Identification of eQTLs associated with lipid metabolism in *Longissimus dorsi* muscle of pigs with different genetic backgrounds

Lourdes Criado-Mesas, María Ballester, Daniel Crespo-Piazueto, Anna Castelló, Ana I. Fernández & Josep M. Folch

Intramuscular fat content and its fatty acid composition affect porcine meat quality and its nutritional value. The present work aimed to identify genomic variants regulating the expression in the porcine muscle (*Longissimus dorsi*) of 45 candidate genes for lipid metabolism and fatty acid composition in three experimental backcrosses based on the Iberian breed. Expression genome-wide association studies (eGWAS) were performed between the muscle gene expression values, measured by real-time quantitative PCR, and the genotypes of 38,426 SNPs distributed along all chromosomes. The eGWAS identified 186 eSNPs located in ten *Sus scrofa* regions and associated with the expression of *ACSM5*, *ACSS2*, *ATF3*, *DGAT2*, *FOS* and *IGF2* (FDR < 0.05) genes. Two expression quantitative trait loci (eQTLs) for *IGF2* and *ACSM5* were classified as cis-acting eQTLs, suggesting a mutation in the same gene affecting its expression. Conversely, ten eQTLs showed trans-regulatory effects on gene expression. When the eGWAS was performed for each backcross independently, only three common trans-eQTL regions were observed, indicating different regulatory mechanisms or allelic frequencies among the breeds. In addition, hotspot regions regulating the expression of several genes were detected. Our results provide new data to better understand the functional regulatory mechanisms of lipid metabolism genes in muscle.

Studies on the traits that determine the quality of pork meat and their derived products have received increasing attention in recent years. The intramuscular fat (IMF) content and its fatty acid (FA) composition are considered determinant for meat quality, playing a central role in the nutritional values of the meat. IMF influences meat flavour, juiciness, tenderness and firmness, which are important traits for consumer acceptance. On the other hand, its FA composition will determine how healthy is the product since it is well-known that some FAs are essential for humans, such as ω-3 and ω-6 polyunsaturated FAs (PUFAs).

During the last years, pig breeding companies have produced commercial pigs that grow faster and have superior carcasses. However, these carcasses have become leaner having less IMF and, therefore, producing a decrease in the meat quality according to consumers. Otherwise, local breeds such as the Iberian pig present a high-fat deposition and FA desaturation values and have a special interest in the production of high-quality dry-cured cuts, such as loin and ham. Often the Iberian pig is crossed with other breeds to improve its reproductive and growth traits, although crossing has been associated with a decrease in meat quality.

Several studies agree that genetic factors can determine intramuscular FA composition in pigs. For example, significant breed effects have been reported for IMF, water binding capacity, colour, and tenderness. Thus, differences according to the genetic background have made the industry aware of it when improving the meat quality of pork.

In recent years, genome-wide association studies (GWAS) have been used to detect genetic variants involved in FA composition traits, unravelling the complex genetic basis of these quantitative traits. In general, genes...
accumulation of triglycerides, so they have a higher risk to gain fat and develop obesity. In a similar way, female
PPARA activation between sexes seems to be influenced by the female hormone estrogen which has been shown
in our study, PPARs members as key regulatory genes for lipid metabolism. In humans, males tend to
present higher activity in lipolytic pathways, with a lower risk to gain fat and develop obesity than females. In
our study, PPARα, PPARδ, and PPARγ1 presented higher expression in males than in females. Differences in
PPARA expression induce genes involved in fatty acid import and oxidation, increasing lipid oxidation and decreasing
triacylglycerol accumulation. In addition, PPARδ have been shown to increase glucose uptake and prevent insu-
lin resistance. Finally, PPARγ1 is a transcriptional co-activator that cooperates with PPARα to promote
mitochondrial oxidative metabolism. On the contrary, human females present higher rates of lipogenesis and
accumulation of triglycerides, so they have a higher risk to gain fat and develop obesity. In a similar way, female
pigs seem to develop obesity more readily than male pigs. In this regard, the higher expression of SREBP1C,
PPARG related to lipogenic pathways were more expressed in BC1_LD (DGAT2 BC1_LD and BC1_DU respectively, and are involved in a wide range of functions. In summary, genes more whereas genes related to lipolytic pathways were higher expressed in BC1_DU (ACSS1, ACSS2) genes. Remarkably, among this second group of co-expressed genes strong correlations for ELF1 and NCOA6 were identified linking both groups of co-expressed and CREG1, PLIN5, PXMP3, NCOA2, and CREG1 which were previously identified in the BC1_LD study. CREG1 and PRKAA1 were identified linking both groups of co-expressed genes. Remarkably, among this second group of co-expressed genes strong correlations for IGF2, NCOA1, NCOA6, and PDHX were identified and the highest node degree corresponded to NCOA6 and PDHX.

Hence, genes involved in lipogenesis are highly correlated in porcine muscle, suggesting a coordinated regulation of its expression. In previous studies of our group, two groups of extreme animals for intramuscular fatty acid composition were analysed by RNA-Seq, showing an increase in fatty acid and glucose uptake, and also an increase in the lipogenesis pathway in the muscle of pigs with higher levels of MUFA and SFA content. Moreover, a transcription factor that regulates the expression of a broad range of lipogenic genes such as SCD, together with the higher expression of the nuclear receptor PPARG, a master regulator of adipogenesis and obesity, may explain the higher number of over-expressed genes observed in females related to lipogenic pathways and fat deposition. Furthermore, it is relevant to highlight the over-expression in females of other nuclear receptors such as LXRα and RXRγ. LXRαs activation in human skeletal cells promote increased uptake, synthesis, utilization and storage of lipid. Finally, the IGF2 gene, which has been involved in muscle growth and fat deposition, showed also a higher expression in females.

A breed effect on the expression of genes involved in energy balance and lipogenesis was reported in a comparison between Iberian and Duroc pigs. In our study, a significant backcross effect on gene-expression levels was detected in 37 out of the 45 genes analysed: ACAA2, ACSS1, ACSS2, ALB, ANGPT1, AQP7, MLXIPL, CPT1B, CREG1, CROT, DGAT1, DGAT2, ELF1, ETS1, FABP5, FOS, HIF1AN, IGF2, LXRα, MGLL, NCOA2, NCOA6, NFKB, PDHX, PIK3R1, PLIN5, PRKAA1, PRKACB, PPARG, PPARGC1A, RXRγ, SCD, SETD7, SLC2A4, SP1 and SREBP1C (Fig. 2). Overall, 18 and 16 out of 45 genes were over-expressed in BC1_LD and BC1_DU respectively, and are involved in a wide range of functions. In summary, genes more related to lipogenic pathways were more expressed in BC1_LD (DGAT2, PPARG, NCOA2, SCD and PRKAA1) whereas genes related to lipolytic pathways were higher expressed in BC1_DU (ACSS1, ACSS2, CPT1B, PPARA and PPARGC1A). Finally, 3 out of 45 genes were over-expressed in BC1_PI and were mainly related to transcriptional regulation and control (ELF1, ETS1 and PLIN5).

The high number of lipid-related genes differentially expressed between breeds may be caused by different selection according to breed in pigs, such as in fat content. In our animal material, the Iberian pig is characterized by a high content of SFA, MUFA and IMF, conferring a good meat quality. On the other hand, Landrace pigs are considered very lean with less IMF and high PUFA content, Duroc pigs are more fat with an improved carcass meat quality and Pietrain animals present a good carcass conformation. Moreover, other studies reported differences in the expression of genes involved in specific metabolic pathways depending on FA traits. These differences in gene expression pattern may indicate differences in gene-expression regulatory mechanisms among breeds, which are described below.
Puig-Oliveras et al. (2016), showed a positive correlation between lipogenic genes and palmitoleic and octadecnoic fatty acids, and a positive correlation between lipolytic genes and PUFAs, specifically with the linoleic fatty acid in the BC1_LD animals.

Genome-wide association studies for gene expression and eQTL identification. An eGWAS was performed with the muscle gene expression values and the genotypes of 38,426 single nucleotide polymorphisms (SNPs) distributed along all chromosomes in 355 3BCs animals. The eGWAS identified 186 expression-SNPs (eSNPs) located in 10 Sus scrofa chromosomes (SSC) regions of SSC1, SSC2, SSC3, SSC6, SSC7, SSC11, SSC13 and SSC16 and associated with the expression of ACSM5, ACSS2, ATF3, DGAT2, FOS and IGF2 (FDR < 0.05) genes (Supplementary Table S1). Ten eQTLs showed trans-regulatory effects on gene expression and two of them, IGF2 and ACSM5, were also classified as cis-acting, suggesting that there is a mutation in the same gene or in a proximal genomic region affecting its expression (Table 1). Both cis and trans-eQTLs were represented in Fig. 4.

Cis-eQTLs. For the IGF2 cis-eQTL region, the IGF2g.3072G > A SNP was the most significantly associated polymorphism (p-value = 3.24 × 10⁻⁴⁴) and explained the 70% of the muscle IGF2 expression variance, approximatively (Fig. 5).
The IGF2:g.3072 G > A substitution has been identified as the causal mutation of an imprinted QTL for muscle growth, fat deposition and heart size and it is maternally imprinted in most animal tissues. The IGF2:g.3072 G > A mutation is located in a well-conserved CpG island, which is hypomethylated and abrogates the binding site for an IGF2 transcriptional repressor called ZBDE6, leading to a three-fold up-regulation of the IGF2 expression in pig skeletal muscle.

An imprinting model was tested for muscle gene expression in 327 animals in which the paternal allele was deduced from progenitor's genotypes (Fig. 6). Animals with the paternally-inherited A allele (A²) of the IGF2:g.3072 G > A polymorphism showed the highest IGF2 gene expression in muscle (AA: NQ mean = 2.29, n = 130 and A²G: NQ mean = 2.65, n = 26) compared to animals with paternally-inherited G allele (A²G: NQ mean = 0.65, n = 122 and GG: NQ mean = 0.78, n = 76).

Therefore, the IGF2:g.3072 G > A SNP genotype and the imprinting model explained the differences observed in IGF2 gene expression in muscle, being the IGF2 genetic variant the major regulator of gene expression in muscle in different genetic backgrounds (see below specific data for each backcross).
A previous study of our group reported that IGF2 polymorphism was also the most significant associated SNP with IGF2 mRNA expression in adipose tissue, but it explained only 25% of the phenotypic variance compared to the 70% explained in muscle tissue, suggesting that other genetic variants, potentially trans-regulation as reported in the current study, may affect the gene expression in adipose tissue. Nevertheless, the IGF2 gene expression followed a maternal imprinting model in both tissues.

The ACSM5 gene, target of the other cis-eQTL region identified, is involved in pathways such as conjugation of carboxylic acids and FA beta-oxidation. A SSC3 cis-eQTL was reported in a previous study of our group analysing the ACSM5 expression in BC1_LD population. The ACSM5 proximal promoter region was amplified and sequenced in ten BC1_LD animals and subsequently three polymorphisms were found. The most proximal 5’ mutation, rs331702081 (hereinafter known as ACSM5.P) was the most significantly associated SNP with the ACSM5 gene expression in the BC1_LD population. Thus, in the current study the ACSM5.P was genotyped in the BC1_DU and BC1_PI populations.

In the eGWAS with all three backcrosses the ACSM5.P SNP presented the strongest association with muscle ACSM5 gene expression (p-value = 1.39 × 10^{-27}) (Fig. 7). The polymorphism located in the promoter region explained approximately the 40% of the phenotypic variance, suggesting the presence of additional genetic factors regulating its gene expression (see below specific data for each backcross). Further analysis should be done to understand the transcriptional regulation of ACSM5 gene.

In a previous study of our group the ACSM5.P mutation has been also described as the most significantly associated SNP with ACSM5 gene expression in backfat adipose tissue of the BC1_LD population. Nonetheless, the correlation between the ACSM5 gene expression in backfat and muscle was 0.60, suggesting that the gene expression in both tissues could be regulated by different genetic variants. In addition, two transcription factors (ARNT and STAT6) that bind only with the A allele is present were identified. Hence, genetic variation on the promoter region of ACSM5 could be a key regulator of the ACSM5 gene expression, at least in muscle and adipose tissues.
Trans-eQTLs. A total of 783 genes were located in the 10 trans-eQTL genomic regions identified in our study. Among them, we identified potential lipid metabolism regulatory genes in three regions (Table 1: interval 1, 9 and 10). The ACSM5 eGWAS revealed a trans-eQTL located in the 18.5 Mb – 53.6 Mb region of SSC3, where the Interleukin 4 Receptor (ILR4) gene was mapped. Polymorphisms in ILR4 have been associated with high density lipoprotein–cholesterol levels, suggesting the possible role of ILAR gene in lipid metabolism in humans. The FBJ Murine Osteosarcoma Viral Oncogene Homolog (FOS) eGWAS revealed a trans-eQTL in the 8.9 Mb – 19.7 Mb region of SSC11, where a gene involved in lipid metabolism was mapped: Forkhead Box O1 (FOXO1). From the FOXO transcription factor family, FOXO1 is the isoform with the highest expression in muscle and has been proposed as a regulator of energy metabolism and the insulin signalling pathway. It is also involved in muscle differentiation and can interact with other transcription factors such as PPARG and HNF4A to regulate insulin gene expression and IMF accumulation. Moreover, FOXO1 was found to regulate FOX gene expression in skeletal muscle, increasing their levels during cancer cachexia in humans. Retinoblastoma 1 (RB1) gene was also a transcription factor mapped in this region and is involved in gene expression control. RB1 plays an important role in cell cycle and cell differentiation and is also considered as a key regulator during adipogenesis. However, it is highly expressed in muscle tissue probably due to its role in muscle differentiation. In humans, RB1 was found co-expressed with FOS gene and is involved in proliferation and apoptosis in myosarcoma. A prediction of a functional integration network was done by GeneMANIA, showing a gene co-expression between FOS and FOXO1, a predicted functional gene relationship between FOXO1 and RB1, and FOXO1 with PPARG and HNF4, protein-protein interactions among FOXO1 and RB1 and finally a FOS, PPARG and RB1 gene pathway.

The Splicing Factor 1 (SF1) gene was mapped in the IGF2 trans-eQTL region located on SSC2 (Table 1) and it was previously described as a candidate gene for IGF2 regulation in adipose tissue. A member of the LXR nuclear receptor family named nuclear receptor subfamily 1 group H member 3 (NR1H3) was also mapped in this trans-eQTL region and chosen as a possible candidate gene due to its involvement in the deposition of lipids in pigs, which may affect lean muscle fat content.

The rest of the trans-eQTL regions were identified for ACS2 (SSC6, SSC7 and SSC13), ATF3 (SSC1 and SSC13), DGAT2 (SSC16) and FOS (SSC1). However, no candidate regulator genes could be identified in these genomic regions. This may be explained by the small intervals size, the lack of gene information in the pig assembly or the presence of other regulators such as enhancers, miRNAs and long-non-coding RNAs among others.

eGWAS analysis for each backcross independently. Expression-GWAS studies were also performed for each backcross independently and 420, 420 and 224 associated eSNPs were identified in the BC1_LD, BC1_DU and BC1_PI animals, respectively (Supplementary Table S2). A total of 26 eQTLs were found in BC1_LD located on SSC1-SSC11, SSC13 and SSC16. Of these, two trans-eQTLs were previously reported in the study with the same BC1_LD animals using the previous Sus scrofa 10.2 assembly (see below). In BC1_DU, 32 eQTLs were detected on SSC1-SSC4, SSC6-SSC7, SSC9, SSC11-SSC13, SSC15, SSC17, and SSC18, and the 25 eQTLs found in BC1_PI were located on SSC1-SSC4, SSC6-SSC7, SSC9, SSC11-SSC13, SSC15, SSC17, and SSC18, and the 25 eQTLs found co-expressed with ACSM5 gene expression in the 3BCs study. Chromosome positions in Mb were presented in the X-axis and the –log10 (p-value) is on the Y-axis. Horizontal lines represent the genome-wide significance level (FDR-based q-value < 0.1 corresponds to blue line and FDR-based q-value < 0.05 to red line).

Cis-eQTLs. The cis-eQTL regions of ACSM5 and IGF2 genes, on SSC33 and SSC2 respectively, appeared segregating in all three backcrosses, which suggest that the Iberian boars and the three founder maternal breeds have different allelic frequencies for the polymorphisms regulating in cis the expression of these genes.

The ACSM5.P polymorphism was segregating at low frequencies, being the ACSM5.P A allele frequency of 0.22 in BC1_LD, 0.09 in BC1_DU and 0.10 in BC1_PI. In the BC1_LD the ACSM5.P SNP was the most significant polymorphism associated with the differences in the mRNA level of ACSM5 and explained around the 58% of the phenotypic variance. As stated before, this result agree with the previous study performed with the same BC1_LD but in which the Sus scrofa 10.2 genome assembly was used. In BC1_PI, rs81475068, rs81278505 and ACSM5.P polymorphisms were located on SSC3 and spanning 0.17 Mb (2.39–2.56 Mb) and were the most significant associated SNPs with ACSM5 gene expression (p-value = 7.32 × 10⁻⁹), explaining approximately the 28% of the phenotypic variance. Contrarily, in BC1_DU rs81327383 was the most significantly associated SNPs (p-value = 2.02 × 10⁻¹²) with ACSM5 mRNA expression although the ACSM5.P polymorphism was also
significant ($p$-value = $3.44 \times 10^{-09}$) and explained a 26% of the gene expression variance. Hence, the lack of allele segregation or the presence of other proximal genetic variants could be involved in these gene expression changes.

In a previous work performed only in BC1_LD animals, the cis-eQTL for the muscle IGF2 gene expression was identified, but the IGF2g.3072 G > A polymorphism was not the most significant associated SNP, which was rs335265872 (called DIA50000846 and located at 6.20 Mb on SSC2 in the Sus Scrofa 10.2 genome assembly)\(^7\).

In the present work, the most significantly associated SNP in BC1_LD ($p$-value = $1.45 \times 10^{-10}$) was rs1322199 located at 3.68 Mb on SSC2 and explaining the 42% of the phenotypic variance. In addition, the IGF2g.3072 G > A polymorphism was significantly associated ($p$-value = $3.03 \times 10^{-10}$) and explained the 22% of the IGF2 mRNA variation. This result may be explained by the low number of homozygous AA animals, being 0.2 the allele frequency of the IGF2g.3072 A allele. On the other hand, the IGF2g.3072 G > A polymorphism was the most significantly associated SNP with IGF2 gene expression in BC1_DU (allele frequency of 0.29) and BC1_PI (allele frequency of 0.23), explaining in both cases a high proportion of the gene expression variance, 58% and 92% respectively. In BC1_DU other genomic regions seem to be also associated with the IGF2 gene expression differences, as the eQTL located in the 107.4–110.8 Mb genomic region of SSC4.

Two more cis-eQTLs were identified only in the BC1_LD population for MGLL and NCOA2 gene expression. The MGLL eQTL was previously described in the same backcross\(^5\). The SSC4 cis-eQTL for NCOA2 gene expression presented four significant associated SNPs, being the rs80803396 the SNP showing the strongest signal ($p$-value = $2.32 \times 10^{-09}$).

**Hotspots identified in trans-eQTLs regions.** All the trans-eQTLs intervals, eSNPs and annotated candidate genes are shown in the supplementary table S1, but only eQTL hotspots are discussed in detail (Supplementary Table S4). In BC1_DU, new trans-eQTLs were identified for ACA2 (SSC1), ACSM5 (SSC1, SSC2, SSC4, SSC6, SSC7, SSC11, and SSC18), ACSM2 (SSC1, SSC11, and SSC18), ETSL (SSC9), IGF2 (SSC4), LPIN1 (SSC4, SSC7, and SSC15), NCOA1 (SSC1), NCOA6 (SSC1), PDHX (SSC1), PPARA (SSC2, SSC15, and SSC17), PRRKA1 (SSC1), and PXMP3 (SSC4) genes in comparison to the 3BCs study. In BC1_LD additional trans-eQTLs were found for ACSM5 (SSC1, SSC6, SSC8, and SSC10), MLXIPL (SSC2, SSC9, and SSC13), CREG1 (SSC2), DGAT2 (SSC2, SSC7, and SSC9), HIF1AN (SSC2, SSC5, and SSC7), MGLL (SSC9 and SSC13), PKR31 (SSC16), PPPAR (SSC2), PPPARC1A (SSC2), and SCD (SSC2) genes in comparison to the 3BCs study. Two of the trans-eQTLs identified in the BC1_LD (FOS at SSC11 and MGLL at SSC13) were previously reported in the same BC1_LD animals with the Sus Scrofa 10.2 genome assembly. Finally, new trans-eQTLs in BC1_PI were detected for ACSM5 (SSC1, SSC8, SSC12, SSC14, and SSC16), ACS2 (SSC18), DGAT2 (SSC12), HIF1AN (SSC6 and SSC9), LXRA (SSC2), PPPAR (SSC7, SSC10, SSC14, SSC16, and SSC17), PPPARC1A (SSC2, SSC6, SSC7, and SSC17), and SCD (SSC17) genes (Fig. 4) in comparison to the 3BCs study.

We only observed three common trans-eQTL regions in the 3BCs study, suggesting the presence of different regulatory mechanisms or frequencies according to breed. Overall, the trans-eQTL regions manifested that the expression of the genes related to lipid metabolism is regulated in a complex way.

In addition, six hotspots regions, two in each backcross, regulating the expression of several genes were detected.

In BC1_LD animals a trans-eQTL hotspot located on SSC2 and spanning 8.7 Mb (119.9–128.7 Mb) was associated with the expression of seven genes: HIF1AN, CREG1, MLXIPL, DGAT2, PPPAR, PPPARC1A, and SCD. After gene annotation of this region no candidate trans-acting regulators modulating the expression of genes on the SSC2 hotspot were found. However, the transcription factor 7 (TCF7) gene was annotated in the CREG1 eQTL region because it was six Mb longer (119.9–136.2 Mb) than the others. TCF7 and its family member transcription factor 7 like 2 (TCF7L2) have been associated with diabetes in humans\(^4\). In addition, TCF7L2 has been described as an indirect regulator of PPARD during adipogenesis\(^4\). In addition, to evaluate potential functional interactions and the co-expression pattern of genes on the SSC2 hotspot, GeneMANIA and PCIT co-expression network analysis were done (Fig. 6). Interactions between DGAT2, PPPAR, PPPARC1A and SCD were found with GeneMANIA (Fig. 8A). In general, meaningful gene-gene interactions were shown by PCIT (Fig. 8B), reinforcing the presence of a common regulatory factor modulating the expression of SSC2 hotspot genes. However, lower correlations were observed for the CREG1 gene, suggesting the presence of an independent regulatory factor modulating its expression. This result is in accordance with the proposal of the TCF7 as a candidate gene of this region, although further validations are needed. Furthermore, HIF1AN presented negative and moderate correlations with DGAT2, MLXIPL, PPPAR, and SCD, suggesting an opposite regulatory effect for this gene. HIF1AN is involved in fatty acid β-oxidation\(^4\), while DGAT2, PPPAR, MLXIPL and SCD are related to de novo lipogenesis, triacylglycerol synthesis and adipogenesis\(^5\).

The strong correlation for SCD, PPPAR and DGAT2 identified in the gene co-expression network in 3BCs, and with MLXIPL and CREG1 have been found associated altogether with the trans-eQTL hotspot on SSC2 in the BC1_LD study but not in the other two backcrosses (BC1_DU and BC1_PI).

The region spanning 3.5 Mb on SSC7 (62.4–65.9 Mb) presented significant associations with the HIF1AN and DGAT2 gene expression. The nuclear factor of kappa light polypeptide gene enhancer in B cells enhancer alpha (NFKBIA) gene was mapped in this region. It is a transcription factor involved in immune response, but also plays a direct role in adipogenesis and fat accumulation\(^5\). NFKBIA was found differentially expressed in different development stages and muscles between Iberian and Iberian × Duroc pigs, suggesting that it is a molecular regulator of metabolism\(^4\). An experimental interaction between HIF1AN and NFKBIA was identified by GeneMANIA and String programs, but no information about DGAT2 interactions was found, so further validation will be needed to corroborate our results. Hence, we can suggest that NFKBIA is involved in muscle lipid metabolism, being an interesting candidate gene to explain the differences in the expression of two genes associated with the SSC7 hotspot in BC1_LD animals. None of these hotspots were identified in the previous study in which the same BC1_LD individuals were analyzed\(^7\). Taking all the eQTL results from BC1_LD together,
discrepancies between our results and the work of Puig-Oliveras et al. (2016)\textsuperscript{17} may be explained by the different genome assemblies used between both works, being Sus Scrofa 10.2 genome assembly in the previous work and Sus Scrofa 11.1 in the present one.

In the BC1_DU animals study two trans-eQTL hotspot regions were found on SSC1, spanning 6 Mb (180.6–203.6 Mb), and on SSC15, spanning 0.3 Mb (103.7–104 Mb). The SSC1 region showed significant associations with the expression of the ACAA2, CREG1, NCOA1, NCOA6, PDHX and PRKAA1 genes. The perilipin 2 (PLIN2) gene was mapped in this region but was only annotated as a candidate gene for ACAA2, NCOA1, NCOA6 and PDHX. PLIN2 was reported to be involved in the uptake and storage of FAs in human skeletal muscle\textsuperscript{53}. Studies in pigs described that a higher PLIN2 gene expression was associated with a higher IMF content in muscle\textsuperscript{54,55}.

In order to deepen in the study of the genes regulated by the same eQTL on SSC1, the PCIT algorithm was used to build a co-expression network. Moderate to high positive correlations, from 0.15 to 0.78, were observed among the genes regulated by the same eQTL (Fig. 9). Lower correlations were observed for ACAA2, a gene encoding an enzyme that catalyzes the last step in mitochondrial fatty acid \(\beta\)-oxidation\textsuperscript{56}, suggesting the presence of another genetic factor regulating its expression. In addition, moderate correlations were found for the rest of the hotspot genes, mainly related to transcriptional regulation and control.

Notably, the second group of genes identified in the gene co-expression network in 3BCs, which showed strong correlations for PRKAA1, PDHX, NCOA1 and NCOA2 among others, coincides with the previously observed SSC1 trans-eQTL hotspot in BC1_DU study, but not in the other two backcrosses (BC1_LD and BC1_PI). LPIN1 and PPARA genes were significantly associated with the SSC15 hotspot region and showed a moderate correlation value (LPIN1-PPARA, \(r = 0.59\) p-value = \(4.97 \times 10^{-13}\)). In this region was mapped a key mitochondrial
enzyme for fatty acid oxidation, AOX1 gene. It has been reported to be associated with FA oxidation in mice adipocytes\(^\text{27}\) and meat quality traits and with muscle development in cattle\(^\text{28}\).

Regarding BC1_PI population, two \textit{trans}-eQTL hotspots regions on SSC7 and SSC17 were observed. The first region, spanning 8.1 Mb on SSC7 (100.1–108.2 Mb), showed a significant association with PPARG and \textit{PPARGC1A} gene expression. \textit{DIO2} gene was mapped in the SSC7 \textit{trans}-eQTL region as a potential candidate gene for lipid metabolism. It has been selected as a muscle candidate gene in an obesity resistance study since it presented differences between lean and fat mouse lines\(^\text{59}\). \textit{DIO2} converts prohormone thyroxine (T4) to the active hormone triiodothyronine (T3), which binds to thyroid hormone receptors (TR). TR heterodimerize with RXR and can compete with PPAR for that binding site affecting gene control and regulation\(^\text{60}\). Hence, \textit{DIO2} may be an indirect regulator of SSC7 hotspot genes.

The second region located on SSC17 and spanning 12.6 Mb (29.2–41.8 Mb), presented a significant association with PPARG, \textit{PPARGC1A} and \textit{SCD}. Three genes were mapped for the SSC17 hotspot: \textit{RBL1}, \textit{FOXA2} and \textit{E2F1}. \textit{RBL1} gene has been associated with the whole body fat metabolism and determines the oxidative state of muscle in mice\(^\text{41}\), \textit{FOXA2} has been described as a transcription factor of several genes involved in the insulin pathway in liver\(^\text{62}\), but no studies in muscle tissue were found. It was reported that \textit{E2F1} is required for \textit{in vivo} skeletal muscle regeneration in mouse\(^\text{62}\) and showed high gene expression levels in Pietrain pigs with high muscle content\(^\text{63}\). Interactions were found between the genes associated with the hotspot (PPARG, \textit{PPARGC1A} and \textit{SCD}) and between the \textit{E2F1}, \textit{RBL1} and \textit{PPARG} genes using GeneMANIA and String. \textit{RBL1} and \textit{E2F1} were selected as promising candidate genes for lipid metabolism in pigs, but further validations are needed to assess the effect of \textit{FOXA2} in muscle tissue.

**Conclusions**

In the present study, gene expression of candidate genes for fatty acid composition in muscle showed sex-dimorphism and breed effects, and gene co-expression in different lipid metabolism pathways was identified. The eGWAS revealed two \textit{cis}-eQTL and ten \textit{trans}-eQTL regions associated with the muscle expression of ACSM5, ACS2, \textit{ATF3}, \textit{DGAT2}, \textit{FOS} and \textit{IGF2} genes. Both \textit{IGF2}:g.3072 G > A and \textit{ACSM5}:P polymorphisms were the most significant SNPs associated with the \textit{IGF2} and \textit{ACSM5} gene expression levels, respectively, in different pig genetic types. Two \textit{trans}-eQTL hotspot regions per backcross regulating the expression of up to seven genes were identified, and a list of candidate genes involved in the regulation of these eQTL regions was reported. Overall, our results increase the knowledge of the genetic basis of gene expression regulation in muscle lipid metabolism.

**Material and Methods**

**Animal material.** A total of 355 pigs (called 3BCs) belonging to three different experimental backcrosses were studied: 114 BC1_LD (25% Iberian and 75% Landrace), 122 BC1_DU (25% Iberian and 75% Duroc), and 119 BC1_PI (25% Iberian and 75% Pietrain)\(^\text{35}\). All animals were maintained under the same intensive conditions and fed \textit{ad libitum} with a cereal-based commercial diet and slaughtered in a commercial abattoir following institutional and national guidelines for the Good Experimental Practices and approved by the Ethical Committee of the Institution (IRTA – Institut de Recerca i Tecnologia Agroalimentàries). In addition, animal care and procedures were carried out according to the Spanish Policy for Animal Protection RD1201/05 and the European Union Directive 86/609 about the protection of animals used in experimentation.

LD samples were collected at slaughterhouse in liquid nitrogen and stored at –80 °C until analysis. Genomic DNA was extracted from diaphragm tissue using the phenol-chloroform method\(^\text{65}\).

**Genotyping.** Animals from BC1_LD and BC1_PI were genotyped using the Porcine SNP60K BeadChip (Illumina, San Diego, USA) and BC1_DU animals were genotyped using the GeneChip Porcine Genome Array (Affymetrix). Only SNPs that mapped against the Sscrofa11.1 assembly and were common to both arrays were selected\(^\text{35}\). Markers that showed a minor allele frequency (MAF) lower than 5% and SNPs with more than 5% of missing genotypes were removed with Plink software\(^\text{66}\). Moreover, based on the information in the prior BC1_LD study\(^\text{37}\), two additional SNPs were genotyped: \textit{ACSM5} (rs331702081) and \textit{IGF2} (IGF2:g.3072 G > A), in the BC1_DU and BC1_PI populations, following the previously described protocols\(^\text{17,22}\). Finally, a total of 38.426 SNPs distributed along all chromosomes, including rs331702081 and IGF2:g.3072 G > A polymorphisms, were used for association studies.

**Gene expression.** Total RNA was obtained from the LD muscle of 355 animals using the RiboPure kit (Ambion), following the fabricator’s instructions. RNA quantification and purity was performed with a NanoDrop ND-1000 spectrophotometer (NanoDrop products) and RNA integrity was assessed by Agilent Bioanalyzer-2100 (Agilent Technologies). The RNA was reverse-transcribed into cDNA using the High-Capacity cDNA Reverse Transcription kit (Applied Biosystems), following the manufacturer’s recommendations. Gene expression was analyzed in 48 genes, of which 45 were target genes and \textit{ACTB}, \textit{HPRT1} and \textit{TBP} were the candidate reference genes, by quantitative real-time PCR (qPCR). Selection of target genes related to lipid metabolism as well as primer design details and sequences was described in Puig-Oliveras et al. (2016)\(^\text{17}\). Gene expression quantification was performed in a 48.48 Microfluidic Dynamic Array IFC Chip (Fluidigm) in a BioMark System succeeding a previously described protocol\(^\text{67}\). Gene expression data was picked up using Fluidigm Real-Time PCR analysis software 3.0.2 (Fluidigm) and analyses were done with DAG Expression software 1.0.4.11\(^\text{68}\), applying the relative standard method curve. \textit{ACTB} and \textit{TBP} were used as the most stable reference genes, while \textit{HPRT1} was discarded, in order to normalize the expression levels of target genes. The normalized quantity (NQ)\(^\text{68}\) values of each sample and assay were used to compare the expression data among animals. Normalization of data was checked through Shapiro-Wilk test in R\(^\text{69}\), and log, transformation of the NQ value was applied if necessary. Sex and breed effects were tested by using a linear model (lm) in R\(^\text{69}\).
Genome-wide association analysis for gene expression. To carry out the genomic association studies between 45 gene expression values and common SNPs genotypes (eGWAS), a previously reported linear mixed model using the GEMMA software was applied.

Gene annotation. Significant associated SNPs were mapped in the Sscrofa11.1 assembly and were annotated with the Ensembl Genes 91 Database using VEP software. BioMart software was used to annotate genomic eQTL intervals considering ±1 Mb around the candidate chromosomal regions. In the three studied BCs study only eQTL intervals containing 2 or more SNPs were annotated, whereas in the individual backcross GWAS annotation was done for eQTL intervals containing 3 or more SNPs.

The identified SNPs were classified depending on their location, as cis if the SNPs were located within 1 Mb of the analyzed gene and as trans if the SNPs were located elsewhere in the genome. The number of significant SNPs belonging to the same interval was considered among associated SNPs less than 10 Mb apart.

Co-expression and functional analysis. The PCIT algorithm was used to calculate weighted gene co-expression networks, through the implementation of first-order partial correlations coefficients combined with information theory approach, in order to identify principal interactions between genes. Only the significant interactions between genes were considered for further steps. Networks were represented with Cytoscape plug-in.

Ingenuity Pathway Analysis software (IPA; Ingenuity Systems) and the Core Analysis function was used to perform functional analysis of genes mapped in the different intervals and for data interpretation in the context of biological processes, pathways and networks. In addition, the iRegulon v1.3. Cytoscape plug-in was used to identify transcription factor (TF) binding sites in silico. ClueGO plug-in was used to integrate and cluster the genes regarding their Gene Ontology and KEGG pathway. Finally, GeneMANIA and String were used to evaluate the functional interaction and networks among genes proteins, respectively.

References
1. Wood, J. D. et al. Fat deposition, fatty acid composition and meat quality: A review. Meat Sci. 78, 343–358 (2008).
2. Simopoulos, A. P. The importance of the ratio of omega-6/omega-3 essential fatty acids. Biomed. Pharmacother. 56, 365–379 (2002).
3. Lopez-Bote, C. J. Sustained utilization of the Iberian pig breed. Meat Sci. 49, (1998).
4. Ventanas, S. & Al, E. Quality traits in muscle biceps femoris and back-fat from purebred Iberian × Duroc crossbred pigs. Meat Sci. 73, 651–659 (2006).
5. Wood, J. D. et al. Effects of fatty acids on meat quality: A review. Meat Sci. 66, 21–32 (2004).
6. Casellas, J. et al. Bayes factor analyses of heritability for serum and muscle lipid traits in Duroc pigs. J. Anim. Sci. 88, 2246–2254 (2010).
7. Ntawubizi, M. et al. Genetic parameters for intramuscular fatty acid composition and metabolism in pigs. J. Anim. Sci. 88, 1286–1294 (2014).
8. Selleri, P. & Monin, G. Genetics of Pig Meat Quality: a Review. J. Muscle Foods 5, 187–219 (1994).
9. Ramayo-Caldas, Y. et al. Genome-wide association study for intramuscular fatty acid composition in an Iberian × Landrace cross. J. Anim. Sci. 90, 2883–2893 (2012).
10. Muñoz, M. et al. Genome-wide analysis of porcine backfat and intramuscular fat fatty acid composition using high-density genotyping and expression data. BMC Genomics 14, 845 (2013).
11. Revilla, M. et al. New insight into the SSC8 genetic determination of fatty acid composition in pigs. Genet. Sel. Evol. 46, 1–10 (2014).
12. Ayuso, M. et al. Comparative analysis of muscle transcriptome between pig genotypes identifies genes and regulatory mechanisms associated to growth, Fatness and metabolism. PLoS One 10, 1–33 (2015).
13. Corominas, J. et al. Polymorphism in the ELOVL6 Gene Is Associated with a Major QTL Effect on Fatty Acid Composition in Pigs. PLoS One 8, 1–12 (2013).
14. Yang, B. et al. Genone-Wide Association Analyses for Fatty Acid Composition in Porcine Muscle and Abdominal Fat Tissues. 8 (2013).
15. Hausman, G. J. et al. Board-invited review: The biology and regulation of preadipocytes and adipocytes in meat animals. J. Anim. Sci. 87, 1218–1246 (2009).
16. Puig-Oliveras, A. et al. Differences in muscle transcriptome among pigs phenotypically extreme for fatty acid composition. PLoS One 9 (2014).
17. Puig-Oliveras, A. et al. Expression-based GWAS identifies variants, gene interactions and key regulators affecting intramuscular fatty acid content and composition in porcine meat. Sci. Rep. 6, 31803 (2016).
18. Ballester, M. et al. Integration of liver gene co-expression networks and eGWAS analyses highlighted candidate regulators implicated in lipid metabolism in pigs. Sci. Rep. 7, 46539 (2017).
19. Revilla, M. et al. Expression analysis of candidate genes for fatty acid composition in adipose tissue and identification of regulatory regions. Sci. Rep. 8, 1–13 (2018).
20. Zhang, Y. et al. Transcriptional profiling of human liver identifies sex-biased genes associated with polygenic dyslipidemia and coronary artery disease. PLoS One 6 (2011).
21. Liu, D. et al. Skeletal muscle gene expression in response to resistance exercise: sex specific regulation. BMC Genomics 11, 659 (2010).
22. Van Laere, A.-S. et al. A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. Nature 425, 832–836 (2003).
23. Varlamov, O., Bethea, C. L. & Roberts, C. T. Sex-specific differences in lipid and glucose metabolism. Front. Endocrinol. (Lausanne). 5, 1–7 (2014).
24. Zhang, X. & Lerman, L. O. Investigating the Metabolic Syndrome. Toxicol. Pathol. 44, 358–366 (2016).
25. Burli, L., Thoresen, G. H. & Berge, R. K. The Role of PPAR alpha Activation in Liver and Muscle. 2010, (2010).
26. Vega, R. B. & Huss, J. M. The Coactivator PGC-1 Cooperates with Peroxisome Proliferator-Activated Receptor _ in Transcriptional Control of Nuclear Genes Encoding Mitochondrial Fatty Acid Oxidation Enzymes. 20, 1868–1876 (2000).
27. Eberlé, D., Hegarty, B., Bossard, P., Ferré, P. & Foufelle, F. SREBP transcription factors: Master regulators of lipid homeostasis. Biochimie 86, 839–848 (2004).
28. Lehrke, M. & Lazar, M. A. The many faces of PPARs. Cell 123, 993–999 (2005).
29. Krook, A. Can the liver X receptor work its magic in skeletal muscle too? Diabetologia 49, 819–821 (2006).
30. Benítez, R. et al. Modulatory effects of breed, feeding status, and diet on adipogenic, lipogenic, and lipolytic gene expression in growing ibeiran and duroc pigs. Int. J. Mol. Sci. 19, 1–20 (2018).
31. Cánovas, A., Quintanilla, R., Arills, M. & Pena, R. N. Muscle transcriptomic profiles in pigs with divergent phenotypes for fatness traits. BMC Genomics 11, (2010).
32. Reverter, A. & Chan, E. K. F. Combining partial correlation and an information theory approach to the reversed engineering of gene co-expression networks. Bioinformatics 24, 2491–2497 (2008).
33. Aslan, O. et al. Variation in the IGF2 gene promoter region is associated with intramuscular fat content in porcine skeletal muscle. Mol. Biol. Rep. 39, 4101–4110 (2012).
34. Markljung, E. et al. ZBED6, a novel transcription factor derived from a domesticated DNA transposon regulates IGF2 expression and muscle growth. PLoS Biol. 7, (2009).
35. Criado-Menas, L. et al. Analysis of porcine IGF2 gene expression in adipose tissue and its effect on fatty acid composition. PLoS One 14, e0220708 (2019).
36. Chang, Y. H., Huang, C. N. & Shiau, M. Y. Association of IL-4 receptor gene polymorphisms with high density lipoprotein cholesterol. Cytokine 59, 309–312 (2012).
37. Barilé, A., Schmoll, D. & Unterman, T. G. FOXO proteins in insulin action and metabolism. Trends Endocrinol. Metab. 16, 183–189 (2005).
38. Ayuso, M. et al. Developmental stage, muscle and genetic type modify muscle transcriptome in pigs: Effects on gene expression and regulatory factors involved in growth and metabolism. PLoS One 11, 1–33 (2016).
39. Judge, S. M. et al. Genome-wide identification of FOXO-dependent gene networks in skeletal muscle during C26 cancer cachexia. 1–17 (2014).
40. Hu, X. et al. Molecular cloning, expression pattern analysis of porcine Rb1 gene and its regulatory roles during primary deridifferentiated fat cells adipogenic differentiation. Gen. Comp. Endocrinol. 214, 77–86 (2015).
41. Huang, D. et al. The possible role of complete loss of myostatin in limiting excessive proliferation of muscle cells (C2C12) via activation of microRNAs. Int. J. Mol. Sci. 20, 643 (2019).
42. Zhang, B. et al. The association of NR1H3 gene with lipid deposition in the pig. Lipids Health Dis. 15, 4–11 (2016).
43. Elhein, S. C., Das, S. K., Hallman, D. M., Hanis, C. L. & Hasstedt, S. J. Gene-wide linkage and admixture mapping of type 2 diabetes in African American families from the American diabetes association GENNID (Genetics of NIDDM) study cohort. Diabetes 58, 268–274 (2009).
44. Cristancho, A. G. et al. Repressor factor 7-like 1 promotes adipogenic competency in precursor cells. Proc. Natl. Acad. Sci. 108, 16271–16276 (2011).
45. Krishnan, J. et al. Dietary obesity-associated hit5α activation in adipocytes restricts fatty acid oxidation and energy expenditure via suppression of the Sirt2-NAD+ system. Genes Dev. 26, 259–270 (2012).
46. Mylonis, I., Simos, G. & Paraskeva, E. Hypoxia-Inducible Factors and the Regulation of Lipid Metabolism. Cells 8, 214 (2019).
47. Knutti, D. & Kralli, A. PGC-1, a versatile coactivator. Trends Endocrinol. Metab. 12, 360–365 (2001).
48. Puig-Oríol, A. et al. A co-association network analysis of the genetic determination of pig conformation, growth and fatness. PLoS One 9, 1–20 (2014).
49. Ahmadian, M. et al. Pparα signaling and metabolism: The good, the bad and the future. Nat. Med. 19, 557–566 (2013).
50. Jiang, Z. et al. Significant associations of stearoyl-CoA desaturase (SCD1) gene with fat deposition and composition in skeletal muscle. Int. J. Biol. Sci. 4, 345–351 (2008).
51. Yu, K. et al. Activating transcription factor 4 regulates adipocyte differentiation via altering the coordinate expression of CCATT/enhancer binding protein α and peroxisome proliferator-activated receptor γ. FEBS J. 281, 2399–2409 (2014).
52. Ren, W. et al. CCAAT/enhancer-binding protein α is a crucial regulator of human fat mass and obesity associated gene transcription and expression. Biomed Res. Int. 2014, (2014).
53. Bickel, P. E., Tansey, J. T. & Welte, M. A. PAT proteins, an ancient family of lipid droplet proteins that regulate cellular lipid stores. Biochim. Biophys. Acta - Mol. Cell Biol. Lipids 1791, 419–440 (2009).
54. Davoli, R. et al. New SNP of the porcine Perilipin 2 (PLIN2) gene, association with carcass traits and expression analysis in skeletal muscle. Mol. Biol. Rep. 38, 1575–1583 (2011).
55. Gandolfi, G. et al. Perilipin 1 and perilipin 2 protein localization and gene expression study in skeletal muscle of European crossbreed pigs with different intramuscular fat contents. Meat Sci. 88, 631–637 (2011).
56. Bartlett, K. & Eaton, S. Mitochondrial-β-oxidation. Eur. J. Biochem. 271, 462–469 (2004).
57. Sun, L., Liu, Z., Cao, W., Zhang, Z. & Sun, C. FABP4 reverses the regulation of leptin on mitochondrial fatty acid oxidation in mice adipocytes. Sci. Rep. 5, 1–12 (2015).
58. Guillocheau, G. M. et al. Survey of allele specific expression in bovine muscle. Sci. Rep. 9, 1–11 (2019).
59. Simončič, M. et al. Obesity resistant mechanisms in the Lean polygenic mouse model as indicated by liver transcriptome and expression of selected genes in skeletal muscle. BMC Genomics 12, 1–12 (2011).
60. Sinha, R. A., Singh, B. K. & Yen, P. M. Thyroid hormone regulation of hepatic lipid and carbohydrate metabolism. Trends Endocrinol. Metab. 25, 538–545 (2014).
61. Scimé, A. et al. Oxidative status of muscle is determined by p107 regulation of PGC-1α. J. Cell Biol. 190, 651–662 (2010).
62. K.A., L. et al. Foxa2 regulates multiple pathways of insulin secretion. J. Clin. Invest. 114, 512–520 (2004).
63. Yan, Z. et al. Highly coordinated gene regulation in mouse skeletal muscle regeneration. J. Biol. Chem. 278, 8826–8836 (2003).
64. Fan, H. The Hydrogoniometer and Assessment of Gleno-Humeral Joint Motion. Gene 496, 8–14 (2011).
65. Sambrook, J., Fritsch, E. & Maniatis, T. Molecular Cloning: A Laboratory Manual 2nd edn. Cold Spring Harbor, N.Y. 0167–708 (2001).
66. Purcell, S. et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am. J. Hum. Genet. 81, 559–575 (2007).
67. Ramayo-Calzas, Y. et al. From SNP co-association to RNA co-expression: Novel insights into gene networks for intramuscular fatty acid composition in porcine. BMC Genomics 15, 232 (2014).
68. Ballester, M., Cordón, R. & Folch, J. M. DAG expression: High-throughput gene expression analysis of real-time PCR data using standard curves for relative quantification. PLoS One 8, 8–12 (2013).
69. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (2018).
70. Zhou, X. & Stephens, M. Genome-wide efficient mixed-model analysis for association studies. Nat. Genet. 44, 821–4 (2012).
71. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. 57, 289–300 (1995).
72. McLaren, W. et al. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. Bioinformatics 26, 3069–2070 (2010).
73. Smalley, D. et al. The BioMart community portal: An innovative alternative to large, centralized data repositories. Nucleic Acids Res. 43, W589–W598 (2015).
74. Watson-Haigh, N. S., Kodarmideon, H. N. & Reverter, A. PCIT: An R package for weighted gene co-expression networks based on partial correlation and information theory approaches. Bioinformatics 26, 411–413 (2009).
Acknowledgements
We want to thank all of the members of the INIA, IRTA, and UAB institutions who contributed to the generation of the animal material used in the current study. This work was supported by the Spanish Ministerio de Economía y Competitividad (MINECO) and the Fondo Europeo de Desarrollo Regional (FEDER) with project references: AGL2014–56369-C2 and AGL2017–82641-R. L. Criado-Mesas was funded with a FPI grant from the AGL2014–56369-C2 project. M. Ballester was financially supported by a “Ramón y Cajal” contract (RYC-2013–12573) from the Spanish Ministerio de Economía y Competitividad. D. Crespo-Piazuelo was funded by a “Formació i Contractació de Personal Investigador Novell” (FI-DGR) Ph.D grant from the Generalitat de Catalunya (ECO/1788/2014). We acknowledge the support of the Spanish Ministerio de Economía y Competitividad for the “Severo Ochoa Programme for Centres of Excellence in R&D” 2016–2019 (SEV-2015–0533) to the Centre for Research in Agricultural Genomics and the CERCA Programme / Generalitat de Catalunya.

Author contributions
J.M.F. and A.I.F. conceived and designed the experiments; J.M.F. was the principal investigator of the project; A.I.F. and J.M.F. collected the animal samples; L.C.M. and D.C.P. performed the pig genomic DNA extraction; A.C. genotyped the samples; L.C.M., M.B. and A.C. designed and performed the gene expression studies; L.C.M. and J.M.F. performed the genome-wide association studies; L.C.M., M.B. and J.M.F. wrote the paper. All authors read and approved the final manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-67015-4.

Correspondence and requests for materials should be addressed to L.C.-M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020