões logo após o nascimento). Na população cruzada, foi extraído a partir de fragmentos de tecidos do pavilhão auricular.

### *Análises de associação*

Para realização da análise de associação genômica (*single*-SNP GWAS) foi também utilizado o software ASReml v. 3.0 (Gilmour *et al.*, 2009), empregando o seguinte modelo:

$$y_c = \mu + SNP_i + animal_j + e_{ij}, \quad (2)$$

em que  $y_c$  é o fenótipo do animal corrigido para todos os efeitos não genéticos estimados pelo modelo 1;  $\mu$  é o valor médio do fenótipo corrigido avaliado na população;  $SNP_i$  é o efeito fixo do SNP  $i$ ;  $animal_j$  é o efeito aleatório genético aditivo do animal  $j$ , assumindo ser distribuído como  $\sim N(0, \mathbf{A}\sigma_a^2)$ , sendo  $\mathbf{A}$  a matriz da média dos numeradores dos coeficientes de parentesco,  $\sigma_a^2$  a variância genética aditiva; e  $e_{ij}$  é o efeito residual aleatório, assumido ser distribuído como  $\sim N(0, \mathbf{I}\sigma_e^2)$ , sendo  $\mathbf{I}$  a matriz identidade e  $\sigma_e^2$  a variância residual.

Os genótipos dos SNPs foram codificados como 0, 1 e 2, sendo 0 para genótipos AA, 1 para AB e 2 para BB.

Os P-valores observados e esperados de cada característica foram plotados em gráficos quantil-quantil com auxílio da função `qqnorm` do Software R (*version 3.0.2*) para análise visual da normalidade e identificação de possível inflação dos P-valores. Utilizando a função `estlambda` do pacote GenABEL do Software R, foi estimado o fator de inflação *lambda* para as distribuições dos P-valores. O controle genômico, dividindo-se os valores F pelo valor de *lambda*, conforme procedimento utilizado por Devlin e Roeder (1999), Bacanu *et al.* (2000), Hinrichs *et al.* (2009) e Rashidi (2016) foi aplicado para todas as características em que os valores de *lambda* foram maiores que 1,0 (Hinrichs *et al.*, 2009).

Após o controle genômico, foi estimada a FDR (*False Discovery Rate*), com o uso do pacote `qvalue` do Software R, utilizando os P-valores corrigidos para evitar resultados falsos positivos devido à múltipla testagem. Foram consideradas significativas as associações entre SNPs e as características avaliadas com  $FDR \leq 0,005$ .

#### *Regiões de QTL e genes candidatos*

As regiões QTL (*locus* de características quantitativas) foram definidas como sendo aquelas formadas pelo grupo de SNPs (localizados em um mesmo cromossomo - SSC) com significativa associação às características estudadas, cujas janelas de 1Mb (anteriores ou posteriores à posição de cada SNP identificado) apresentaram alguma sobreposição.

A variância de cada característica explicada pelas regiões de QTL identificadas foi definida como a variância genética aditiva do SNP mais significativo ( $\sigma_{SNP}^2$ ) na

respectiva região de QTL. Com base nesse valor, foi estimado o percentual da variância fenotípica explicada por esse SNP ( $\% \sigma_{SNP}^2$ ). Ambos valores foram estimados como:

$$\sigma_{SNP}^2 = 2pq\alpha^2$$

$$\% \sigma_{SNP}^2 = \frac{\sigma_{SNP}^2}{\sigma_p^2} \times 100 ,$$

em que,  $p$  e  $q$  são as frequências alélicas observadas,  $\alpha$  é o efeito da substituição alélica do SNP mais significativo em cada região QTL e  $\sigma_p^2$  é a variância fenotípica da característica.

A identificação da posição dos genes em relação às janelas de 1Mb anterior e posterior às regiões QTL identificadas neste estudo, foi realizada com o auxílio do Genome Gene Viewer (Sequence Viewer 3.20.4 - [https://www.ncbi.nlm.nih.gov/genome/gdv/browser/?context=genome&acc=GC\\_F\\_000003025.5](https://www.ncbi.nlm.nih.gov/genome/gdv/browser/?context=genome&acc=GC_F_000003025.5), acessado em 24 de abril de 2017). Foram definidos como candidatos aqueles genes que, em estudos anteriores, foram identificados e relacionados com processos metabólicos ou mecanismos moleculares que possam estar envolvidos com as características de ADG, ADFI ou FCR.

Buscando identificar estudos prévios de associações entre as mesmas regiões QTL identificadas no presente estudo e as mesmas características estudadas, foi realizada pesquisa utilizando o Animal QTLdb, baseado no *Sus scrofa* build 10.2 (<https://www.animalgenome.org/cgi-bin/QTLdb/SS/browse>, acessado em 24 de abril de 2017).

## Resultados

### *Parâmetros genéticos e componentes de variância*

Os componentes de variância foram estimados após a pré-correção dos fenótipos (Tabela 2). Para as características de ADG e FCR as variâncias fenotípicas foram de, respectivamente, 5.813,95 (g/dia)<sup>2</sup> e 0,019 (g/g)<sup>2</sup>, com herdabilidades de 0,19 e 0,22, respectivamente. Para a característica de ADFI a variância fenotípica foi de 34.398,9 (g/dia)<sup>2</sup>, com herdabilidade de 0,29, portanto, maior quando comparada às demais características alvos deste estudo.

### *Controle do GWAS*

Durante os procedimentos de controle genômico, os valores *lambda* para o efeito SNP em todas as características apresentaram valores acima do limite de 1,0. (Hinrichs *et al.*, 2009), motivando a necessidade de ajustá-los para a inflação dos P-valores. Após o controle genômico, os valores de *lambda* passaram de 1,456 para 1,001 para a característica ADG, de 1,615 para 1,001 para ADFI e de 1,359 para 1,001 para FCR (Figura 1).

### *SNPs associados ao ADG*

A análise da associação dos SNPs com a característica ADG considerou significativos aqueles com limiar para FDR  $\leq 0,005$  (Figura 2), sendo identificados 8 SNPs no SSC1 localizados entre 177,01 e 185,47 Mb e 9 no SSC12, localizados entre 2,87 e 3,22 Mb (Tabela 3). No SSC1, ALGA0006599 (localizado em 177,01 Mb) foi o SNP mais significativo, com  $-\log_{10}$  (P-valor) de 7,09, com efeito de 18,87 g de ADG, explicando 3,06% da variância fenotípica para essa característica. No SSC12, WU\_10.2\_12\_2922849 (localizado em 2,92 Mb) foi o SNP mais significativo, apresentando  $-\log_{10}$  (P-valor) de 6,69, com efeito de 16,97 g de ADG, explicando 2,29% da variância fenotípica para essa característica.

### *SNPs associados ao ADFI*

Para a característica de ADFI, foram identificados 8 SNPs no SSC1 com significativo efeito sobre a característica (FDR  $< 0,005$ ), localizados entre 173,26 e 185,47 Mb. WU\_10.2\_1\_180284104 foi o SNP mais significativo, com  $-\log_{10}$  (P-valor) de 8,64, apresentando efeito de substituição alélica de -53,66 g/dia, explicando 3,26% da variância fenotípica para a característica.

### *SNPs com efeitos significativos para ADG e ADFI*

No SSC1 foram identificados 5 SNPs localizados na região QTL entre 177,01 e 185,47 Mb, com efeito significativo (FDR  $\leq 0,005$ ) sobre as características ADG e ADFI. Apesar disso, não foram encontrados marcadores com efeito significativo para a característica FCR, mesmo quando utilizado um limiar de FDR menos crítico (FDR  $\leq 0,20$ ).

## Discussão

Neste estudo, foram identificadas duas regiões de QTL localizadas nos SSC1 e SSC12 associadas à característica ADG e uma região localizada no SSC1 associada a ADFI, entretanto não foram encontradas associações com FCR. No SSC1, foram identificados SNPs com associações significativas para as duas características, ADG e ADFI.

### *Controle genômico*

No GWAS, em geral, são utilizados grandes números de marcadores e pequenas diferenças nas frequências alélicas entre subgrupos fenotípicos podem obter significância estatística. Resultados falsos positivos são observados quando a diferença entre a frequência alélica entre subgrupos fenotípicos ocorre em razão de erros na medição do fenótipo, ou quando não estão relacionadas às características estudadas (Hinrichs *et al.*, 2009).

Em GWAS, assume-se que apenas um pequeno número de marcadores estarão associados com a característica estudada e que o teste estatístico para os outros marcadores deveria seguir a distribuição sob a hipótese nula de não associação entre o marcador e a característica. Entretanto, diferenças artificiais na frequência alélica devido à estratificação da população e erros de genotipagem, afetariam todos os SNPs. Assim sendo, o teste estatístico seria inflado em todo o genoma (Yang *et al.*, 2011).

A estratificação populacional pode levar à identificação de falsos positivos pelas diferenças na frequência alélica dentro das famílias. Price *et al.* (2010) descreveram que uma forma de se evitar falsos positivos devido à estratificação populacional, é a aplicação de modelos mistos e a utilização da matriz de parentesco. Dessa forma, são consideradas a variância e a covariância entre os indivíduos avaliados.

Para avaliar se as estratégias adotadas durante a realização da GWAS foram suficientes para evitar falsos positivos, recomenda-se a aplicação do controle genômico para estimar o fator de inflação (*lambda*) dos P-valores (Hinrichs *et al.*, 2009; Price *et al.*, 2010). Quando o *lambda* é menor ou igual a 1, nenhum ajuste é necessário (Hinrichs *et al.*, 2009; Price *et al.*, 2010), entretanto quando maior que 1, é sugerida a multiplicação do valor estatístico pelo fator  $1/\lambda$  (Devlin e

Roeder, 1999; Clayton *et al.*, 2005). Esse procedimento reduz a dispersão dos dados, sendo desnecessário remover grande número de informações. Embora com o uso do ajuste da inflação exista uma perda no poder dos SNPs, muitas vezes ela é preferível à sua exclusão.

No presente estudo, um rigoroso controle dos dados genômicos foi aplicado para se excluir eventuais erros de genotipagem. Modelos mistos associados à uma matriz de numeradores dos coeficientes de parentesco também foram aplicados para corrigir eventual estratificação populacional. Entretanto, mesmo com todos esses passos, os valores de *lambda* para todas as características ficaram acima de 1, indicando inflação dos P-valores. Por essa razão, os P-valores foram ajustados de acordo com os valores de *lambda* para maior confiabilidade dos resultados descritos neste trabalho (Figura 1).

#### *Regiões QTL associadas a ADG e ADFI e genes candidatos*

Neste estudo, identificamos duas regiões QTL apresentando significativa associação com a característica ADG, sendo uma com 8 SNPs localizada entre 177,01 e 185,47 Mb no SSC1, e outra com 9 SNPs, localizada entre 2,87 e 3,22 Mb no SSC12. Para a característica ADFI, foi identificada associação significativa com a região QTL localizada entre 173,26 e 185,47 Mb no SSC1.

Onteru *et al.* (2013) identificaram associação fortemente significativa entre a característica ADG e a região QTL localizada no SSC1 em 177,0 Mb (janela de 1Mb – SNPs ALGA0006599 e INRA0004954) que explicou 2,4% da variância fenotípica para essa característica. Neste trabalho, identificamos 3 SNPs localizados nessa mesma região QTL com associação significativa à característica ADG, sendo ALGA0006599 o SNP com associação mais significativa, explicando 3,06% da variância fenotípica para essa característica.

Alguns genes identificados anteriormente foram localizados nas mesmas janelas de 1Mb do SSC1 em que se identificou SNPs associados ao ADFI ou ADG (Tabela 4).

Os genes *CDH19* e *CDH7* (Caderina 19 e Caderina 7), genes posicionados respectivamente entre 172,92 e 173,01 Mb e entre 173,58 e 173,67 Mb, têm sido muito estudados em humanos. Os membros da família das caderinas compõem um grupo de moléculas cálcio dependentes que permitem a ligação e promovem

a adesão entre células vizinhas (Kools *et al.*, 2000), desempenhando importante papel na formação e manutenção da integridade de diversos tecidos. Em frangos de corte, CDH7 está envolvido no desenvolvimento da medula espinhal em estágios embrionários iniciais (Lin *et al.*, 2014).

Os genes *RNF152* (entre 177,06 e 177,07 Mb) e *PMAIP1* (entre 178,85 e 178,86 Mb), localizados no SSC1, foram anteriormente correlacionados com o aumento de gordura intramuscular em suínos (Rothschild *et al.*, 2014) e localizam-se dentro da janela de 1Mb anterior de ALGA0006599 e posterior a WU\_10.2\_1\_180358223 (Tabela 4).

Diversos trabalhos têm associado os genótipos *MC4R* (gene receptor de melanocortina-4) localizado no SSC1 com o ganho de peso diário, consumo de ração e gordura (Kim *et al.*, 2000; Davoli *et al.*, 2012). O gene *MC4R* possui papel importante no controle do consumo de alimento e homeostase energética (Chao *et al.*, 2012). Em um estudo realizado com linhagens Large White durante sete gerações com diferentes objetivos de seleção (conversão alimentar magra, crescimento magro com alimentação *ad libitum* e crescimento magro com alimentação restrita e consumo diário de ração), a associação do polimorfismo Asp298Asn (substituição do ácido aspártico por asparagina na posição 298 do gene *MC4R* do suíno) mostrou diferentes efeitos sobre animais de alto e baixo mérito genético nas características avaliadas, porém, quando analisado sobre todas as linhagens, o polimorfismo Asp298Asn foi significativamente associado ( $P < 0,05$ ) à espessura de toucinho medida por ultrassom, ADG e ADFI (Houston *et al.*, 2004).

Baseado em estudos realizados com várias raças e linhagens de suínos, a significância do efeito do genótipo do *locus MC4R* sobre crescimento e gordura pode variar, dependendo da raça ou linhagem (Chao *et al.*, 2012).

Neste trabalho, foi identificada forte associação entre o SNP ALGA0006623 (posicionado a 178,02 Mb no SSC1) com a característica ADG. Este SNP localiza-se a aproximadamente 0,5 Mb do gene *MC4R* (posicionado entre 178,55 – 178,56 Mb no SSC1). Além disso, toda a região QTL compreendida pela janela de 1 Mb dos SNPs identificados entre 177,01 e 180,36 Mb apresentam sobreposição com a janela de 1Mb do gene *MC4R*. Essa mesma região QTL apresentou forte associação com a característica ADFI, corroborando os

achados prévios que indicavam a associação desse gene com essa mesma característica.

Devido à propriedade intelectual (patente) relacionada ao *MC4R*, os *chips* comerciais utilizados neste estudo não possuem a mutação causal e nem mesmo SNPs fortemente ligados ao gene *MC4R* (Fan *et al.*, 2011; Onteru *et al.*, 2013). Esse pode ter sido o motivo pelo qual não tenhamos identificado SNPs com associação significativa com ADG e ADFI localizados em posição ainda mais próxima ao gene *MC4R* no SSC1. Caso houvesse marcadores mais próximos ao *MC4R* disponíveis nos *chips* comerciais, provavelmente seriam identificadas associações ainda mais significativas com esse gene e efeitos de substituição alélica ainda maiores.

O SNP H3GA0003232 (184,27 Mb no SSC1), que demonstrou significativa associação com a característica ADFI, localiza-se a menos de 0,5 Mb da posição do gene *FEM1B* (entre 183,92 e 183,93 Mb). Em mamíferos os genes *FEM1* são expressos em tecidos de relevância à fisiologia da glicose, incluindo pâncreas e músculo esquelético. Em estudo realizado com ratos, houve evidências de que *FEM1B* pode estar envolvido na homeostase da glicose (Lu *et al.*, 2005).

O gene *FEM1B* localiza-se na região compreendida pela janela de 1Mb dos SNPs localizados entre 184,27 e 185,47 no SSC1.

Como descrito anteriormente, neste estudo também encontrou-se significativa associação entre a região QTL localizada entre 2,87 e 3,22 Mb no SSC12 com a característica de ADG, sendo WU\_10.2\_12\_2922849 o SNP mais significativo, explicando 2,29% da variância fenotípica da característica.

Um estudo prévio realizado por Rohrer (2000) com suínos retrocruzados Meishan-White, associou o SSC12 com o ADG entre 8 e 18 semanas de idade, porém, desconhecemos estudos de GWAS que tenham associado esta região QTL com a característica de ADG.

O gene *GAA* (ácido alfa-glicosidase), localizado entre 2,30 e 2,41 Mb no SSC12, foi um dos genes que ficou dentro da janela de 1Mb da região QTL associada com a característica ADG. Em humanos, esse gene codifica a enzima ácido alfa-glicosidase e sua função essencial é a degradação do glicogênio em glicose nos lisossomos (Reuser *et al.*, 1985).



### *Associação com FCR*

Embora tenha sido observada a associação genômica para as características diretamente envolvidas no cálculo da FCR, algumas vezes envolvendo os mesmos marcadores (o que aconteceu com 5 marcadores), não se encontrou associação significativa e específica dessa característica com alguma região genômica, mesmo com a tentativa de utilização de um limiar menos crítico ( $FDR \leq 0,20$ ). Uma das hipóteses que pode explicar essa situação é o fato que os mesmos SNPs que apresentaram associação significativa com ADG e ADFI, mostraram efeitos significativos na mesma direção para essas características (exemplos: ALGA0006599 com efeito de substituição alélica de 18,87 g/dia para ADG e 44,89 g/dia para ADFI; MARC0013872 com efeito de -17,34 g/dia para ADG e -54,18 para ADFI) (Tabela 4).

Os resultados obtidos neste trabalho permitiram observar que as características de ADG e ADFI são controladas por várias regiões QTL com pequeno efeito sobre essas características. Essa observação também pode explicar o fato de não havermos encontrado associação significativa de nenhum SNP com FCR, considerando  $FDR \leq 0,005$  ou mesmo em limiares menos críticos.

O crescimento é uma característica complexa que é definida através de vários processos biológicos que acontecem no animal. Entender os aspectos genéticos que envolvem os componentes associados ao crescimento corporal (conteúdo muscular, gordura, ossos, vísceras, entre outros) e sua dinâmica de crescimento ao longo da vida são fundamentais. Além disso, mais de um terço da variação no consumo de ração e eficiência alimentar estão relacionadas a processos que independem de crescimento e composição corporal (Knap e Wang, 2012).

Características complexas como crescimento e consumo são geralmente resultado de sistemas dinâmicos (Howard *et al.*, 2015). Ainda de acordo com Knap e Wang (2012), a utilização das características que compõem a FCR dentro dos objetivos de seleção (como exemplo: ADG, ADFI e espessura de toucinho), permite maior controle sobre cada elemento relevante do sistema, separando a FCR em suas características componentes funcionais para que possam ser manipuladas de maneira independente.

Dentro desse aspecto, a mensuração precisa das características que compõem a EA têm demonstrado poder auxiliar no entendimento dos mecanismos

envolvidos com as características primárias que compõem a FCR. Com esse objetivo, a utilização de métodos não invasivos, como a Tomografia Computadorizada (CT), Ressonância Magnética (MR) e a Absorciometria radiográfica de dupla energia, têm sido utilizados como ferramentas para investigar aspectos genéticos ligados ao crescimento (Kušec *et al.*, 2016). Utilizando-se dessas ferramentas é possível estimar, pelo uso de imagens e com o auxílio de ferramentas estatística adequadas, o conteúdo corporal dos animais *in vivo* de maneira acurada. Kongsro (2014) relatou estimativa de ganho de 30% no progresso genético do percentual de carne magra em suínos pelo uso de técnica não invasiva para coleta de informações em animais vivos.

Grande número de estudos sobre seleção genômica tem demonstrado e apresentado diversas possibilidades ao uso dessa nova ferramenta (Jonas e de Koning, 2015), porém ainda há grande número de possibilidades a serem exploradas. Neste estudo, identificou-se regiões QTL fortemente associadas às características de ADG e ADFI. Esse fato, associado à relevância dessas características para os programas de melhoramento genético em suínos, nos fazem sugerir a realização de novos estudos que possam avaliar o potencial de usar os marcadores significativos aqui identificados para ADG e ADFI em modelos de predição que aliem seleção assistida por marcadores (MAS) com seleção genômica ampla (GBLUP), conforme descrito por Lopes *et al.* (2017).

### **Conclusões**

Neste estudo, identificou-se regiões de QTL com forte associação com características de ADG e ADFI no SSC1 e com a característica ADG no SSC12 em uma população de suínos cruzados. Essas regiões foram previamente descritas em populações de linhagens puras de suínos, porém, este é o primeiro estudo a confirmar a relevância dessas regiões de QTL em uma população de suínos cruzados.

A não identificação de associação genômica com a característica FCR pode justificar-se pela complexidade dessa característica. Este resultado evidencia que GWAS para fenótipos mais refinados (como ADG e ADFI) tem maior poder de identificação de QTL do que GWAS para características que correspondem a uma função de duas outras características (como FCR).

Neste estudo, também identificamos alguns genes candidatos, que poderão ser avaliados quanto ao seu potencial de utilização como marcadores para as características de ADG e ADFI, em modelos de predição aliando a seleção genômica ampla com a seleção assistida por marcadores.

### **Agradecimentos**

Agradecemos ao Topigs Norsvin Research Center (Beuningen, the Netherlands), por disponibilizar a base de dados fenotípicos e informação genômica, bem como pelo apoio na análise dos dados.

### **Referências**

- Bacanu S-A, Devlin B and Roeder K 2000. The power of genomic control. *The American Journal of Human Genetics* 66, 1933–1944.
- Chao Z, Wang F, Deng C–Y, Wei L–M, Sun R–P, Liu H–L, Liu Q–W and Zheng X–L 2012. Distribution and linkage disequilibrium analysis of polymorphisms of *MC4R*, *LEP*, *H-FABP* genes in the different populations of pigs, associated with economic traits in DIV2 line. *Molecular Biology and Reproduction* 39, 6329–6335.
- Clayton DG, Walker NM, Smyth DJ, Pask R, Cooper JD, Maier LM, Smink LJ, Lam AC, Ovington NR, Stevens HE, Nutland S, Howson JMM, Faham M, Moorhead M, Jones HB, Falkowski M, Hardenbol P, Willis TD and Todd JA 2005. Population structure, differential bias and genomic control in a large-scale, case-control association study. *Nature Genetics* 37, 1243-1246.
- Davoli R, Braglia S, Valastro V, Annarratone C, Comella M, Zambonelli P, Nisi I, Gallo M, Buttazzoni L and Russo V 2012. Analysis of *MC4R* polymorphism in Italian Large White and Italian Duroc pigs: Association with carcass traits. *Meat Science* 90, 887–892.
- Devlin B and Roeder K 1999. Genomic control for association studies. *Biometrics* 55, 997-1004.
- Do DN, Strathe AB, Ostersen T, Pant SD and Kadarmideen HN 2014a. Genome-wide association and pathway analysis of feed efficiency in pigs reveal candidate genes and pathways for residual feed intake. *Frontiers in Genetics* 5, 307. doi: 10.3389/fgene.2014.00307.
- Do DN, Ostersen T, Strathe AB, Mark T, Jensen J and Kadarmideen HN 2014b. Genome-wide association and systems genetic analyses of residual feed intake,

- daily feed consumption, backfat and weight gain in pigs. *BMC Genetics* 15, 27. doi: 10.1186/1471-2156-15-27.
- Do DN, Janss LLG, Jensen J and Kadarmideen HN 2015. SNP annotation-based whole genomic prediction and selection: An application to feed efficiency and its component traits in pigs. *Journal of Animal Science* 93, 2056–2063.
- Fan B, Onteru SK, Du Z-Q, Garrick DJ, Stalder KJ and Rothschild MF 2011. Genome-wide association study identifies loci for body composition and structural soundness traits in pigs. *Plos One* 6, e14726.
- Gilmour AR, Gogel B, Cullis B and Thompson R 2009. ASReml user guide release 3.0. VSN International Ltd, Hemel Hempstead, UK.
- Groenen MAM, Archibald AL, Uenishi H, Tuggle CK, Takeuchi Y, Rothschild MF, Rogel-Gaillard C, Park C, Milan D, Megens H-J, Li S, Larkin DM, Kim H, Frantz LAF, Caccamo M, Ahn H, Aken BL, Anselmo A, Anthon C, Auvil L, Badaoui B, Beattie CW, Bendixen C, Berman D, Blecha F, Blomberg J, Bolund L, Bosse M, Botti S, Bujie Z, Bystrom M, Capitanu B, Carvalho-Silva D, Chardon P, Chen C, Cheng R, Choi S-H, Chow W, Clark RC, Clee C, Crooijmans RPMA, Dawson HD, Dehais P, De Sapio F, Dibbits B, Drou N, Du Z-Q, Eversole K, Fadista J, Fairley S, Faraut T, Faulkner GJ, Fowler KE, Fredholm M, Fritz E, Gilbert JGR, Giuffra E, Gorodkin J, Griffin DK, Harrow JL, Hayward A, Howe K, Hu Z-L, Humphray SJ, Hunt T, Hornshøj H, Jeon J-T, Jern P, Jones M, Jurka J, Kanamori H, Kapetanovic R, Kim J, Kim J-H, Kim KW, Kim T-H, Larson G, Lee K, Lee K-T, Leggett R, Lewin HA, Li Y, Liu W, Loveland JE, Lu Y, Lunney JK, Ma J, Madsen O, Mann K, Matthews L, McLaren S, Morozumi T, Murtaugh MP, Narayan J, Nguyen DT, Ni P, Oh S-J, Onteru S, Panitz F, Park E-W, Park H-S, Pascal G, Paudel Y, Perez-Enciso M, Ramirez-Gonzalez R, Reecy JM, Rodriguez-Zas S, Rohrer GA, Rund L, Sang Y, Schachtschneider K, Schraiber JG, Schwartz J, Scobie L, Scott C, Searle S, Servin B, Southey BR, Sperber G, Stadler P, Sweedler JV, Tafer H, Thomsen B, Wali R, Wang J, Wang J, White S, Xu X, Yerle M, Zhang G, Zhang J, Zhang J, Zhao S, Rogers J, Churcher C and Schook LB 2012. Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* 491, 393-398.
- Guo Y, Huang Y, Hou L, Ma J, Chen C, Ai H, Huang L and Ren J 2017. Genome-wide detection of genetic markers associated with growth and fatness in four pig populations using four approaches. *Genetics Selection Evolution* 49, 21.
- Hayes BJ, Bowman PJ, Chamberlain AJ and Goddard ME 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. *Journal of Dairy Science* 92, 433–443.

- Hayes BJ, Harris AL and Goddard ME 2013. The future of livestock breeding: genomic selection for efficiency, reduced emissions intensity, and adaptation. *Trends in Genetics* 29, 206-214.
- Hickey JM, Kinghorn BP, Tier B, Wilson JF, Dunstan N, van der Werf JHJ 2011. A combined long-range phasing and long haplotype imputation method to impute phase for SNP genotypes. *Genetics Selection Evolution* 43, 12. doi: 10.1186/1297-9686-43-12.
- Hidalgo AM, Bastiaansen JW, Lopes MS, Harlizius B, Groenen MA and de Koning DJ 2015. Accuracy of predicted genomic breeding values in purebred and crossbred pigs. *G3: Genes| Genomes| Genetics* 5, 1575-1583.
- Hinrichs AL, Larkin EK and Suarez BK 2009. Population stratification and patterns of linkage disequilibrium. *Genetic Epidemiology* 33, S88–S92. doi: 10.1002/gepi.20478.
- Houston RD, Cameron ND and Rance KA 2004. A melanocortin-4 receptor (*MC4R*) polymorphism is associated with performance traits in divergently selected large white pig populations. *Animal Genetics* 35, 386–390.
- Howard JT, Jiao S, Tiezzi F, Huang Y, Gray KA and Maltecca C 2015. Genome-wide association study on legendre random regression coefficients for the growth and feed intake trajectory on Duroc boars. *BMC Genetics* 16, 59. doi: 10.1186/s12863-015-0218-8.
- Jiao S, Maltecca C, Gray KA and Cassady JP 2014. Feed intake, average daily gain, feed efficiency, and real-time ultrasound traits in Duroc pigs: II. Genomewide association. *Journal of Animal Science* 92, 2846-2860.
- Jonas E and de Koning D-J 2015. Genomic selection needs to be carefully assessed to meet specific requirements in livestock breeding programs. *Frontiers in Genetics* 6, 49. doi: 10.3389/fgene.2015.00049.
- Kim KS, Larsen N, Short T, Plastow G and Rothschild MF 2000. A missense variant of the porcine melanocortin-4 receptor (*MC4R*) gene is associated with fatness, growth, and feed intake traits. *Mammalian Genome* 11, 131–135.
- Knap PW and Wang L 2012. Pig breeding for improved feed efficiency. In *Feed efficiency in swine* (ed. J. F. Patience), pp. 167-179. Wageningen Academic Publishers, Wageningen, the Netherlands.
- Kongsro J 2014. Genetic gain on body composition in pigs by computed tomography (CT): return on investment. In *FAIM III - Farm Animal Imaging Copenhagen 2014* (ed. Maltin C, Craigie C and Bünger L), pp. 28-30. Scotland's Rural College, Edinburgh, Scotland, UK.

- Kools P, Imschoot G, Roy F 2000. Characterization of three novel human Cadherin genes (*CDH7*, *CDH19*, and *CDH20*) clustered on chromosome 18q22–q23 and with high homology to chicken Cadherin-7. *Genomics* 68, 283-295.
- Kušec G, Sholz AM, Baulain U, Djurkin Kušec I and Bernau M 2016. Non-invasive techniques for exact phenotypic assessment of carcass composition and tissue growth in domestic animals. *Acta Agriculturae Slovenica* (suppl. 5), 12-17.
- Lin J, Wang C and Redies C 2014. Restricted expression of classic cadherins in the spinal cord of the chicken embryo. *Frontiers in Neuroanatomy* 8, 1-17.
- Lopes MS, Silva FF, Harlizius B, Duijvesteijn N, Lopes PS, Guimarães SEF and Knol EF 2013. Improved estimation of inbreeding and kinship in pigs using optimized SNP panels. *BMC Genetics* 14, 92. doi: 10.1186/1471-2156-14-92.
- Lopes MS, Bovenhuis H, van Son N, Nordbø Ø, Grindflek EH, Knol EF and Bastiaansen JWM 2017. Using markers with large effect in genetic and genomic predictions. *Journal of Animal Science* 95, 59-71.
- Lu D, Ventura-Holman T, Li J, McMurray RW, Subauste JS and Maher JF 2005. Abnormal glucose homeostasis and pancreatic islet function in mice with inactivation of the *Fem1b* Gene. *Molecular and Cellular Biology* 25, 6570-6577.
- Onteru SK, Gorbach DM, Young JM, Garrick DJ, Dekkers JCM and Rothschild MF 2013. Whole genome association studies of residual feed intake and related traits in the pig. *Plos One* 8, e61756.
- Patience JF, Rossoni-Serão MC and Gutiérrez NA 2015. A review of feed efficiency in swine: biology and application. *Journal of Animal Science and Biotechnology* 6, 33. doi: 10.1186/s40104-015-0031-2.
- Price AL, Zaitlen NA, Reich D and Patterson N 2010. New approaches to population stratification in genome-wide association studies. *Nature Review Genetics* 11, 459-463.
- Qiao R, Gao J, Zhang Z, Li L, Xie X, Fan Y, Cui L, Ma J, Ai H, Ren J and Huang L 2015. Genome-wide association analyses reveal significant loci and strong candidate genes for growth and fatness traits in two pig populations. *Genetic Selection Evolution* 47, 17. doi: 10.1186/s12711-015-0089-5.
- Rashidi H 2016. Breeding against infectious diseases in animals. PhD thesis, Wageningen University, Wageningen, the Netherlands.
- Resende MDV, Silva FF, Lopes PS and Azevedo CF 2012. Seleção genômica ampla (GWS) via modelos mistos (REML/BLUP), inferência bayesiana (MCMC), regressão aleatória multivariada e estatística espacial. Universidade Federal de Viçosa/Departamento de Estatística, Viçosa, Minas Gerais, Brasil.

- Reuser AJJ, Kroos M, Oude Elferink RPJO and Tager JM 1985. Defects in synthesis, phosphorylation, and maturation of acid alpha-glucosidase in glycogenosis type II. *The Journal of Biological Chemistry* 260, 8336-8341.
- Rohrer GA 2000. Identification of quantitative trait loci affecting birth characters and accumulation of backfat and weight in a Meishan-White Composite resource population. *Journal of Animal Science* 78, 2547–2553.
- Rothschild MF, Kim E-S and Plastow GS 2014. Novel tools provide new opportunities for genetic improvement of swine. In *Proceedings of the 10th World Congress of Genetics Applied to Livestock Production*, 17-22 August 2014, Vancouver, BC, Canada.
- Sahana G, Kadlecová V, Hornshøj H, Nielsen B and Christensen OF 2013. A genome-wide association scan in pig identifies novel regions associated with feed efficiency trait. *Journal of Animal Science* 91, 1041-1050.
- Sato S, Uemoto Y, Kikuchi T, Egawa S, Kohira K, Saito T, Sakuma H, Miyashita S, Arata S, Kojima T and Suzuki K 2016. SNP- and haplotype-based genome-wide association studies for growth, carcass, and meat quality traits in a Duroc multigenerational population. *BMC Genetics* 17, 60. doi: 10.1186/s12863-016-0368-3.
- Veroneze R, Bastiaansen JW, Knol EF, Guimarães SE, Silva FF, Harlizius B, Lopes MS and Lopes PS 2014. Linkage disequilibrium patterns and persistence of phase in purebred and crossbred pig (*Sus scrofa*) populations. *BMC genetics* 15, 126. doi: 10.1186/s12863-014-0126-3.
- Wang K, Liu D, Hernandez-Sanchez J, Chen J, Liu C, Wu Z, Fang M and Li N 2015. Genome wide association analysis reveals new production trait genes in a male Duroc population. *Plos One* 10, e0139207.
- Yang J, Weedon MN, Purcell S, Lettre G, Estrada K, Willer CJ, Smith AV, Ingelsson E, O'Connell JR, Mangino M, Mägi R, Madden PA, Heath AC, Nyholt DR, Martin NG, Montgomery GW, Frayling TM, Hirschhorn JN, McCarthy MI, Goddard ME, Visscher PM and the GIANT Consortium 2011. Genomic inflation factors under polygenic inheritance. *European Journal of Human Genetics* 19, 807–812.

**Tabela 1** Estatística descritiva dos grupos de informações fenotípicas e genotípicas para as características de ganho de peso médio diário (ADG), consumo de ração médio diário (ADFI) e conversão alimentar (FCR)

Descrição	ADG (g/dia)			ADFI (g/dia)			FCR (g/g)		
	P	P+G	Todos	P	P+G	Todos	P	P+G	Todos
Animais (nr)	2.578	2.386	4.964	350	2.385	2.735	349	2.383	2.732
Média	925,6±96,8	950,8±97,6	937,7±98,0	2,328±257,1	2,334±247,1	2,334±248,3	2,464±0,183	2,462±0,196	2,462±0,194
Mínimo	579	603	579	1.634	1.539	1.539	1.885	1.798	1.798
Máximo	1.258	1.280	1.280	3.046	3.083	3.083	3.034	3.237	3.237

P = animais apenas com informação fenotípica; P+G = animais com informação fenotípica e genotípica.



**Tabela 2** Componentes de variância e herdabilidades estimadas para as características de ganho de peso médio diário (ADG), consumo de ração médio diário (ADFI) e conversão alimentar (FCR) utilizando fenótipos pré-corrigidos

Característica	ADG	ADFI	FCR
Herdabilidade	0,19±0,04	0,29±0,04	0,22±0,04
Variância fenotípica	5.813,95	34.398,9	0,019

**Tabela 3** Caracterização dos SNPs com efeito significativo sobre as características de ganho de peso médio diário (ADG) e consumo de ração médio diário (ADFI)

Carac. SNP	SSC	Posição (Mb)	MAF	Efeito	$\sigma_{SNP}^2$	% $\sigma_{SNP}^2$	$-\log_{10}$ (P-valor)
<b>ALGA0006599</b>		177,01	0,49	18,87	178,0	3,06	7,09
WU_10.2_1_177201808		177,20	0,48	-19,12	182,5	3,14	6,83
ALGA0006623		178,02	0,48	19,40	188,0	3,23	6,87
<b>MARC0013872</b>	1	179,33	0,23	-17,34	106,8	1,84	5,88
WU_10.2_1_180054118		180,05	0,43	-16,33	131,0	2,25	6,17
<b>WU_10.2_1_180284104</b>		180,28	0,26	-18,67	135,7	2,33	7,01
<b>WU_10.2_1_180358223</b>		180,36	0,49	-15,09	113,8	1,96	5,98
<b>ASGA0005216</b>		185,47	0,37	-15,36	110,3	1,90	6,07
ADG							
WU_10.2_12_2865907		2,87	0,35	16,10	117,2	2,02	6,03
ALGA0064359		2,88	0,34	-16,02	116,0	1,99	5,98
WU_10.2_12_2922849		2,92	0,36	16,97	132,9	2,29	6,69
WU_10.2_12_3079587		3,08	0,34	16,02	116,0	1,99	5,98
ALGA0118028	12	3,10	0,38	-15,98	120,7	2,08	5,89
WU_10.2_12_3118577		3,12	0,38	-15,98	120,7	2,08	5,89
WU_10.2_12_3139862		3,14	0,38	-15,98	120,7	2,08	5,89
WU_10.2_12_3211882		3,21	0,34	16,01	115,8	1,99	5,98
WU_10.2_12_3221756		3,22	0,34	16,01	115,8	1,99	5,98
ADFI							
ALGA0006539		173,26	0,48	41,07	842,3	2,45	6,72
ALGA0006534		173,30	0,48	41,07	842,3	2,45	6,72
<b>ALGA0006599</b>		177,01	0,49	44,89	1007,1	2,93	6,11
<b>MARC0013872</b>	1	179,33	0,23	-54,18	1042,7	3,03	8,46
<b>WU_10.2_1_180284104</b>		180,28	0,26	-53,66	1120,0	3,26	8,64
<b>WU_10.2_1_180358223</b>		180,36	0,49	-41,07	843,1	2,45	6,64
H3GA0003232		184,27	0,44	40,26	796,8	2,32	6,93
<b>ASGA0005216</b>		185,47	0,37	-41,66	810,7	2,36	6,60

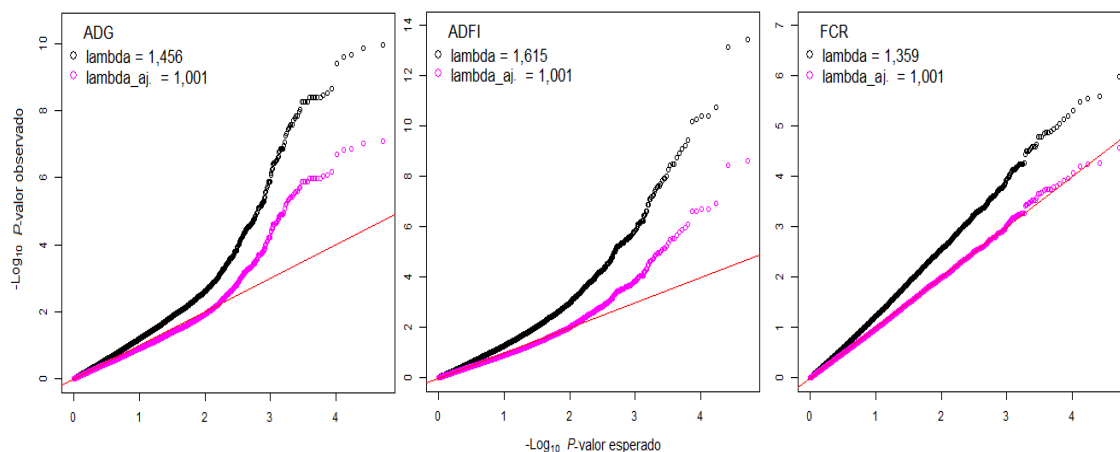
Carac. = Característica; SSC= Cromossomo *Sus scrofa*; MAF = valor da menor frequência alélica; Efeito = efeito de substituição alélica;  $\sigma_{SNP}^2$  = variância fenotípica do SNP; %  $\sigma_{SNP}^2$  = percentual da variância fenotípica do SNP

SNPs com nome em negrito apresentaram efeito significativo para ADG e ADFI.

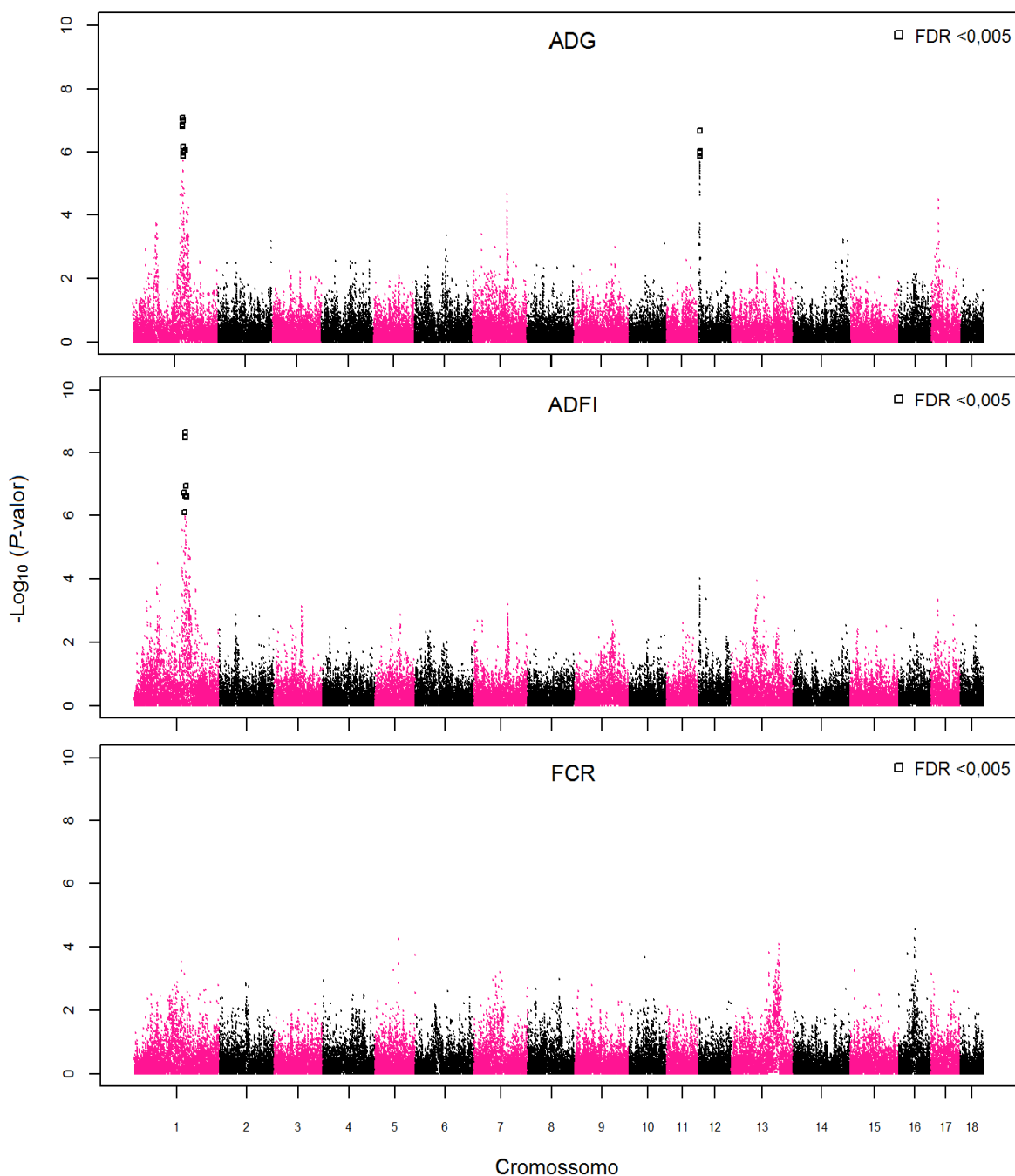
**Tabela 4** SNPs com efeito significativo sobre as características de ganho de peso médio diário (ADG) e consumo de ração médio diário (ADFI), regiões QTL associadas a estas características e genes candidatos

SSC	SNP	Posição do SNP (Mb)	Característica associada	Região QTL(Mb)	Gene candidato	Posição do gene (Mb)
	ALGA0006539	173,26	ADFI	172,26 ..174,30	<i>CDH19</i>	172,92 ..173,01
	ALGA0006534	173,30	ADFI		<i>CDH7</i>	173,58 ..173,67
	ALGA0006599	177,01	ADG / ADFI			
	WU_10.2_1_177201808	177,20	ADG		<i>RNF152</i>	177,06 ..177,07
	ALGA0006623	178,02	ADG			
1	MARC0013872	179,33	ADG / ADFI	176,01 ..181,36	<i>MC4R</i>	178,55 ..178,56
	WU_10.2_1_180054118	180,05	ADG		<i>PMAIP1</i>	178,85 ..178,86
	WU_10.2_1_180284104	180,28	ADG / ADFI			
	WU_10.2_1_180358223	180,36	ADG / ADFI			
	H3GA0003232	184,27	ADFI	183,27 ..186,47	<i>FEM1B</i>	183,92 ..183,93
	ASGA0005216	185,47	ADG / ADFI			
	WU_10.2_12_2865907	2,87	ADG			
	ALGA0064359	2,88	ADG			
	WU_10.2_12_2922849	2,92	ADG			
	WU_10.2_12_3079587	3,08	ADG			
12	ALGA0118028	3,10	ADG	1,87 .. 4,22	<i>GAA</i>	2,30 .. 2,41
	WU_10.2_12_3118577	3,12	ADG			
	WU_10.2_12_3139862	3,14	ADG			
	WU_10.2_12_3211882	3,21	ADG			
	WU_10.2_12_3221756	3,22	ADG			

SSC = Cromossomo *Sus scrofa*



**Figura 1** Gráficos quantil-quantil dos P-valores dos SNPs para as características de ganho de peso médio diário (ADG), consumo de ração médio diário (ADFI) e conversão alimentar (FCR). Os círculos na cor preta representam os P-valores, os círculos na cor magenta representam os P-valores ajustados e as linhas vermelhas representam os P-valores esperados sob hipótese nula.



**Figura 2** Estudo de associação genômica de características relacionadas à eficiência alimentar na fase de recria/terminação, representado pelo  $-\log_{10}$  (P-valor) da associação do SNP com as características de ganho de peso médio diário (ADG), consumo de ração médio diário (ADFI) e conversão alimentar (FCR) (eixo y), sobre a posição física dos SNPs nos 18 cromossomos autossômicos (eixo x). Os SNPs associados com as respectivas características (com  $FDR \leq 0,005$ ) são representados por pequenos quadros(□).