Identification of a missense variant in the porcine AGPAT gene family that affects intramuscular fat content through whole-genome sequencing

Eduard Molinero

Universitat de Lleida - Agrotecnio-CERCA Center

Ramona N. Pena

Universitat de Lleida - Agrotecnio-CERCA Center

Joan Estany

Universitat de Lleida - Agrotecnio-CERCA Center

Roger Ros-Freixedes (✉ roger.ros@udl.cat)

Universitat de Lleida - Agrotecnio-CERCA Center

Research Article

Keywords: AGPAT5, association study, backfat thickness, fatty acids, intramuscular fat, pig, whole-genome sequencing

Posted Date: February 2nd, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1321611/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Animal Genetics on September 15th, 2022. See the published version at <https://doi.org/10.1111/age.13258>.

Identification of a missense variant in the porcine *AGPAT* gene family that affects intramuscular fat content through whole-genome sequencing

Eduard Molinero, Ramona N. Pena, Joan Estany, Roger Ros-Freixedes*

Department of Animal Science, University of Lleida-Agrotecnio-CERCA Center, Lleida, Catalonia, Spain.

*Corresponding author: RRF - roger.ros@udl.cat

Abstract

The 1-acylglycerol-3-phosphate O-acyltransferases (AGPAT) are enzymes that catalyze the conversion of lysophosphatidic acid to phosphatidic acid, which is a precursor of triacylglycerol, the main fat reservoir in mammals. In pigs, there are 5 annotated genes in the *AGPAT* gene family. We used whole-genome sequencing to identify genetic variants in the *AGPAT* gene family in pigs and then we validated their association with fat content and fatty acid composition. Among the 6,639 variants identified across the *AGPAT* gene family through whole-genome sequencing of 205 Duroc pigs, we preselected a missense single nucleotide polymorphism in exon 6 of *AGPAT5* (rs196952262, A>G) based on minor allele frequency, association analyses, and linkage disequilibrium. The effect of this variant was validated in 1,105 pigs from the same Duroc line genotyped using a high-resolution melt

protocol. The A allele showed a positive additive effect for intramuscular fat ($+1.12\% \pm 0.21$, $p < 0.001$, for *gluteus medius* and $+0.89\% \pm 0.33$, $p < 0.01$, for *longissimus*). We also observed significant effects on fatty acid composition that were, at least in part, independent of the increased intramuscular fat content. Accounting for intramuscular fat content, the A allele resulted in more monounsaturated fatty acids ($+0.34\% \pm 0.15$, $p < 0.05$, for *longissimus*) and a greater monounsaturated/polyunsaturated fatty acids ratio ($+0.11 \pm 0.04$, $p < 0.01$, for *gluteus medius* and $+0.13 \pm 0.05$, $p < 0.05$, for *longissimus*). Although further studies would be needed before the causality of the variant on intramuscular fat content and fatty acid composition can be confirmed, this variant can be used as a marker in assisted selection for modulating pig fat deposition and fatty acid content.

Introduction

Intramuscular fat content (IMF) is related to organoleptic attributes and consumer acceptance of pork (Fernandez *et al.*, 1999; Huff-Lonergan *et al.*, 2002; Schwab *et al.*, 2006). In turn, intramuscular fatty acid composition has implications on the nutritional value of pork. Despite that saturated (SFA) and monounsaturated fatty acids (MUFA) provide more favorable organoleptic and technological attributes than polyunsaturated fatty acids (PUFA) (Cameron *et al.*, 2000; Wood *et al.*, 2008; Bekhit *et al.*, 2013; Wood and Enser, 2017), the intake of saturated fatty acids has been associated to increased risk of cardiovascular disease (Christophersen and Haug, 2011; Calder, 2015; Kapoor *et al.*, 2021). Intramuscular fat content and fatty acid composition are especially relevant traits for consumer preferences for high-quality products (Estany *et al.*, 2017) such as premium fresh pork or dry-cured products, and, as a consequence, they have been included in the selection objectives of pig lines aimed at this production. However, selection for IMF has been hindered by its unfavorable positive genetic correlation to backfat thickness (Newcom *et al.*, 2005; Suzuki *et al.*, 2005; Schwab *et al.*, 2010; Ros-

Freixedes *et al.*, 2013), which is typically included in selection objectives of pig breeding programs as an indicator of carcass composition on production efficiency. Similarly, the technical handicaps for routinely measuring fatty acid composition in live selection candidates and the strong correlation structure between fatty acid contents (Ros-Freixedes and Estany, 2014; Zhang *et al.*, 2019) have limited selection for fatty acid composition. To overcome these difficulties, a lot of efforts have been put into finding genetic variants that affect IMF and fatty acid composition (e.g., Puig-Oliveras *et al.*, 2016; Ros-Freixedes *et al.*, 2016; Won *et al.*, 2018; Ding *et al.*, 2019; Pena *et al.*, 2019; Zhang *et al.*, 2021) that can be used for marker-assisted selection.

In mammals, the main fat reservoir is in the form of triacylglycerol. Within skeletal muscle, triacylglycerols are mainly stored in the adipocytes but they can also be stored as droplets in the myocyte cytoplasm (Machann *et al.*, 2004; Gardan *et al.*, 2006). Triacylglycerols are primarily synthesized by sequential esterification of a glycerol backbone mediated first by glycerol 3-phosphate acyltransferase (GPAT), then by 1-acylglycerol-3-phosphate O-acyltransferase (AGPAT), and later by diacylglycerol acyltransferase (DGAT) (Coleman and Lee, 2004) (Figure 1). Each of these enzymes has multiple isoforms that are encoded by a family of genes. In pig, there are 4 annotated *GPAT* genes, 5 annotated *AGPAT* genes, and 2 annotated *DGAT* genes.

A recent study showed that a single nucleotide polymorphism (SNP) in the *DGAT2* gene is associated with intramuscular palmitoleic acid content in pigs (Solé *et al.*, 2021). Because the *AGPAT* genes are involved in the same metabolic pathway of triacylglycerol synthesis, those genes are plausible candidates for underlying the genetic variation of fat content and composition traits. In particular, the AGPAT enzymes catalyze the conversion of lysophosphatidic acid to phosphatidic acid. The phosphatidic group is hydrolyzed by a phosphatidic phosphatase (LPIN and PLPP) to produce diacylglycerol that later the DGAT enzyme takes as a substrate. Phosphatidic acid is not only a precursor of triacylglycerol, but also a precursor of various glycerophospholipids and of signaling

65 molecules involved in multiple regulatory processes, such as phosphatidylinositol, which is involved
66 in insulin signaling (Coleman and Lee, 2004). Mutations in the *AGPAT* genes have been associated to
67 body fat mass in humans (*AGPAT2*; Agarwal *et al.*, 2002), mice (*AGPAT1* and *AGPAT2*; Cortés *et al.*,
68 2009; Vogel *et al.*, 2011; Agarwal *et al.*, 2017), and buffalo (*AGPAT1* and *AGPAT6*; Xiaoya Ma *et al.*,
69 2022), and to milk fat content in goat (*AGPAT6*; He *et al.*, 2011). However, only scarce information
70 can be found on *AGPAT* genes in pigs. One study in Berkshire pigs described a missense SNP
71 (rs19695226) in the *AGPAT5* gene that contributes to meat color, cooking loss and carcass temperature
72 (Park *et al.*, 2017).

73 Whole-genome sequencing is a powerful tool for variant detection. Whole-genome sequence data
74 include rare and population-specific variants, including causative mutations (Daetwyler *et al.*, 2014;
75 Nicod *et al.*, 2016; Schaid *et al.*, 2018). The objective of this study was to use whole-genome
76 sequencing to identify variants in the *AGPAT* genes and then to validate the association of preselected
77 variants with carcass and meat quality traits, mainly backfat thickness and IMF content and
78 composition, in pigs.

79 **Materials and Methods**

80 **Ethics statement**

81 Pigs used in this study were raised and slaughtered in commercial units complying with the regulations
82 and good practice guidelines on the protection of animals kept for farming purposes, during transport
83 and slaughter (Royal Decree 37/2014, Spain). Tissue samples used in this study were collected at the
84 slaughterhouse. The Ethical Committee on Animal Experimentation of the University of Lleida
85 approved all experimental procedures.

86 **Animals and phenotypes**

87 A total of 1,105 Duroc pigs from 182 sires and 585 dams of the same line were used in this experiment.
88 Animals were raised in 16 batches between 2002 and 2019 following a common protocol for data
89 recording and tissue sampling (Ros-Freixedes *et al.*, 2013; Ros-Freixedes *et al.*, 2016). Pigs from each
90 batch were raised from 75 days until slaughter age (210 days, SD 9.4) under identical conditions.
91 During this time, animals had ad libitum access to commercial feed (Esporc, Riudarenes, Girona,
92 Spain). At 180 days of age (SD 8.8), pigs were weighted and backfat and loin thickness were measured
93 between the third and fourth last ribs by a portable ultrasonic scanner (Piglog 105, Frontmatec,
94 Kolding, Denmark). All pigs were slaughtered in the same abattoir, where carcass weight, carcass
95 backfat and carcass loin thickness at 6 cm off the midline between the third and fourth last ribs were
96 recorded using an on-line ultrasound automatic scanner (AutoFOM, SFK-Technology, Denmark).

97 After chilling for about 24 h at 2 °C, samples of the muscles *gluteus medius* (n=1,105) and *longissimus*
98 (n=492) were collected, vacuum-packaged, and stored at -20 °C until required. Samples of
99 subcutaneous fat (n=354) were also collected and stored in the same way. The IMF content in *gluteus*
100 *medius* and *longissimus*, as well as the fatty acid composition of both muscles and of subcutaneous fat,
101 were determined in duplicate by quantitative gas chromatography (Bosch *et al.*, 2009). Fatty acids were
102 expressed as percentages relative to total fatty acid content. The proportion of SFA (C14:0, C16:0,
103 C18:0 and C20:0), MUFA (C16:1n-7, C18:1n-7, C18:1n-9 and C20:1n-9), and PUFA (C18:2n-6,
104 C18:3n-3, C20:2n-6 and C20:4n-6) were calculated.

105 **Whole-genome sequencing**

106 Genomic DNA from a subset of 205 pigs was isolated from *gluteus medius* samples using a standard
107 protocol. The DNA samples were submitted to Centre Nacional d'Anàlisi Genòmica (CNAG-CRG,
108 Barcelona, Spain) for sequencing. The short-insert paired-end libraries for the whole-genome
109 sequencing were prepared with PCR free protocol using KAPA HyperPrep kit (Roche, Basel,

Switzerland) with some modifications. In short, in 0.4-1.0 µg of genomic DNA was sheared on a Covaris™ LE220-Plus focused-ultrasonicator (Covaris, Brighton, UK) in order to reach the fragment size of ~400 bp. The fragmented DNA was size-selected for the fragment size of 220-550 bp with AMPure XP beads (Agencourt, Beckman Coulter, Nyon, Switzerland). The size-selected genomic DNA fragments were end-repaired, adenylated and ligated to adaptors with unique dual indexes and unique molecular identifiers that were compatible with the Illumina platform (Integrated DNA Technologies, Leuven, Belgium). The libraries were quality controlled on an Agilent 2100 Bioanalyzer with the DNA 7500 assay (Agilent, Madrid, Spain) for size and quantified by Kapa Library Quantification Kit for Illumina platforms (Roche).

The libraries were sequenced on NovaSeq6000 (Illumina, San Diego, CA, USA) in paired-end mode with a read length of 2x151+17+8 bp following the manufacturer's protocol for dual indexing. Image analysis, base calling and quality scoring of the run were processed using the manufacturer's software Real Time Analysis (RTA 3.4.4, Illumina) and followed by generation of FASTQ sequence files. A minimum of 20 Gb of sequencing data was generated per sample.

Sequencing data processing and identification of variants

DNA sequence reads were pre-processed using Trimmomatic (Bolger *et al.*, 2014) to remove adapter sequences from the reads. We mapped the reads to the reference genome Sscrofa11.1 (GenBank accession: GCA_000003025.6; Warr *et al.*, 2020) using the BWA-MEM algorithm (Li, 2013). Duplicates were marked with Picard (<http://broadinstitute.github.io/picard>). SNPs and short insertions and deletions (indels) were identified with the variant caller GATK HaplotypeCaller (GATK 3.8.0) (DePristo *et al.*, 2011; Poplin *et al.*, 2017) using default settings. The average realized sequencing coverage was 7.9x (SD 2.4x). Variant discovery with GATK HaplotypeCaller was performed separately for each individual and then a joint variant set for all the individuals in each population was

133 obtained by extracting the variant positions from all the individuals. We retained all biallelic variants
134 for further analyses with VCFtools (Danecek *et al.*, 2011).

135 Variants in the porcine *AGPAT* genes were retrieved including all variants in the transcription units
136 and 500 bp upstream of the proximal promoter of the gene (for *AGPAT1*: SSC7, 24,204,902 to
137 24,213,693 bp; for *AGPAT2*: AEMK02000682.1, 936,421 to 1,106,715 bp; for *AGPAT3*: SSC13,
138 206,750,603 to 206,906,353 bp; for *AGPAT4*: SSC1, 6,772,006 to 6,883,555 bp; and for *AGPAT5*:
139 SSC15, 37,800,616 to 37,853,472 bp). Note that the *AGPAT2* gene is located in an unplaced scaffold
140 in the current genome assembly. Although the *GPAT4* was previously named *AGPAT6*, we did not
141 include it in our study because it is no longer considered a member of the *AGPAT* family and its
142 ortholog showed no acylglycerol acyltransferase activity in humans (Chen *et al.*, 2008).

143 **Preselection of candidate variants for fat traits**

144 To preselect those variants with greater evidence of association with the studied traits among all called
145 variants, we performed an association study on the 205 sequenced pigs. To do so, we used a single-
146 marker regression approach based on a univariate linear mixed model that accounted for the genomic
147 relationship matrix. We used the GEMMA 0.96 software (Zhou and Stephens, 2012) to fit the following
148 model:

$$149 \quad \mathbf{y} = \mathbf{W}\mathbf{b} + \mathbf{x}_j\alpha_j + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

150 where $\mathbf{y}$ is the vector of phenotypic values (backfat thickness, IMF, or fatty acid composition); $\mathbf{W}$ is
151 the incidence matrix for fixed effects; $\mathbf{b}$ is the vector of the fixed effects; $\mathbf{x}_j$ is the vector of genotypes
152 of the j^{th} SNP; α_j is the allele substitution effect of the j^{th} SNP; $\mathbf{Z}$ is the incidence matrix for the random
153 individual polygenic effects; $\mathbf{u}$ is the vector of the random individual polygenic effects; and $\mathbf{e}$ is the
154 vector of residual terms. The fixed effects included the intercept, the fattening batch (14 levels), the

genotypes of two polymorphisms with major impact on fat content and fatty acid composition (rs709596309 for *LEPR* and rs80912566 for *SCD*; 3 genotypes each) (Ros-Freixedes *et al.*, 2016), and the slaughter age as a covariate. Even though *DGAT2* is involved in the same metabolic pathway, we did not include it in the model since no interaction was observed. The random individual polygenic effects and residual terms were assumed to follow normal distributions $\mathbf{u} \sim N(0, \mathbf{K}\sigma_u^2)$ and $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$, respectively, where $\mathbf{K}$ is the genomic relationship matrix, σ_u^2 is the additive genetic variance, $\mathbf{I}$ is an identity matrix, and σ_e^2 is the residual variance. We focused the analyses on the variants with minor allele frequency equal or greater than 0.20 so that all genotypes were sufficiently represented in the 205 pigs. Those variants with p-value equal or lower than 0.001 for any of the tested traits were prioritized as candidate variants. Among those, we prioritized candidate variants in promoters, 5' and 3' untranslated regions (UTR) and exons based on their predicted functional annotation using VEP (McLaren *et al.*, 2016). Candidate variants were also examined for linkage disequilibrium to avoid redundancy using PLINK 1.9 (Chang *et al.*, 2015). Based on these criteria, a single variant was preselected for further validation.

Genotyping of the candidate variant

The SNP rs196952262 (A>G) in the *AGPAT5* gene (SSC15 at 37,843,344 bp) was genotyped in 1,105 pigs using 5'-GTCCCTTCGAAAGCCACTGT-3' as forward primer and 5'-CACCAAGAATAAAGGCAACCCA-3' as reverse primer. Amplifications were performed by real-time PCR (QuantStudio3, Applied Biosystems, Thermo Scientific, Waltham, MA, USA) with High-Resolution Melt analysis (Luminaris Colour HRM Master Mix, Thermo Scientific) using 20 ng of genomic DNA and 0.4 μ M of each primer in 5 μ L final volume reaction. Thermocycling conditions were 95 °C for 10 min, and 40 cycles of 95 °C for 15 sec, 60 °C for 1 min, followed by a high-resolution melting curve starting with a denaturation at 95 °C for 15 sec, annealing at 60 °C for 1 min and a slow ramp at 0.015 °C/sec up to 95 °C. High Resolution Melt software v3.1 (Applied Biosystems, Thermo

Scientific) was used for the analysis of melting data and sample genotyping. All pigs were also genotyped for SNPs in genes *SCD* (rs80912566, C>T, on SSC14) and *LEPR* (rs709596309, C>T, on SSC6) following the protocols described in (Estany *et al.*, 2014; Solé *et al.*, 2022) and (Ros-Freixedes *et al.*, 2016), respectively. In a subset of 807 pigs, the genotype for *DGAT2* (ss7315407085, G>A, on SSC9) was also available (Solé *et al.*, 2021).

Validation of the *AGPAT5* variant

The effect of the *AGPAT5* rs196952262 SNP genotype on production traits (body weight, carcass weight, backfat thickness, loin thickness, and IMF) and fatty acid composition was estimated using a model with the same fixed effects as above (with 16 batches) plus the genotype for *AGPAT5* (GG, AG and AA). The effect of the genotype was tested using the F-statistic. Multiple pairwise comparisons among *AGPAT5* genotypes were tested with the Tukey HSD test. The additive and dominant effects of the *AGPAT5* SNP were also estimated by replacing the genotype with two covariates coded as (-1, 0, 1) and (0, 1, 0) for the GG, AG and AA genotypes, respectively. The IMF content was added as a covariate to test whether the effect of the *AGPAT5* SNP on fatty acid composition was due to changes in fat content. Finally, we also tested the interaction of the *AGPAT5* SNP with *LEPR* and *SCD*, as well as with *DGAT2*, by adding this latter SNP to previous models. All the analyses were performed using the statistical package JMP Pro 15 (SAS Institute Inc., Cary, NC, USA).

Results

Preselection of candidate variants for fat traits

Our analysis of the genetic variation of the *AGPAT* gene family revealed that *AGPAT5* contained the strongest candidate variants for fat content and fatty acid composition traits. We found large differences in the number of variants located in the transcription unit of the genes of the *AGPAT* family (Table 1).

In total, 6,639 variants were detected across the 205 pigs, 2,220 of which had minor allele frequency equal or greater than 0.20 (referred to here as ‘common’ variants). Of these, 171 were prioritized as candidate variants because they showed associations with $p\text{-value} \leq 0.001$ for at least one of the tested traits. Most of these candidate variants (166 out of 210 common variants) were located on *AGPAT5*, only 1 (out of 120 common variants) was located on *AGPAT2*, and 4 (out of 442 common variants) on *AGPAT4*. The five candidate variants in *AGPAT2* and *AGPAT4* were intronic.

Of the 210 common variants detected in *AGPAT5*, only 11 were in the promoter, 5’ and 3’ untranslated and exonic regions of the gene (Table 2). Three of these variants were located in exons and were predicted to be missense mutations. One of these (SNP3) was predicted to be missense for both annotated transcripts for the gene, while the other two were missense for only one of the two transcripts. SNP3 was also the missense variant that showed the highest association with the tested traits ($p \leq 0.001$). Regarding the other 8 variants, one was located in the promoter region, one in the 5’-UTR and 6 (3 SNPs and 3 indels) in the 3’-UTR. Linkage disequilibrium of the candidate variants showed that all markers were in moderate linkage disequilibrium ($r^2 > 0.7$) with SNP3 (Table 2). Taking into account these results, together with the fact that SNP3 (rs196952262) was previously reported in Berkshire pigs as associated to meat quality traits (Park *et al.*, 2017), we preselected this variant as a tag SNP for further validation.

Validation of the AGPAT5 variant

The effect of the *AGPAT5* SNP rs196952262 on fat content was validated using data from 1,105 pigs (Table 3). The *AGPAT5* SNP was associated with production traits at 180 d of age, with the A allele showing a positive additive effect on backfat thickness ($+0.60 \text{ mm} \pm 0.27$, $p < 0.05$) and a negative effect on loin thickness ($-0.85 \text{ mm} \pm 0.35$, $p < 0.05$). These effects were less clearly detected when measured on carcass around 4 weeks later and only loin thickness showed a similar trend (-0.96 mm

224 ± 0.52 , $p < 0.10$). The A allele also was positively associated with IMF, both in *gluteus medius* ($+1.12\%$
225 ± 0.31 , $p < 0.001$) and in *longissimus* ($+0.89\% \pm 0.33$, $p < 0.01$).

226 Differences in intramuscular fatty acid composition were also detected across *AGPAT5* genotypes. The
227 A allele had a consistent effect in both muscles (Tables 4 and 5). The A allele had a positive additive
228 effect on MUFA ($+0.35\% \pm 0.13$, $p < 0.01$, in *gluteus medius* and $+0.50\% \pm 0.16$, $p < 0.01$, in
229 *longissimus*), mainly due to its effect on C18:1n-9. Also, the A allele showed a negative effect on
230 PUFA ($-0.45\% \pm 0.13$, $p < 0.001$, in *gluteus medius*, and $-0.54\% \pm 0.18$, $p < 0.01$, in *longissimus*),
231 with decreased values for C18:2n-6, C18:3n-3 and C20:4n-6. As a result, the A allele showed a positive
232 effect on the MUFA/PUFA ratio (0.22 ± 0.04 , $p < 0.001$, in *gluteus medius*, and 0.34 ± 0.09 , $p < 0.001$,
233 in *longissimus*). Since the effects of the A allele on MUFA and PUFA counteract each other, the
234 *AGPAT5* genotype did not influence the SFA/(MUFA+PUFA) ratio. We observed no differences in
235 fatty acid composition for subcutaneous fat (Supplementary Table S1).

236 With the exception of MUFA and C18:1n-9, the effects of the A allele on fatty acid composition were
237 mostly due to changes in fat content, since they were not detected when adjusted for IMF
238 (Supplementary Tables S2 and S3). The effects of the A allele on MUFA, which was driven mainly by
239 its effect on C18:1n-9, was in part independent to IMF, although this was more evident in *longissimus*
240 ($+0.34\% \pm 0.15$, $p < 0.05$, for MUFA and $+0.27\% \pm 0.15$, $p < 0.10$, for C18:1n-9) than in *gluteus*
241 *medius* ($+0.21\% \pm 0.12$, $p < 0.10$, for MUFA and $+0.16\% \pm 0.10$, $p = 0.11$, for C18:1n-9). As a result,
242 the A allele displayed a positive additive effect on the MUFA/PUFA ratio in both *gluteus medius*
243 ($+0.11 \pm 0.04$, $p < 0.01$) and *longissimus* ($+0.13 \pm 0.05$, $p < 0.05$) even after adjusting for IMF.

244 We detected a significant interaction between *AGPAT5* and *LEPR* for IMF (Supplementary Table S4).
245 Although the effect of *AGPAT5* on IMF showed the same trend within each *LEPR* genotype, the
246 additive effect of the A allele of *AGPAT5* was larger in the *LEPR*-TT pigs ($+1.47\% \pm 0.72$, $p < 0.05$),

247 which are fatter, than in the *LEPR*-CC or CT pigs ($+0.99\% \pm 0.34$, $p < 0.01$). We also confirmed the
248 effect of *AGPAT5* on IMF in the *DGAT2*-GG pigs (Supplementary Table S5), although statistical
249 evidence for its effect within the other *DGAT2* genotypes was limited to assess the interaction between
250 these two genes. We detected some significant interactions for fatty acid composition between
251 *AGPAT5* and genes *LEPR* (Supplementary Table S4) and *SCD* (Supplementary Table S6), but they did
252 not show clear patterns that were biologically meaningful. There were no interactions between
253 *AGPAT5* and *DGAT2* for fatty acid composition (Supplementary Table S5).

254 **Discussion**

255 In this study, we identified variants in the family of *AGPAT* genes using whole-genome sequencing
256 and preselected a candidate variant for fat content and composition traits. Using a large number of
257 individuals, we validated the effect of the rs196952262 SNP as a tag SNP for the haplotype in *AGPAT5*.
258 In the following we will discuss: (1) the suitability of whole-genome sequencing for preselecting
259 candidate variants in the *AGPAT* gene family associated to traits of interest, and (2) the effect of the
260 *AGPAT5* SNP on fat content and composition.

261 **Preselection of candidate variants in the *AGPAT* gene family**

262 Whole-genome sequencing is a powerful tool for detecting large numbers of variants that could be
263 associated with complex traits. However, the identification of casual variants remains challenging
264 because the high density of variants that are in high linkage disequilibrium hinders the disentangling
265 of the causal variant. We found the strongest evidence of association with fat content and composition
266 traits for *AGPAT5*. In contrast, we did not find any candidate variant in *AGPAT1* or *AGPAT3*, and the
267 few candidate variants in *AGPAT2* and *AGPAT4* were discarded because they were located in introns.
268 A typical criterion for prioritizing candidate variants is to limit the search to coding and promoter
269 regions, since the prediction of the potential effects of variants located in non-coding regions from

DNA sequence is not straightforward (Johnsson and Jungnickel, 2021). However, in some instances non-coding variants, which may have regulatory functions, have been proposed as candidate variants (Van Laere *et al.*, 2003; Ryan *et al.*, 2012; Solé *et al.*, 2021), while variants in coding regions have often been found to have a small impact on complex traits (Koufariotis *et al.*, 2018; Xiang *et al.*, 2019). Despite that mutations in *AGPAT2* that affected triacylglycerol synthesis and storage have been described in humans (Agarwal *et al.*, 2002), we did not detect any orthologous polymorphism in pigs with significant association to the studied traits.

The results for *AGPAT5* are consistent with the fact that this gene is the most expressed of the *AGPAT* gene family in skeletal muscle in pigs. Using data from a previous RNA-Seq experiment (Solé *et al.*, 2022), we found that *AGPAT5* and *AGPAT2* are the most expressed genes in muscle *semimembranosus* in pig (Supplementary Figure S1). Those results are also in line with previous gene expression data reported in pig skeletal muscle (Freeman *et al.*, 2012). In contrast, expression analysis in humans showed that *AGPAT1* and *AGPAT2* were highly expressed in adipocytes and skeletal muscle (Agarwal *et al.*, 2002; Prasad, Garg and Agarwal, 2011), while in mice *AGPAT1* and *AGPAT2* and *AGPAT4* showed the highest overall expression (Vergnes *et al.*, 2006). Although our expression data were insufficient to detect any potential expression differences between *AGPAT5* genotypes (Supplementary Figure S2), the fact that in pigs *AGPAT5* is the most expressed gene of the *AGPAT* family, in contrast to human and mice, supports the contribution of this isoform in the conversion of lysophosphatidic into phosphatidic acid in pig skeletal muscle, and hence, in triacylglycerol synthesis. If these results were confirmed, they could help understand differences in the metabolic pathways of fat deposition of pigs compared to other mammals.

We preselected the rs196952262 SNP as a tag variant for the *AGPAT5* gene based on several criteria. Coincidentally, the same SNP was previously reported to be associated with meat quality in Berkshire pigs (Park *et al.*, 2017) for traits such as meat color, cooking loss and carcass temperature, albeit not

for backfat thickness. In particular, these authors found that the A allele was associated to a lighter pork color (greater CIE L* parameter). Although these authors did not analyze IMF, the lighter color associated to the A allele is compatible with greater levels of IMF (Schwab *et al.*, 2006; Suárez-Mesa *et al.*, 2021), which adds yet another layer of evidence from an independent population in support of this variant.

Effect of the *AGPAT5* SNP on fat content and composition

We validated the association of rs196952262 SNP from *AGPAT5* on fat content and composition. This association had not been identified in previous analyses in Duroc. We observed that the A allele had a positive additive effect on backfat thickness at 180 d and on IMF in both muscles. Alterations on the gene activity led to lower triacylglycerol synthesis, and therefore to a reduced fat accumulation (Cortés *et al.*, 2009). Although we expected to observe the same additive effect on carcass backfat thickness, we were not able to detect this effect on backfat thickness, likely because of differences in age and measuring methods for live and carcass backfat thickness, but collectively results indicate an effect on overall carcass fatness. The rs196952262 SNP explained 1.2% and 1.4% of the genetic variance for backfat and loin thickness at 180 days, respectively, but much less at 210 days (0.03% and 0.6%, respectively). The rs196952262 SNP had a higher impact on IMF than on backfat thickness. The SNP explained 2.3% and 2.5% of the genetic variance for IMF in *gluteus medius* and *longissimus*, respectively. Moreover, we observed differences in fatty acid composition that were specific of intramuscular fat and did not affect subcutaneous fat, which might further indicate a lower impact on backfat thickness. Since the A allele also had a negative effect on loin thickness and thus, attention should be paid to any unfavorable correlated responses in carcass lean content.

The A allele segregated at high frequency in Iberian pigs (0.81, n=18), which is consistent with the body fatness that characterizes this breed, and at intermediate frequency in much leaner breeds such as

Pietrain (0.47, n=28) and a Large White × Landrace crossbred (0.50, n=49). Despite this, it segregated at lower frequencies in other heavily marbled breeds such as Duroc, both the population studied here (0.24) and in two other Duroc populations (0.15, n=26, and 0.25, n=20), and Berkshire (0.21; Park *et al.*, 2017). The lower frequency in Duroc pigs indicated that selection assisted by this marker could increase the population average IMF in *gluteus medius* by up to +1.7% if the A allele was fixed.

The rs196952262 SNP was also associated with intramuscular fatty acid composition. The effects that we observed for *AGPAT5* on fatty acid composition were consistent with the faster accumulation of SFA and MUFA relative to PUFA as fat reserves grow (De Smet *et al.*, 2004; Ros-Freixedes *et al.*, 2016; Zhang *et al.*, 2019). However, we found some evidence of an effect of *AGPAT5* on fatty acid composition that was, at least in part, independent to its effect on fat content. In particular, *AGPAT5* showed an IMF-independent effect on MUFA, mainly driven by C18:1n-9. The rs196952262 SNP explained 0.5% of the genetic variance for MUFA in *gluteus medius* and 1.0% in *longissimus*. In humans, functional studies of the AGPAT proteins showed that both AGPAT3 and AGPAT5 had a higher affinity for C18:1n-9 as a substrate, which could explain our results (Prasad *et al.*, 2011). As a consequence of this, we found significant differences for the MUFA/PUFA ratio between genotypes and the SNP could be used as a marker for selecting pork with a higher monounsaturated fatty acid profile.

The effect of *AGPAT5* on IMF may become more noticeable in fatter pigs. Thus, *AGPAT5* may show significant interactions with genes that affect overall fatness, such as *LEPR*. In turn, this can produce changes on fatty acid composition, to the extent that composition indirectly reflects changes in fat content. However, there were no clear interactions between *AGPAT5* and *SCD* despite that both genes affect MUFA. Even though the *AGPAT5* gene intervenes in the same metabolic pathway as *DGAT2*, in a previous study we found no effect of *DGAT2* on IMF (Solé *et al.*, 2021) and therefore the interaction of these two genes for IMF seems unlikely. In contrast, *DGAT2* has a specific effect on

341 intramuscular C16:1n-7 (Solé *et al.*, 2021) but we did not detect any effect of *AGPAT5* on this fatty
342 acid.

343 In conclusion, our analysis of the genetic variation in the sequence of the genes of the *AGPAT* family
344 revealed that *AGPAT5* contained the strongest candidate variant for fat content and fatty acid
345 composition traits. The A allele of the rs196952262 variant increases IMF and MUFA but it negatively
346 affects loin thickness. The effect of *AGPAT5* on IMF is more noticeable in fatter pigs and, therefore,
347 *AGPAT5* could interact with genes that affect overall fatness, such as *LEPR*. Although further studies
348 would be needed before the causality of the variant can be confirmed, this variant can be used as a
349 selection marker for modulating pig fat deposition and fatty acid composition of pork.

350

351 **Data Availability Statement**

352 The original contributions presented in the study are included in the article and supplementary files.
353 Further inquiries can be directed to the corresponding author.

354 **Conflict of Interest**

355 The authors declare that the research was conducted in the absence of any commercial or financial
356 relationships that could be construed as a potential conflict of interest.

357 **Author Contributions**

358 JE and RRF designed the study; EM and RNP processed the samples; EM and RRF performed the
359 statistical analyses; EM and RRF wrote the first draft; RNP and JE contributed to the interpretation of
360 the results and provided comments on the manuscript. All authors read and approved the final
361 manuscript.

Funding

This research was supported by the Spanish Ministry of Science, Innovation & Universities and the EU Regional Development Funds (grant RTI2018-101346-B-I00). EM is recipient of a UdL-Santander Predoc scholarship.

Acknowledgments

We acknowledge the personnel at Selecció Batallé for their cooperation for the recording of on-farm data and sample collection. We gratefully acknowledge Pilar Sopena from the Animal Breeding group, University of Lleida, for laboratory assistance.

References

- Agarwal, A.K., Arioglu, E., de Almeida, S., Akkoc, N., Taylor, S.I., Bowcock, A.M., Barnes, R.I. and Garg, A. (2002) AGPAT2 is mutated in congenital generalized lipodystrophy linked to chromosome 9q34, *Nat. Genet.* 31, 21–23. doi:10.1038/ng880.
- Agarwal, A.K., Tunison, K., Dalal, J.S., Nagamma, S.S., Hamra, F.K., Sankella, S., Shao, X., Auchus, R.J. and Garg, A. (2017) Metabolic, Reproductive, and Neurologic Abnormalities in Agpat1-Null Mice, *Endocrinology*. 158, 3954–3973. doi:10.1210/en.2017-00511.
- Bekhit, A.E.-D.A., Hopkins, D.L., Fahri, F.T. and Ponnampalam, E.N. (2013) Oxidative Processes in Muscle Systems and Fresh Meat: Sources, Markers, and Remedies, *Compr. Rev. Food Sci. Food Saf.* 12, 565–597. doi:10.1111/1541-4337.12027.
- Bolger, A.M., Lohse, M. and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data, *Bioinform.* 30, 2114–2120. doi:10.1093/bioinformatics/btu170.
- Bosch, L., Tor, M., Reixach, J. and Estany, J. (2009) Estimating intramuscular fat content and fatty acid composition in live and post-mortem samples in pigs, *Meat Sci.* 82, 432–437. doi:10.1016/j.meatsci.2009.02.013.
- Calder, P.C. (2015) Functional Roles of Fatty Acids and Their Effects on Human Health, *J. Parenter. Enteral Nutr.* 39, 18S-32S. doi:10.1177/0148607115595980.
- Cameron, N.D., Enser, M., Nute, G.R., Whittington, F.M., Penman, J.C., Fiskén, A.C., Perry, A.M. and Wood, J.D. (2000) Genotype with nutrition interaction on fatty acid composition of

390 intramuscular fat and the relationship with flavour of pig meat, *Meat Sci.* 55, 187–195.
391 doi:10.1016/S0309-1740(99)00142-4.

392 Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M. and Lee, J.J. (2015) Second-
393 generation PLINK: rising to the challenge of larger and richer datasets, *Gigascience.* 4, s13742-015-
394 0047–8. doi:10.1186/s13742-015-0047-8.

395 Chen, Y.Q., Kuo, M.-S., Li, S., Bui, H.H., Peake, D.A., Sanders, P.E., Thibodeaux, S.J., Chu, S.,
396 Qian, Y.-W., Zhao, Y., Bredt, D.S., Moller, D.E., Konrad, R.J., Beigneux, A.P., Young, S.G. and
397 Cao, G. (2008) AGPAT6 Is a Novel Microsomal Glycerol-3-phosphate Acyltransferase *, *J. Biol.*
398 *Chem.* 283, 10048–10057. doi:10.1074/jbc.M708151200.

399 Christophersen, O.A. and Haug, A. (2011) Animal products, diseases and drugs: a plea for better
400 integration between agricultural sciences, human nutrition and human pharmacology, *Lipids Health*
401 *Dis.* 10, p. 16. doi:10.1186/1476-511X-10-16.

402 Coleman, R.A. and Lee, D.P. (2004) Enzymes of triacylglycerol synthesis and their regulation, *Prog.*
403 *Lipid Res.* 43, 134–176. doi:10.1016/S0163-7827(03)00051-1.

404 Cortés, V.A., Curtis, D.E., Sukumaran, S., Shao, X., Parameswara, V., Rashid, S., Smith, A.R., Ren,
405 J., Esser, V., Hammer, R.E., Agarwal, A.K., Horton, J.D. and Garg, A. (2009) Molecular
406 Mechanisms of Hepatic Steatosis and Insulin Resistance in the AGPAT2-Deficient Mouse Model of
407 Congenital Generalized Lipodystrophy, *Cell Metab.* 9, 165–176. doi:10.1016/j.cmet.2009.01.002.

408 Daetwyler, H.D., Capitan, A., Pausch, H., Stothard, P., van Binsbergen, R., Brøndum, R.F., Liao, X.,
409 Djari, A., Rodriguez, S.C., Grohs, C., Esquerré, D., Bouchez, O., Rossignol, M.-N., Klopp, C.,
410 Rocha, D., Fritz, S., Eggen, A., Bowman, P.J., Coote, D., Chamberlain, A.J., Anderson, C.,
411 VanTassell, C.P., Hulsegge, I., Goddard, M.E., Guldbrandtsen, B., Lund, M.S., Veerkamp, R.F.,
412 Boichard, D.A., Fries, R. and Hayes, B.J. (2014) Whole-genome sequencing of 234 bulls facilitates
413 mapping of monogenic and complex traits in cattle, *Nat. Genet.* 46, 858–865. doi:10.1038/ng.3034.

414 Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E.,
415 Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., Durbin, R., and 1000 Genomes Project Analysis
416 Group (2011) The variant call format and VCFtools, *Bioinform.* 27, 2156–2158.
417 doi:10.1093/bioinformatics/btr330.

418 De Smet, S., Raes, K., and Demeyer, D. (2004) Meat fatty acid composition as affected by fatness
419 and genetic factors: a review, *Anim. Res.* 53, 81–98. doi:10.1051/animres:2004003.

420 DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A.,
421 del Angel, G., Rivas, M.A., Hanna, M., McKenna, A., Fennell, T.J., Kernytsky, A.M., Sivachenko,
422 A.Y., Cibulskis, K., Gabriel, S.B., Altshuler, D. and Daly, M.J. (2011) A framework for variation
423 discovery and genotyping using next-generation DNA sequencing data, *Nat. Genet.* 43, 491–498.
424 doi:10.1038/ng.806.

425 Ding, R., Yang, M., Quan, J., Li, S., Zhuang, Z., Zhou, S., Zheng, E., Hong, L., Li, Z., Cai, G.,
426 Huang, W., Wu, Z. and Yang, J. (2019) Single-Locus and Multi-Locus Genome-Wide Association
427 Studies for Intramuscular Fat in Duroc Pigs, *Front. genet.* 10. Available at:
428 <https://www.frontiersin.org/article/10.3389/fgene.2019.00619>.

429 Estany, J., Ros-Freixedes, R., Tor, M. and Pena, R.N. (2014) A Functional Variant in the Stearoyl-
430 CoA Desaturase Gene Promoter Enhances Fatty Acid Desaturation in Pork, PLoS One. 9, e86177.
431 doi:10.1371/journal.pone.0086177.

432 Estany, J., Ros-Freixedes, R., Tor, M. and Pena, R.N. (2017) TRIENNIAL GROWTH AND
433 DEVELOPMENT SYMPOSIUM: Genetics and breeding for intramuscular fat and oleic acid content
434 in pigs1, J. Anim. Sci. 95, 2261–2271. doi:10.2527/jas.2016.1108.

435 Fernandez, X., Monin, G., Talmant, A., Mourot, J. and Lebret, B. (1999) Influence of intramuscular
436 fat content on the quality of pig meat — 2. Consumer acceptability of m. longissimus lumborum,
437 Meat Sci. 53, 67–72. doi:10.1016/S0309-1740(99)00038-8.

438 Freeman, T.C., Ivens, A., Baillie, J.K., Beraldi, D., Barnett, M.W., Dorward, D., Downing, A.,
439 Fairbairn, L., Kapetanovic, R., Raza, S., Tomoiu, A., Alberio, R., Wu, C., Su, A.I., Summers, K.M.,
440 Tuggle, C.K., Archibald, A.L. and Hume, D.A. (2012) A gene expression atlas of the domestic pig,
441 BMC Biol. 10, 90. doi:10.1186/1741-7007-10-90.

442 Gardan, D., Gondret, F. and Louveau, I. (2006) Lipid metabolism and secretory function of porcine
443 intramuscular adipocytes compared with subcutaneous and perirenal adipocytes, Am. J. Physiol.
444 Endocrinol. Metab. 291, E372–E380. doi:10.1152/ajpendo.00482.2005.

445 he, C., Wang, C., Chang, Z.H., Guo, B.L., Li, R., Yue, X., Lan, Y., Chen, H. and Lei, C. (2011)
446 AGPAT6 polymorphism and its association with milk traits of dairy goats, Genet. Mol. Res. 10,
447 2747–56. doi:10.4238/2011.November.4.8.

448 Huff-Lonergan, E., Baas, T.J., Malek, M., Dekkers, J.C.M., Prusa, K. and Rothschild, M.F. (2002)
449 Correlations among selected pork quality traits, J. Anim. Sci. 80, 617–627.
450 doi:10.2527/2002.803617x.

451 Johnsson, M. and Jungnickel, M.K. (2021) Evidence for and localization of proposed causative
452 variants in cattle and pig genomes, Genet. Sel. Evol. 53, 67. doi:10.1186/s12711-021-00662-x.

453 Kapoor, B., Kapoor, D., Gautam, S., Singh, R. and Bhardwaj, S. (2021) Dietary Polyunsaturated
454 Fatty Acids (PUFAs): Uses and Potential Health Benefits, Curr. Nutr. Rep. 10, 232–242.
455 doi:10.1007/s13668-021-00363-3.

456 Koufariotis, L.T., Chen, Y.-P.P., Stothard, P. and Hayes, B.J. (2018) Variance explained by whole
457 genome sequence variants in coding and regulatory genome annotations for six dairy traits, BMC
458 Genom. 19, 237. doi:10.1186/s12864-018-4617-x.

459 Li, H. (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM,
460 arXiv preprint arXiv:1303.3997 [Preprint].

461 Machann, J., Häring, H., Schick, F. and Stumvoll, M. (2004) Intramyocellular lipids and insulin
462 resistance, Diabetes Obes. Metab. 6, 239–248. doi:10.1111/j.1462-8902.2004.00339.x.

463 McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., Flicek, P. and
464 Cunningham, F. (2016) The Ensembl Variant Effect Predictor, Genome Biol. 17, 122.
465 doi:10.1186/s13059-016-0974-4.

466 Newcom, D.W., Baas, T.J., Schwab, C.R. and Stalder, K.J. (2005) Genetic and phenotypic
467 relationships between individual subcutaneous backfat layers and percentage of longissimus
468 intramuscular fat in Duroc swine, *J. Anim. Sci.* 83, 316–323. doi:10.2527/2005.832316x.

469 Nicod, J., Davies, R.W., Cai, N., Hassett, C., Goodstadt, L., Cosgrove, C., Yee, B.K., Lionikaite, V.,
470 McIntyre, R.E., Remme, C.A., Lodder, E.M., Gregory, J.S., Hough, T., Joynson, R., Phelps, H., Nell,
471 B., Rowe, C., Wood, J., Walling, A., Bopp, N., Bhomra, A., Hernandez-Pliego, P., Callebert, J.,
472 Aspden, R.M., Talbot, N.P., Robbins, P.A., Harrison, M., Fray, M., Launay, J.-M., Pinto, Y.M.,
473 Blizard, D.A., Bezzina, C.R., Adams, D.J., Franken, P., Weaver, T., Wells, S., Brown, S.D.M.,
474 Potter, P.K., Klenerman, P., Lionikas, A., Mott, R. and Flint, J. (2016) Genome-wide association of
475 multiple complex traits in outbred mice by ultra-low-coverage sequencing, *Nat. Genet.* 48, 912–918.
476 doi:10.1038/ng.3595.

477 Park, W.B., An, S.M., Yu, G.E., Kwon, S., Hwang, J.H., Park, D.H., Kang, D.G., Kim, T.W., Park,
478 H.C., Ha, J. and Kim, C.W. (2017) The rs196952262 Polymorphism of the AGPAT5 Gene is
479 Associated with Meat Quality in Berkshire Pigs, *Korean J Food Sci Anim Resour.* 37, 926–930.
480 doi:10.5851/kosfa.2017.37.6.926.

481 Pena, R.N., Noguera, J.L., García-Santana, M.J., González, E., Tejeda, J.F., Ros-Freixedes, R. and
482 Ibáñez-Escriche, N. (2019) Five genomic regions have a major impact on fat composition in Iberian
483 pigs, *Sci. Rep.* 9, 2031. doi:10.1038/s41598-019-38622-7.

484 Poplin, R., Ruano-Rubio, V., DePristo, M.A., Fennell, T.J., Carneiro, M.O., Van der Auwera, G.A.,
485 Kling, D.E., Gauthier, L.D., Levy-Moonshine, A. and Roazen, D. (2017) Scaling accurate genetic
486 variant discovery to tens of thousands of samples, *BioRxiv*, 201178.

487 Prasad, S.S., Garg, A. and Agarwal, A.K. (2011) Enzymatic activities of the human AGPAT isoform
488 3 and isoform 5: localization of AGPAT5 to mitochondria [S], *J. Lipid Res.* 52, 451–462.
489 doi:10.1194/jlr.M007575.

490 Puig-Oliveras, A., Revilla, M., Castelló, A., Fernández, A.I., Folch, J.M. and Ballester, M. (2016)
491 Expression-based GWAS identifies variants, gene interactions and key regulators affecting
492 intramuscular fatty acid content and composition in porcine meat, *Sci. Rep.* 6, 31803.
493 doi:10.1038/srep31803.

494 Ros-Freixedes, R. and Estany, J. (2014) On the compositional analysis of fatty acids in pork, *J.*
495 *Agric. Biol. Environ. Stat.* 19, 136–155.

496 Ros-Freixedes, R., Gol, S., Pena, R.N., Tor, M., Ibáñez-Escriche, N., Dekkers, J.C.M. and Estany, J.
497 (2016) Genome-Wide Association Study Singles Out SCD and LEPR as the Two Main Loci
498 Influencing Intramuscular Fat Content and Fatty Acid Composition in Duroc Pigs, *PLoS One.* 11,
499 e0152496. doi:10.1371/journal.pone.0152496.

500 Ros-Freixedes, R., Reixach, J., Bosch, L., Tor, M. and Estany, J. (2013) Response to selection for
501 decreased backfat thickness at restrained intramuscular fat content in Duroc pigs1, *J. Anim. Sci.* 91,
502 3514–3521. doi:10.2527/jas.2013-6282.

503 Ryan, M.T., Hamill, R.M., OHalloran, A.M., Davey, G.C., McBryan, J., Mullen, A.M., McGee, C.,
504 Gispert, M., Southwood, O.I. and Sweeney, T. (2012) SNP variation in the promoter of the

505 PRKAG3 gene and association with meat quality traits in pig, *BMC Genet.*, 13, 66.
506 doi:10.1186/1471-2156-13-66.

507 Schaid, D.J., Chen, W. and Larson, N.B. (2018) From genome-wide associations to candidate causal
508 variants by statistical fine-mapping, *Nat. Rev. Genet.* 19, 491–504. doi:10.1038/s41576-018-0016-z.

509 Schwab, C.R., Baas, T.J. and Stalder, K.J. (2010) Results from six generations of selection for
510 intramuscular fat in Duroc swine using real-time ultrasound. II. Genetic parameters and trends¹, *J.*
511 *Anim. Sci.* 88, 69–79. doi:10.2527/jas.2008-1336.

512 Schwab, C.R., Baas, T.J., Stalder, K.J. and Mabry, J.W. (2006) Effect of long-term selection for
513 increased leanness on meat and eating quality traits in Duroc swine¹, *J. Anim. Sci.* 84, 1577–1583.
514 doi:10.2527/2006.8461577x.

515 Solé, E., González-Prendes, R., Oliinychenko, Y., Tor, M., Ros-Freixedes, R., Estany, J. and Pena,
516 R.N. (2022) Transcriptome shifts triggered by vitamin A and SCD genotype interaction in Duroc
517 pigs, *BMC Genom.* 23, 16. doi:10.1186/s12864-021-08244-3.

518 Solé, E., Ros-Freixedes, R., Tor, M., Pena, R.N. and Estany, J. (2021) A sequence variant in the
519 diacylglycerol O-acyltransferase 2 gene influences palmitoleic acid content in pig muscle, *Sci. Rep.*
520 11, 14797. doi:10.1038/s41598-021-94235-z.

521 Suárez-Mesa, R., Ros-Freixedes, R., Tor, M., Reixach, J., Pena, R.N. and Estany, J. (2021) Leptin
522 receptor and fatty acid desaturase-2 gene variants affect fat, color and production profile of dry-cured
523 hams, *Meat Sci.* 173, 108399. doi:10.1016/j.meatsci.2020.108399.

524 Suzuki, K., Irie, M., Kadowaki, H., Shibata, T., Kumagai, M. and Nishida, A. (2005) Genetic
525 parameter estimates of meat quality traits in Duroc pigs selected for average daily gain, longissimus
526 muscle area, backfat thickness, and intramuscular fat content, *J. Anim. Sci.* 83, 2058–2065.
527 doi:10.2527/2005.8392058x.

528 Van Laere, A.-S., Nguyen, M., Braunschweig, M., Nezer, C., Collette, C., Moreau, L., Archibald,
529 A.L., Haley, C.S., Buys, N., Tally, M., Andersson, G., Georges, M. and Andersson, L. (2003) A
530 regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig, *Nature.* 425,
531 832–836. doi:10.1038/nature02064.

532 Vergnes, L., Beigneux, A.P., Davis, R., Watkins, S.M., Young, S.G. and Reue, K. (2006) *Agpat6*
533 deficiency causes subdermal lipodystrophy and resistance to obesity, *J. Lipid Res.* 47, 745–754.
534 doi:10.1194/jlr.M500553-JLR200.

535 Vogel, P., Read, R., Hansen, G., Wingert, J., DaCosta, C.M., Buhring, L.M. and Shadoan, M. (2011)
536 Pathology of Congenital Generalized Lipodystrophy in *Agpat2*^{−/−} Mice, *Vete. Pathol.* 48, 642–654.
537 doi:10.1177/0300985810383870.

538 Warr, A., Affara, N., Aken, B., Beiki, H., Bickhart, D.M., Billis, K., Chow, W., Eory, L., Finlayson,
539 H.A., Flicek, P., Girón, C.G., Griffin, D.K., Hall, R., Hannum, G., Hourlier, T., Howe, K., Hume,
540 D.A., Izuogu, O., Kim, K., Koren, S., Liu, H., Manchanda, N., Martin, F.J., Nonneman, D.J.,
541 O'Connor, R.E., Phillippy, A.M., Rohrer, G.A., Rosen, B.D., Rund, L.A., Sargent, C.A., Schook,
542 L.B., Schroeder, S.G., Schwartz, A.S., Skinner, B.M., Talbot, R., Tseng, E., Tuggle, C.K., Watson,

543 M., Smith, T.P.L. and Archibald, A.L. (2020) An improved pig reference genome sequence to enable
544 pig genetics and genomics research, *Gigascience*. 9, giaa051. doi:10.1093/gigascience/giaa051.

545 Won, S., Jung, J., Park, E. and Kim, H. (2018) Identification of genes related to intramuscular fat
546 content of pigs using genome-wide association study, *Asian-australas. J. Anim. Sci.* 31, 157–162.
547 doi:10.5713/ajas.17.0218.

548 Wood, J.D. and Enser, M. (2017) Chapter 20 - Manipulating the Fatty Acid Composition of Meat to
549 Improve Nutritional Value and Meat Quality, ed. *New Aspects of Meat Quality*. Woodhead
550 Publishing, 501–535. doi:10.1016/B978-0-08-100593-4.00023-0.

551 Wood, J.D., Enser, M., Fisher, A.V., Nute, G.R., Sheard, P.R., Richardson, R.I., Hughes, S.I. and
552 Whittington, F.M. (2008) Fat deposition, fatty acid composition and meat quality: A review, *Meat*
553 *Sci.*, 78, 343–358. doi:10.1016/j.meatsci.2007.07.019.

554 Xiang, R., Berg, I. van den, MacLeod, I.M., Hayes, B.J., Prowse-Wilkins, C.P., Wang, M.,
555 Bolormaa, S., Liu, Z., Rochfort, S.J., Reich, C.M., Mason, B.A., Vander Jagt, C.J., Daetwyler, H.D.,
556 Lund, M.S., Chamberlain, A.J. and Goddard, M.E. (2019) Quantifying the contribution of sequence
557 variants with regulatory and evolutionary significance to 34 bovine complex traits, *PNAS*, 116,
558 19398. doi:10.1073/pnas.1904159116.

559 Xiaoya Ma, Anqin Duan, Xingrong Lu, Md Mahmodul Hasan Sohel, Hamdy Abdel-Shafy, Ahmed
560 Amin, Shasha Liang, and Tingxian Deng (2022) Novel Insight Into The Potential Role of AGPAT
561 Gene Family in Triacylglycerols Synthesis in Buffalo [Preprint]. doi:10.21203/rs.3.rs-382481/v1.

562 Zhang, Y., Zhang, J., Gong, H., Cui, L., Zhang, W., Ma, J., Chen, C., Ai, H., Xiao, S., Huang, L. and
563 Yang, B. (2019) Genetic correlation of fatty acid composition with growth, carcass, fat deposition
564 and meat quality traits based on GWAS data in six pig populations, *Meat Sci.*, 150, 47–55.
565 doi:10.1016/j.meatsci.2018.12.008.

566 Zhang, Z., Zhang, Z., Oyelami, F.O., Sun, H., xu, Z., Ma, P., Wang, Q. and Pan, Y. (2021)
567 Identification of genes related to intramuscular fat independent of backfat thickness in Duroc pigs
568 using single-step genome-wide association, *Anim. Genet.* 52, 108–113. doi:10.1111/age.13012.

569 Zhou, X. and Stephens, M. (2012) Genome-wide efficient mixed-model analysis for association
570 studies, *Nat. Genet.* 44, 821–824. doi:10.1038/ng.2310.

571

572 **Figure captions**

573

574 **Figure 1.** Triacylglycerol synthesis pathway. GPAT: glycerol-3-phosphate acyltransferase; AGPAT:
575 1-acyl-glycerol-3-phosphate O-acyltransferase; LPIN: lipin; PLPP: phospholipid phosphatase;
576 DGAT: diacylglycerol acyltransferase; and P_i: phosphate.

577

578 **Tables**

579 **Table 1.** Number of total called, common and candidate variants in the gene transcription unit plus
580 500-bp upstream of the transcription start sites of the pig *AGPAT* gene family.

Gene	Total called variants	Common variants	Candidate variants
<i>AGPAT1</i>	55	0	0
<i>AGPAT2</i>	1469	120	1
<i>AGPAT3</i>	2045	1448	0
<i>AGPAT4</i>	2722	442	4
<i>AGPAT5</i>	348	210	166
Total	6639	2220	171

581 Common variants are variants with a minor allele frequency ≥ 0.2 . Candidate variants are variants that showed an
582 association with at least one trait ($p \leq 0.001$).

Table 2. Summary of candidate non-intronic variants identified in the pig *AGPAT5* gene

Variant	Position (bp)	Location	Major/minor allele	MAF ^a	Association study ^b	Predicted consequence ^c	Linkage disequilibrium ^d
SNP1	37,800,993	Promoter	C/T	0.266	2.75	-	0.85
SNP2	37,801,250	5'-UTR	A/C	0.263	4.23	-	0.79
SNP3	37,843,344	Exon 6	G/A	0.259	3.38	missense (2/2)	1.00
SNP4	37,850,594	Exon 8	G/C	0.263	2.88	missense (1/2)	0.77
SNP5	37,850,596	Exon 8	G/T	0.263	2.88	missense (1/2)	0.77
INDEL6	37,851,301	3'-UTR	G/GAAAC	0.278	3.99	-	0.76
SNP7	37,852,012	3'-UTR	C/A	0.281	3.61	-	0.70
SNP8	37,852,233	3'-UTR	C/T	0.231	3.26	-	0.70
SNP9	37,852,886	3'-UTR	C/T	0.278	4.30	-	0.76
INDEL10	37,853,314	3'-UTR	G/GT	0.276	3.74	-	0.80
INDEL11	37,853,443	3'-UTR	G/GA	0.272	3.13	-	0.78

^a MAF: minor allele frequency.

^b P-value, expressed as $-\log(\text{p-value})$, of the most significant association of the variant with any of the studied traits.

^c Predicted consequence on the anngene transcripts. In parentheses, the number of transcripts affected over the total number of transcripts.

^d Linkage disequilibrium between the variant and SNP3, measured as the squared coefficient of correlation between alleles (r^2).

Table 3. Least-square means ($\pm$ standard errors) for live and carcass weight and composition by *AGPAT5* (rs196952262, A>G) genotype.

Trait	<i>AGPAT5</i> genotype			Additive effect ¹		Dominant effect	
	AA (n=71)	AG (n=390)	GG (n=644)	a	p-value	d	p-value
Live measurements (180 d)							
Body weight, kg	110.5 $\pm$ 1.5	110.4 $\pm$ 0.6	109.5 $\pm$ 0.5	0.51 $\pm$ 0.78	0.51	0.38 $\pm$ 0.95	0.69
Backfat thickness, mm	20.5 $\pm$ 0.5	19.7 $\pm$ 0.2	19.3 $\pm$ 0.2	0.60 $\pm$ 0.27	0.02	-0.20 $\pm$ 0.32	0.53
Loin thickness, mm	43.2 $\pm$ 0.7 ^a	45.0 $\pm$ 0.3 ^b	44.9 $\pm$ 0.2 ^b	-0.85 $\pm$ 0.35	0.02	0.98 $\pm$ 0.43	0.02
Carcass measurements (210 d)							
Carcass weight, kg	96.9 $\pm$ 1.2	98.0 $\pm$ 0.52	96.8 $\pm$ 0.4	0.01 $\pm$ 0.63	0.98	1.15 $\pm$ 0.78	0.14
Backfat thickness, mm	24.0 $\pm$ 0.4	24.3 $\pm$ 0.18	23.8 $\pm$ 0.1	0.09 $\pm$ 0.22	0.66	0.36 $\pm$ 0.27	0.19
Loin thickness, mm	42.2 $\pm$ 1.0 ^{ab}	42.8 $\pm$ 0.43 ^a	44.1 $\pm$ 0.3 ^b	-0.96 $\pm$ 0.52	0.07	-0.32 $\pm$ 0.64	0.62
Lean content, %	40.4 $\pm$ 0.7	40.1 $\pm$ 0.3	40.9 $\pm$ 0.2	-0.29 $\pm$ 0.36	0.43	0.57 $\pm$ 0.45	0.21
Intramuscular fat, % dry matter							
<i>M. gluteus medius</i>	20.42 $\pm$ 0.59 ^a	18.85 $\pm$ 0.26 ^b	18.18 $\pm$ 0.20 ^b	1.12 $\pm$ 0.31	<0.001	-0.45 $\pm$ 0.39	0.24
<i>M. longissimus</i> ²	15.65 $\pm$ 0.64 ^a	14.45 $\pm$ 0.31 ^{ab}	13.86 $\pm$ 0.26 ^b	0.89 $\pm$ 0.33	0.007	-0.31 $\pm$ 0.41	0.46

¹ Additive allele substitution of G by A.² Sample size was 38, 187 and 267 for AA, AG, and GG, respectively.Bold font indicates statistical significance ($p \leq 0.05$).^{a,b} Within trait, means with different superscripts indicate statistically significant differences ($p \leq 0.05$).

Table 4. Least-square means ($\pm$ standard errors) for fatty acid composition in muscle *gluteus medius* by *AGPAT5* (rs196952262, A>G) genotype.

Trait	<i>AGPAT5</i> genotype			Additive effect ¹		Dominant effect	
	AA (n=71)	AG (n=390)	GG (n=644)	a	p-value	d	p-value
Fatty acid, %							
SFA	37.92 $\pm$ 0.24	37.83 $\pm$ 0.11	37.73 $\pm$ 0.08	0.10 $\pm$ 0.13	0.46	0.01 $\pm$ 0.16	0.95
C16:0	24.39 $\pm$ 0.15	24.43 $\pm$ 0.07	24.31 $\pm$ 0.05	0.05 $\pm$ 0.08	0.60	0.08 $\pm$ 0.10	0.42
C18:0	11.73 $\pm$ 0.11	11.56 $\pm$ 0.05	11.59 $\pm$ 0.04	0.07 $\pm$ 0.06	0.22	-0.10 $\pm$ 0.07	0.16
MUFA	50.37 $\pm$ 0.25 ^a	49.75 $\pm$ 0.11 ^b	49.66 $\pm$ 0.09 ^b	0.35 $\pm$ 0.13	0.006	-0.26 $\pm$ 0.16	0.11
C16:1n-7	3.70 $\pm$ 0.07	3.72 $\pm$ 0.03	3.66 $\pm$ 0.02	0.02 $\pm$ 0.04	0.61	0.03 $\pm$ 0.05	0.47
C18:1n-7	4.26 $\pm$ 0.04	4.30 $\pm$ 0.02	4.29 $\pm$ 0.02	-0.01 $\pm$ 0.02	0.47	0.03 $\pm$ 0.03	0.28
C18:1n-9	42.02 $\pm$ 0.19 ^a	41.66 $\pm$ 0.10 ^{ab}	41.50 $\pm$ 0.09 ^b	0.26 $\pm$ 0.11	0.01	-0.09 $\pm$ 0.13	0.48
PUFA	11.71 $\pm$ 0.24 ^a	12.41 $\pm$ 0.11 ^b	12.61 $\pm$ 0.08 ^b	-0.45 $\pm$ 0.13	<0.001	0.24 $\pm$ 0.16	0.13
C18:2n-6	9.43 $\pm$ 0.19 ^a	9.89 $\pm$ 0.08 ^{ab}	10.03 $\pm$ 0.06 ^b	-0.30 $\pm$ 0.10	0.002	0.16 $\pm$ 0.12	0.18
C18:3n-3	0.56 $\pm$ 0.01	0.57 $\pm$ 0.01	0.58 $\pm$ 0.01	-0.01 $\pm$ 0.01	0.10	0.01 $\pm$ 0.01	0.30
C20:4n-6	1.20 $\pm$ 0.07 ^a	1.37 $\pm$ 0.03 ^b	1.46 $\pm$ 0.02 ^c	-0.13 $\pm$ 0.03	<0.001	0.045 $\pm$ 0.04	0.30
Fatty acid ratio							
MUFA/SFA	1.35 $\pm$ 0.02	1.33 $\pm$ 0.01	1.34 $\pm$ 0.01	0.01 $\pm$ 0.01	0.55	-0.01 $\pm$ 0.01	0.44
MUFA/PUFA	4.57 $\pm$ 0.09 ^a	4.23 $\pm$ 0.04 ^b	4.14 $\pm$ 0.03 ^b	0.22 $\pm$ 0.04	<0.001	-0.13 $\pm$ 0.06	0.03
SFA/(MUFA+PUFA) (x10)	6.16 $\pm$ 0.06	6.14 $\pm$ 0.03	6.11 $\pm$ 0.02	0.02 $\pm$ 0.03	0.49	0.00 $\pm$ 0.04	0.95

SFA: saturated fatty acids, MUFA: monounsaturated fatty acids, PUFA: polyunsaturated fatty acids, C16:0: palmitic acid, C18:0: stearic acid, C16:1n-7: palmitoleic acid, C18:1n-7: vaccenic acid, C18:1n-9: oleic acid, C18:2n-6: linoleic acid, C18:3n-3: linolenic acid, and C20:4n-6: arachidonic acid.

¹ Additive allele substitution of G by A.

Bold font indicates statistical significance ($p \leq 0.05$).

^{a-c} Within trait, means with different superscripts indicate statistically significant differences ($p \leq 0.05$).

Table 5. Least-square means ($\pm$ standard errors) for fatty acid composition in muscle *longissimus* by *AGPAT5* (rs196952262, A>G) genotype

Trait	<i>AGPAT5</i> genotype			Additive effect ¹		Dominant effect	
	AA (n=38)	AG (n=187)	GG (n=267)	a	p-value	d	p-value
Fatty acid, %							
SFA	39.56 $\pm$ 0.34	39.51 $\pm$ 0.16	39.48 $\pm$ 0.14	0.04 $\pm$ 0.17	0.82	-0.01 $\pm$ 0.22	0.97
C16:0	25.40 $\pm$ 0.20	25.29 $\pm$ 0.10	25.29 $\pm$ 0.18	0.06 $\pm$ 0.10	0.59	-0.06 $\pm$ 0.13	0.65
C18:0	12.46 $\pm$ 0.17	12.45 $\pm$ 0.08	12.41 $\pm$ 0.07	0.02 $\pm$ 0.09	0.80	0.02 $\pm$ 0.11	0.87
MUFA	51.22 $\pm$ 0.32 ^a	50.50 $\pm$ 0.16 ^{ab}	50.22 $\pm$ 0.13 ^b	0.50 $\pm$ 0.16	0.002	-0.22 $\pm$ 0.20	0.29
C16:1n-7	4.00 $\pm$ 0.09	3.92 $\pm$ 0.04	3.97 $\pm$ 0.03	0.02 $\pm$ 0.04	0.72	-0.06 $\pm$ 0.06	0.31
C18:1n-7	4.34 $\pm$ 0.06	4.33 $\pm$ 0.03	4.34 $\pm$ 0.03	-0.00 $\pm$ 0.03	0.97	-0.01 $\pm$ 0.04	0.77
C18:1n-9	42.07 $\pm$ 0.29 ^a	41.73 $\pm$ 0.15 ^{ab}	41.30 $\pm$ 0.14 ^b	0.39 $\pm$ 0.16	0.01	0.05 $\pm$ 0.21	0.82
PUFA	9.21 $\pm$ 0.34 ^a	9.99 $\pm$ 0.17 ^{ab}	10.30 $\pm$ 0.14 ^b	-0.54 $\pm$ 0.18	0.002	0.23 $\pm$ 0.22	0.31
C18:2n-6	7.18 $\pm$ 0.24 ^a	7.73 $\pm$ 0.12 ^{ab}	7.93 $\pm$ 0.10 ^b	-0.37 $\pm$ 0.13	0.003	0.18 $\pm$ 0.16	0.26
C18:3n-3	0.34 $\pm$ 0.01	0.37 $\pm$ 0.01	0.37 $\pm$ 0.01	-0.01 $\pm$ 0.01	0.03	0.01 $\pm$ 0.01	0.14
C20:4n-6	1.31 $\pm$ 0.09 ^a	1.46 $\pm$ 0.05 ^{ab}	1.57 $\pm$ 0.04 ^b	-0.13 $\pm$ 0.05	0.006	0.02 $\pm$ 0.06	0.7
Fatty acid ratio							
MUFA/SFA	1.31 $\pm$ 0.02	1.29 $\pm$ 0.01	1.29 $\pm$ 0.01	0.01 $\pm$ 0.01	0.22	-0.01 $\pm$ 0.01	0.57
MUFA/PUFA	5.78 $\pm$ 0.18 ^a	5.26 $\pm$ 0.09 ^b	5.11 $\pm$ 0.07 ^b	0.34 $\pm$ 0.09	<0.001	-0.18 $\pm$ 0.12	0.12
SFA/(MUFA+PUFA) (x10)	6.59 $\pm$ 0.09	6.57 $\pm$ 0.04	6.57 $\pm$ 0.04	0.01 $\pm$ 0.05	0.82	-0.01 $\pm$ 0.06	0.91

SFA: saturated fatty acids, MUFA: monounsaturated fatty acids, PUFA: polyunsaturated fatty acids, C16:0: palmitic acid, C18:0: stearic acid, C16:1n-7: palmitoleic acid, C18:1n-7: vaccenic acid, C18:1n-9: oleic acid, C18:2n-6: linoleic acid, C18:3n-3: linolenic acid, and C20:4n-6: arachidonic acid.

¹ Additive allele substitution of G by A.

Bold font indicates statistical significance ($p \leq 0.05$).

^{a,b} Within trait, means with different superscripts indicate statistically significant differences ($p \leq 0.05$).

Figures

Figure 1

Triacylglycerol synthesis pathway. GPAT: glycerol-3-phosphate acyltransferase; AGPAT: 1-acyl-glycerol-3-phosphate O-acyltransferase; LPIN: lipin; PLPP: phospholipid phosphatase; DGAT: diacylglycerol acyltransferase; and P_i: phosphate.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AGPAT5frontiersEMG20220202preprintSUPP.pdf](#)