animal

The effect of divergent selection for intramuscular fat on the domestic rabbit genome

B. S. Sosa-Madrid[1] ⓘ, L. Varona[2] ⓘ, A. Blasco[1] ⓘ, P. Hernández[1] ⓘ, C. Casto-Rebollo[1] ⓘ and N. Ibáñez-Escriche[1]† ⓘ

[1]*Institute for Animal Science and Technology, Universitat Politècnica de València, 46022 Valencia, Spain;* [2]*Unidad de Genética Cuantitativa y Mejora Animal, Universidad de Zaragoza, Instituto Agroalimentario de Aragón (IA2), 50013 Zaragoza, Spain*

*An experiment of divergent selection for intramuscular fat was carried out at Universitat Politècnica de València. The high response of selection in intramuscular fat content, after nine generations of selection, and a multidimensional scaling analysis showed a high degree of genomic differentiation between the two divergent populations. Therefore, local genomic differences could link genomic regions, encompassing selective sweeps, to the trait used as selection criterion. In this sense, the aim of this study was to identify genomic regions related to intramuscular fat through three methods for detection of selection signatures and to generate a list of candidate genes. The methods implemented in this study were Wright's fixation index, cross population composite likelihood ratio and cross population – extended haplotype homozygosity. Genomic data came from the 9th generation of the two populations divergently selected, 237 from Low line and 240 from High line. A high single nucleotide polymorphism (**SNP**) density array, Affymetrix Axiom OrcunSNP Array (around 200k SNPs), was used for genotyping samples. Several genomic regions distributed along rabbit chromosomes (OCU) were identified as signatures of selection (SNPs having a value above cut-off of 1%) within each method. In contrast, 8 genomic regions, harbouring 80 SNPs (OCU1, OCU3, OCU6, OCU7, OCU16 and OCU17), were identified by at least 2 methods and none by the 3 methods. In general, our results suggest that intramuscular fat selection influenced multiple genomic regions which can be a consequence of either only selection effect or the combined effect of selection and genetic drift. In addition, 73 genes were retrieved from the 8 selection signatures. After functional and enrichment analyses, the main genes into the selection signatures linked to energy, fatty acids, carbohydrates and lipid metabolic processes were ACER2, PLIN2, DENND4C, RPS6, RRAGA (OCU1), ST8SIA6, VIM (OCU16), RORA, GANC and PLA2G4B (OCU17). This genomic scan is the first study using rabbits from a divergent selection experiment. Our results pointed out a large polygenic component of the intramuscular fat content. Besides, promising positional candidate genes would be analysed in further studies in order to bear out their contributions to this trait and their feasible implications for rabbit breeding programmes.*

Implications

Intramuscular fat content is an essential factor in meat quality because it affects nutritional, sensory and technological properties of meat, such as tenderness, flavour and juiciness of meat. In this study, we applied method of selection signatures to identify genomic regions modified by a divergent selection experiment for intramuscular fat in rabbits. Results revealed several selection signatures across the rabbit genome with genes linked to lipid metabolism. These findings will help to increase our understanding of intramuscular fat genomic regulation and could be used to apply in genomic evaluation programmes for intramuscular fat.

Introduction

Selection and mutation trigger shifts in the genome architecture of traits, gathering the history of particular populations at a genomic level (Oleksyk *et al.*, 2010). Genomic regions harbouring genes influenced by a selective process can be detected by the methods for the identification of selection signatures. These methods can be categorised in four groups depending on the assumptions behind the null hypothesis of

† E-mail: noeibes@dca.upv.es

absence of selection (Qanbari and Simianer, 2014): (i) based on classical analyses of genetic variability (e.g. F_{st} – Wright's fixation index, π – nucleotide diversity), (ii) reduction of local variation in genomic regions (e.g. ROH – run of homozygosity), (iii) modification of allelic frequency spectrum (e.g. TD – Tajima's D, Fay and Wu H test, CLR – composite likelihood ratio) (iv) and linkage disequilibrium decay (e.g. |iHS| – integrated haplotype score, EHH – extended haplotype homozygosity, varLD – variation of linkage disequilibrium); see reviews by Oleksyk et al. (2010) and Qanbari and Simianer (2014). The choice of methods depends on the type of selective events, timescale of selective events, the density of the genotyping data and the number of populations available for each particular study. A combination of methods for selection signatures can provide a clearer evidence of the genomic regions considered as selection signatures (Utsunomiya et al., 2013).

The identification of genomic regions containing genes affected by natural and artificial selection can be a difficult task, because selection is a complex phenomenon involving a potentially large number of traits (Mallick et al., 2009). Conversely, populations derived from divergent selection experiments for one trait provide a valuable biological material for detecting those signatures, as the genetic divergence between them is linked to one particular trait (Qanbari and Simianer, 2014). In this sense, several studies of divergent selection were used to detect genomic regions associated with selection events in poultry: for BW (Johansson et al., 2010), feather pecking behaviour (Grams et al., 2015) and antibody response (Lillie et al., 2017), and pigs: for intramuscular fat (Kim et al., 2015) and feed efficiency (Mauch et al., 2018).

In rabbits, an experiment of divergent selection for intramuscular fat was carried out at the Universitat Politècnica de València attaining a high selection response (Martínez-Álvaro et al., 2016). The genomic information from these two rabbit lines establishes an outstanding material to disentangle the genetic architecture of intramuscular fat content through genome-wide scan studies for the detection of selection signatures.

The aim of this study was to identify genomic regions using three methods to detect selection signatures that exploit genomic information from divergent populations and based on distinct hypotheses. The first is F_{st} (Qanbari and Simianer, 2014), based on conventional genetic differentiation, the second is the cross population – composite likelihood ratio (XP-CLR; Chen et al., 2010), which analyses the modifications on the allele frequency spectrum and the last one is the cross population – extended haplotype homozygosity (XP-EHH; Sabeti et al., 2007), focused on the differences on the extension of linkage disequilibrium between populations. The final objective was to generate a list of potential candidate genes associated with intramuscular fat content.

Material and methods

Animals, genotyping data and quality control
The two rabbit lines divergently selected for intramuscular fat came from a synthetic line (base population) reared at Universitat Politècnica de València (Zomeño et al., 2013). Each line was composed of 8 to 10 sires and 40 to 60 does per generation. Further details of the divergent selection experiment for intramuscular fat are presented in Martínez-Álvaro et al. (2016). After nine generations of selection, the response was 3.1 phenotypic SDs (41% of the mean from the base population), estimated as the phenotypic difference between the two divergently selected lines: GA9 – High line at 9th generation and GB9 – Low line at 9th generation (Sosa-Madrid et al., 2020). In addition, this selection response was corroborated by estimating the genetic means of each line at 9th generation under Bayesian inference. The model was the same as one described by Martínez-Álvaro et al. (2016), including the fixed effects (line, month, sex and parity order), a common litter random effect and a residual random effect.

Muscle samples were collected for genotyping. A total of 480 individual rabbits (240 from each line) at 9th generation were genotyped with the Affymetrix Axiom OrcunSNP Array, around 200k single nucleotide polymorphism (SNP). In addition, we genotyped 96 ancestors at 8th generation (10 sires and 38 dams by each line): GA8 - High line at 8th generation and GB8 - Low line at 8th generation. Quality control of the SNP data was performed using 'Axiom Analysis Suite v. 3.0.1.4' by using the following criteria: (i) individual call rate > 0.97, (ii) SNP call rate > 0.95, (iii) SNP minor allele frequency (MAF) > 0.05, and (iv) only autosomal SNPs with known positions were used. An exploratory analysis of sensitivity to MAF threshold on the results of selection signatures was carried out using 0.001, 0.01 and 0.05 MAF values. The number of SNPs after quality control and the results of selection signatures showed negligible changes between the MAF thresholds. Hence, a MAF threshold of 0.05 was chosen in order to control the rate of false-positive selection signatures and the effect of genotyping errors on the results. After filtering, we imputed the missing genotypes and inferred haplotype phases using population and genealogical information with FImpute (Sargolzaei et al., 2014). The imputation was carried out because the genomic data had a high quality, for example, all individuals presented a missing genotyping rate less than 0.02. A total of 5144 SNPs were imputed, showing an accuracy greater than 0.98 (Pearson's correlation). The final data set consisted of 89 968 genotyped SNPs from 477 rabbits (240 from the High and 237 from the Low lines, respectively).

Divergence between lines
At first, a multidimensional scaling (MDS) analysis with all genomic data was carried out to corroborate the divergence between lines. The command cmdscale() from R package stats was implemented for the MDS analysis. In addition, linkage disequilibrium was computed as Pearson's squared correlation coefficient (R^2) across the rabbit genome using PLINK (Purcell et al., 2007).

Detection of selection signatures
The data were analysed using the following methods for the detection of selection signatures, taking advantage of selection in the two divergent lines after nine generations of selection:

Fixation index. This F_{st} parameter was computed for each SNP as:

$$F_{st} = \frac{H_T - H_S}{H_T} = \frac{2\overline{p}\,(1-\overline{p}) - \frac{\sum 2\,p_i\,(1-p_i)\,n_i}{\sum n_i}}{2\overline{p}\,(1-\overline{p})}$$

where H_S is the average expected heterozygosity of rabbit lines, H_T is the expected heterozygosity of the total population, $\overline{p}$ is the average allele frequency across lines, p_i is the allele frequency for each line i, and n_i is the number of rabbits (individuals) in each of the i lines. The F_{st} values were clustered over sliding windows of predefined size (250, 500 and 1000 kb) surrounding every SNP. The F_{st} normalisation was carried out in order to correct the F_{st} values due to the heterogeneous distribution of SNPs, after quality control, along the rabbit genome. The equation used was

$$\text{normalised\_}F_{st} = \frac{\overline{X}_j - \mu}{s/\sqrt{n_j}}$$

This is based on the number of SNP within window j: n_j, the SD using all data: S, the deviation from F_{st} average of a given window j: $\overline{X}_j$, and the F_{st} total mean: μ (Beissinger *et al.*, 2015).

Cross population – composite likelihood ratio test. The XP-CLR method computes the likelihood ratio of selection signatures by comparing the spatial distribution of allele frequencies in an observed window to the frequency spectrum of the whole genome between two populations (Chen *et al.*, 2010). The High line was used as the objective population and the Low line was used as the reference population. In this analysis, *XP-CLR* software available at http://genetics.med.harvard.edu/reich/Reich_Lab/Software.html was employed to compute the XP-CLR. After several exploratory analyses on XP-CLR score and its parameters, we defined a grid size of 2000 base pairs, sliding window size of three levels (250, 500 and 1000 kb), maximum number of SNPs within a window 200 and a correlation value between two adjacent SNPs weighted with a cut-off of $R^2 > 0.95$ (author's recommendation). A shortcoming of CLR-based methods is that the correlation of marginal likelihood terms in the composite likelihood function is ignored. Thus, these methods overestimate the amount of information that is available in the data, which can prompt false-positive signals of selection. To control for this issue, the XP-CLR method assigns weights to each marginal likelihood function in proportion to their statistical independence from all of the others (low or null correlations amongst the functions of marginal likelihoods). When $R^2 > 0.95$, CLR scores for two SNPs are down-weighted. After performing the analyses for every level of sliding window, the XP-CLR score for every SNP was chosen as the value of the nearest grid to each SNP. More details of the parameters of this method are described in Chen *et al.* (2010).

Cross population – extended haplotype homozygosity test. The EHH profiles are defined as the probability that two randomly chosen haplotypes are identical by descent for the entire interval from the core region to a given point. The XP-EHH test compares the integrated EHH profiles between two populations around the same SNP, detecting ongoing selection or nearly fixed sites (overrepresented haplotypes) unveiling selection in one population (Sabeti *et al.*, 2007). As in the XP-CLR score, we defined the High line as the objective population and Low line as the reference population. First, we calculated the integrated haplotype score (IHH) for both lines. Then, the statistic was calculated at each SNP position as:

$$\text{XP} - \text{EHH}_{\text{high}-\text{low}} = \ln(\text{IHH}_{\text{high}}/\text{IHH}_{\text{low}})$$

in which XP-EHH$_{\text{high-low}}$ is the XP-EHH between the High and Low intramuscular fat lines, IHH$_{\text{high}}$ is the integrated haplotype score of the High line and IHH$_{\text{low}}$ is the integrated haplotype score of the Low line (Sabeti *et al.*, 2007; Qanbari and Simianer, 2014). The maximum of extended haplotype was restricted in 250, 500 and 1000 kb in order to compare with the other methods and to evaluate the sensitivity of the methods to the window size. The computation of XP-EHH score was carried out using *selscan* software (Szpiech and Hernandez, 2014), and finally, normalisation of these data was performed as recombination rates vary widely across the rabbit genome within and between populations. This normalisation was carried out setting all such log-ratios have zero mean and unit variance. Each log of XP-EHH values was subtracted the mean and divided by SD using all log of EHH values. The EHH statistic can be interpreted as a measure of selection solely after appropriate normalisation for genome-wide difference in haplotype length between populations as the distribution of recombination sites and the recombination rate are not steady (Sabeti *et al.*, 2007).

Enrichment analysis of functional annotation, and gene ontology terms

In this study, we used a cut-off of 1% (the 99th percentile of all values) for every method to retain interesting signatures of selection. In order to determine the genomic regions of interest for searching genes and functional annotations, we used the physical position of the SNPs exceeding the cut-off (± 250 kb) in at least two methods. This distance criterion was chosen based on the results of the relationship between window size and the three methods of selection signatures used in this study. We considered that criterion for searching genes under the assumption that an outstanding signature of selection must be detected in at least two methods, showing better evidence in this way and taking into account that the methods use different null hypotheses of absence of selection. Each method models the genomic information under mathematical procedures that imply different shifts on the genome: simple genetic differentiation on each marker genetic, F_{st}; shifts on multi-locus including the alleles distribution, XP-CLR and linkage disequilibrium decay including the haplotypes extension, XP-EHH. Selection signatures can be better identified when the modelling of absence or presence of selection mirrors the history of selection for particular populations

Figure 1 (colour online) Single nucleotide polymorphism (SNP) density within 1-Mb window size for each chromosome (Chr), using 89 968 SNPs after quality control for the intramuscular fat genomic data in rabbits.

(Kim *et al.*, 2015; González-Rodríguez *et al.*, 2016). In addition, we considered a cut-off of 1% in order to reduce false-positive selection signatures (Mallick *et al.*, 2009).

The genomic regions for each rabbit chromosome (OCU) were defined based on the OryCun2.0 rabbit genome assembly (Carneiro *et al.*, 2014b). The genes comprised within those genomic regions were identified using *BIOMART (Ensemble 98)*, available at https://www.ensembl.org/index.html; then, the functional annotation and gene ontology (GO) terms enrichment analysis was performed using *Enrichr* (Kuleshov *et al.*, 2016). The gene functional analysis was carried out using the rabbit and mouse annotation databases. Parameters recommended by the developers of *Enrichr* (i.e. *P*-value < 0.05 for Fisher exact test and a high combined score, greater than 70) were used for the identification of genes in this study. We focused on the genes related to biological functions (GO terms) of energy metabolism and lipid metabolism.

Results and discussion

The average physical distance between SNPs was 23.51 kb after quality control. However, the density of the SNPs through the rabbit genomic map was heterogeneous (Figure 1). The average SNP density for each 1-Mb window was 41.87 with a SD of 22.36, ranging between 0 and 93. The OCU14,

OCU20 and OCU21 had a low density with the average number of SNPs 32.29, 35.56 and 28.38 per Mb, respectively. In OCU14, two large gaps without any SNP marker were found (54.0 to 64.0 and 89.0 to 95.0 Mb). Despite the gaps, the SNP density used in our study was in line with other studies (Gurgul *et al.*, 2018; Ma *et al.*, 2019). The heterogeneity of SNP density confirmed the need of data normalisation for the methods to detect selection signatures.

The linkage disequilibrium was very high, with estimated R^2 values of 0.81, 0.68 and 0.52 at physical distances of 40, 250 and 1000 kb, respectively. The number of linkage disequilibrium blocks was 2309 encompassing 89 346 SNPs and showing a distance of 1 Mb for the longest linkage disequilibrium block. Moreover, the MDS displayed a noticeable genomic differentiation between the individuals (8th and 9th generation) from the High and Low lines; 17.59% and 2.75% of variance explained by two first dimensions, respectively (Figure 2). As expected, this analysis also supported the close relationship between the parents (8th generation) and their offspring (9th generation) in both lines.

Genome-wide scan for populations from divergent selection

The identification of potential signatures of selection was performed with three methods (F_{st}, XP-CLR and XP-EHH) and three window sizes (250, 500 and 1000 kb). The
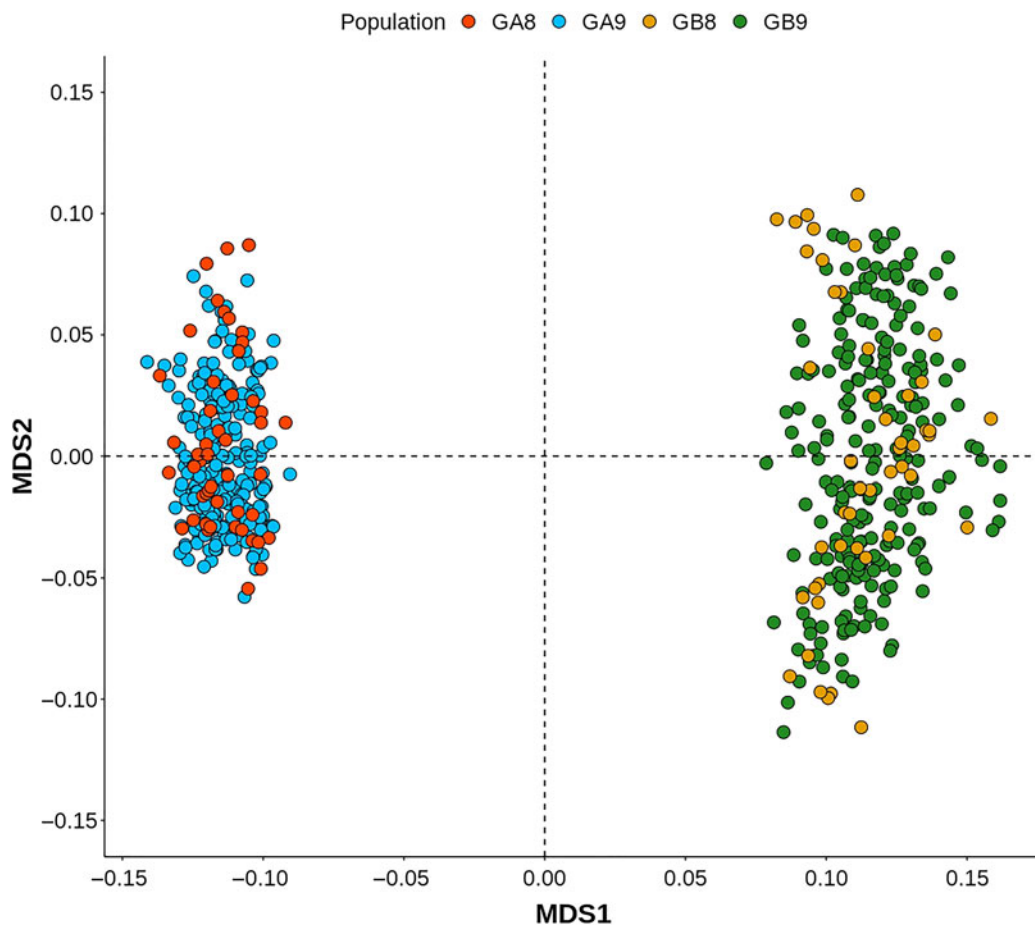
Figure 2 (colour online) Multidimensional scaling plot of intramuscular fat genomic data in rabbits. The plot displays the first component (MDS1), the second component (MDS2) and intramuscular fat lines: (left panel) High line at 8th generation (GA8 in red), High line at 9th generation (GA9 in light blue), (right panel) Low line at 8th generation (GB8 in yellow), Low line at 9th generation (GB9 in dark green).

Table 1 *Correlations between the levels of window size within methods of selection signatures in rabbits*

	XP-CLR 250 kb	XP-CLR 500 kb	XP-CLR 1000 kb	XP-EHH 250 kb	XP-EHH 500 kb	XP-EHH 1000 kb	F_{st} 250 kb	F_{st} 500 kb	F_{st} 1000 kb
XP-CLR 250 kb	1	0.8798	0.8142	−0.0284	−0.0298	−0.0437	0.1804	0.1697	0.1545
XP-CLR 500 kb		1	0.9587	−0.0427	−0.0463	−0.0634	0.1804	0.1731	0.1571
XP-CLR 1000 kb			1	−0.0442	−0.0486	−0.0669	0.1787	0.1726	0.1572
XP-EHH 250 kb				1	0.8951	0.7189	−0.0195	−0.0226	−0.0238
XP-EHH 500 kb					1	0.8229	−0.0259	−0.0283	−0.0302
XP-EHH 1000 kb						1	−0.0403	−0.0424	−0.0463
F_{st} 250 kb							1	0.9719	0.9215
F_{st} 500 kb								1	0.9648
F_{st} 1000 kb									1

XP-CLR = cross population − composite likelihood ratio; XP-EHH = cross population − extended haplotype homozygosity; F_{st} = Wright's fixation index.

correlations between window sizes within every method were very high for F_{st} method (>0.92) and lower for the XP-CLR and XP-EHH methods, especially between 250 and 1000 kb (Table 1). However, all correlations within the methods were high, between 0.72 and 0.97, confirming that the results were robust to variations in window size. Thus, hereinafter we will refer exclusively to the results

obtained with a window size of 500 kb. As expected, the correlations of the results between methods were very low, with an average of 0.08 using absolute values (Table 1). These correlations agreed with a genome scan study using pigs divergently selected for intramuscular fat, in which the correlations were less of 0.12 between methods to detect selection signatures (iHS, F_{st} and Rsb; Kim *et al.*, 2015). It can be

explained because every method entails a distinct hypothesis (Qanbari and Simianer, 2014; González-Rodríguez *et al.*, 2016), capturing different selection signals depending on the timescale of selective events (Utsunomiya *et al.*, 2013). Hence, we decided to analyse the results separately and those SNPs exceeding a cut-off of 1% in at least two methods were used to establish selection signatures and to search for candidate genes.

The average of 500 kb windows for F_{st} (non-normalised), XP-CLR (non-normalised) and XP-EHH (in absolute value) was 0.0973, 1.7228 and 0.5878, respectively. The average of F_{st} by computing each SNP was 0.10. This average of F_{st} was higher than the results obtained in other studies between several populations of domestic European rabbit, $F_{st} = 0.08$ (Carneiro *et al.*, 2014a), and also, from an experiment of divergent selection for uterine capacity in rabbits, $F_{st} = 0.05$ (Sosa-Madrid *et al.*, 2017). To our knowledge, no comparison can be made for XP-EHH and XP-CLR because until date this is the first study using these methods for detecting signatures of selection in rabbits.

The results of the genome-wide scans are shown by Manhattan plots in Figure 3. The results of each method individually showed several chromosomes with SNPs exceeding the cut-off of 1%. However, the joint results of the three methods did not evidence a genomic region clearly linked to a signature of selection. None of SNPs had values exceeding the cut-off of 1% in all methods (Figure 4). In contrast, several SNPs (80) associated with selection signatures overlapped between at least two methods. The overlapping results between XP-CLR and F_{st} were greater than the others, harbouring SNPs in OCU16 (31 SNPs) and OCU17 (24 SNPs); see Table 2. It can be explained because F_{st} and XP-CLR are based on differences in allele frequency and could detect dramatic shifts of opposite symmetrical allele frequencies for the SNPs located in the vicinity of an important causative variant. Conversely, XP-EHH is based on haplotype lengths comparison and was designed to compare a selected population with a reference population (non-selected; see Sabeti *et al.*, 2007). Then, if the extension of a selected haplotype occurs in both lines with similar strength, the power of detection for XP-EHH could be lower than when regions were selected in one of the divergent lines but not in the another.

In total, eight genomic regions of the rabbit genome were identified as selection signatures (Table 2). This number is low in comparison with most signature selection studies (13 to 224 regions) that used populations from divergent artificial selection experiments for BW, antibody response and feather pecking behaviour in poultry (Johansson *et al.*, 2010; Grams *et al.*, 2015; Lillie *et al.*, 2017) and for intramuscular fat and backfat thickness in pigs (15 regions; Kim *et al.*, 2015). Nevertheless, most of these studies identified selection signatures using only one method, unlike our study.

The selection signatures identified in our study can be a consequence of the selection of a polygenic trait with a high heritability such as the intramuscular fat (Martínez-Álvaro *et al.*, 2016) or due to the effect of genetic drift. The last hypothesis could be plausible because of the reduced number of sires used in the first generations, the mating structure (a female was mated with the same sire during its production life; Zomeño *et al.*, 2013), and the increase of two sire families in the last generations in each line. In many ways, detection of selective sweeps in smaller populations is more difficult than in large populations as extensive drift can obscure and weaken the selection signatures (Mallick *et al.*, 2009; Johansson *et al.*, 2010). However, genetic drift would generate random shifts of allelic frequency across the rabbit genome and our results showed the existence of consecutives SNPs with high scores of selection signatures within methods which is a direct evidence of selection. For instance, cluster 5 in OCU16 presented high scores of normalised F_{st} (up to 17.85, Table 2) and a substantial length (487.49 kb) encompassing 31 SNPs. They could be identified because we employed windows for detecting selection signatures instead of evaluating each SNP of the rabbit array. In addition, under the F_{st} method and a cut-off of 1%, a cluster in OCU13 encompassing 45 SNPs was identified by this study in a relevant genomic region (83.8 to 86.0 Mb) associated with intramuscular fat in rabbits, according to a genome-wide association study (**GWAS**) using the two lines of divergent selection (Sosa-Madrid *et al.*, 2020). This region showed SNPs with high normalised F_{st}, reaching values up to 20.33 (0.51 as F_{st} mean of 500 kb windows), albeit with only one of these SNPs, Affx-151937959, agreed with the relevant SNPs reported by GWAS (Sosa-Madrid *et al.*, 2020). This SNP showed a low MAF (0.09), but the surrounding SNPs presented very high MAF (up to 0.44). The methods of selection signatures can validate GWAS results assuming the presence of major genes affecting a selected trait. Otherwise, these methods would reveal new associated genomic regions, unlike GWAS results, when the selected trait has a large polygenic component influencing several genomic regions (Qanbari and Simianer, 2014).

On the other hand, divergent selection for intramuscular fat did not lead to fixation of alternate alleles of any of the SNPs studied. The selection response was very high (3.1 SD); hence, we expected some SNPs associated with causal variants had their alternate alleles fixed to nearby fixation in one of the opposite divergent lines (e.g. frequencies in High line: 0A/1T and in Low line: 1A/0T). These SNPs would show MAF values of 0.5 using all samples (both lines). However, the SNPs did not show both conditions. All these results would suggest several soft selective sweeps caused by short-term divergent selection of intramuscular fat instead of few hard selective sweeps, controlling this trait (Oleksyk *et al.*, 2010).

Underlying selected genes and gene ontology terms for divergent selection

Potential candidate genes were explored within the genomic regions identified as signatures of selection using a cut-off of 1%. The number of genes disclosed for each method was 579, 443 and 368 for XP-CLR, XP-EHH and F_{st}, respectively (see Supplementary Table S1). From these genes, 73 were detected by at least 2 methods of selection signatures.

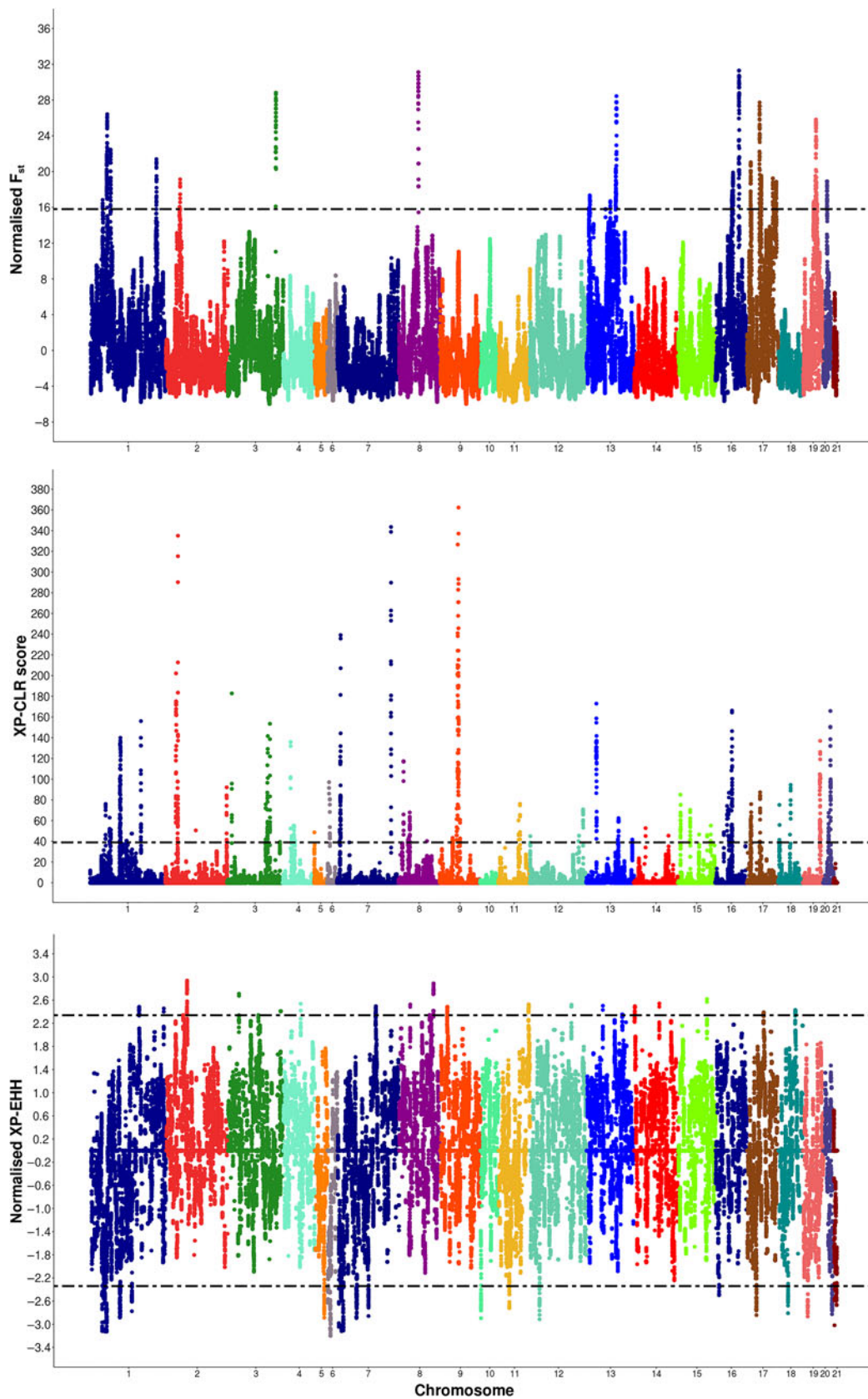**Figure 3** (colour online) Manhattan plot of 500 kb window for every method of selection signature in rabbits: normalised Wright's fixation index, F_{st} (top); cross population – composite likelihood ratio test, XP-CLR score (middle) and normalised cross population - extended haplotype homozygosity test, XP-EHH (bottom). The dashed line denotes the cut-off of 1% (F_{st}: 15.81, XP-CLR: 38.94, and XP-EHH: $+/-2.34$).
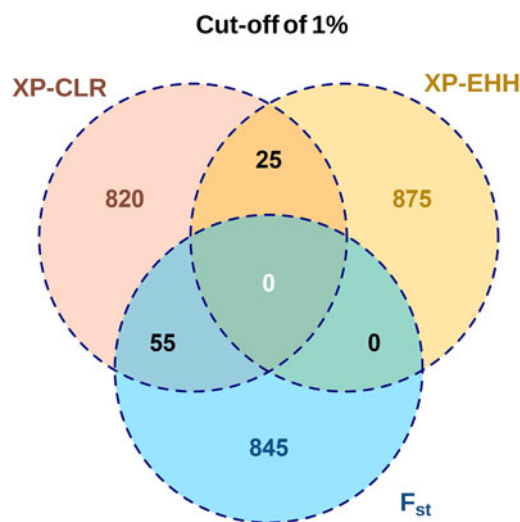
Figure 4 (colour online) Venn diagram of methods of selection signatures in rabbits: cross population – composite likelihood ratio test (XP-CLR), cross population – extended haplotype homozygosity test (XP-EHH) and Wright's fixation index (F_{st}).

These genes were grouped in 63 protein coding and 10 non-protein coding genes (see Supplementary Table S2). The results of the first 10 biological processes of the GO term enrichment analysis through *Enrichr* are presented in Supplementary Table S3. The GO term enrichment analysis did not identified pathways related to biological processes affecting the intramuscular fat.

A deep search of biological functions for the 73 genes disclosed 12 promising candidate genes related to lipid and carbohydrate metabolism which are important pathways to modulate the intramuscular fat (Table 3). Genes involved in the lipid metabolism were alkaline ceramidase 2 (*ACER2*), Perilipin 2 (*PLIN2*), Vimentin (*VIM*), Ras-related GTP binding A (*RRGA*), ribosomal protein S6 (*RPS6*), RNA Polymerase-Associated Protein RTF1 Homolog (*RTF1*), solute-carrier gene family 24 member 2 (*SLC24A2*) and phospholipase A2 group IVB (*PLA2G4B*). From these, it is worth to highlight *ACER2* (OCU1), *VIM* (OCU16) and *PLIN2* (OCU1), which are tightly related to lipid droplets and storage, being crucial in disease such as obesity, diabetes and atherosclerosis. *ACER2* encodes ceramidases which break down ceramides to sphingosine and free fatty acids at alkaline pH. *VIM* can cause an excessive endosomal cholesterol accumulation due to an imbalance of its iterations with other proteins (Walter *et al.*, 2009). *PLIN2* bears an essential role over long-chain fatty acid transport. Genomic studies reported *PLIN2* associations with intramuscular fat content (Gandolfi *et al.*, 2011) and its composition in pigs (Gol *et al.*, 2016). Moreover, gene expression studies for intramuscular fat in pig identified differentially expressed genes such as *RTF1* in OCU17 (Damon *et al.*, 2012) and *SLC24A2* in OCU1 (Li *et al.*, 2010). *RRGA* and *RPS6* in OCU1 could stimulate the lipogenesis and the lipid accumulation via activation of the mammalian target of rapamycin signalling pathways (Wipperman *et al.*, 2019). *PLA2G4B* in OCU17 is linked to phospholipid catabolic processes because of its phospholipase A2 (PL2) activity.

This enzyme has been widely studied using knockout and transgenic mice, showing to be important for the fatty acid pathway, for example, for oleic acid (Aloulou *et al.*, 2012).

Phenotypic variation of intramuscular fat between divergent lines could also be due to differences in regulation of lipid and carbohydrates (glycogen) metabolisms. This latter is important for intramuscular fat as the glycolytic products could be used to synthesise fatty acids, being incorporated into cholesterol esters, triacylglycerol and phospholipids in hepatocytes, increasing the lipid stores (Rui, 2014). Genes involved in the carbohydrate metabolism were RAR related orphan receptor A (*RORA*), glucosidae alpha neutral C (*GANC*), ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 6 (*ST8SIA6*) and DENN domain containing 4C (*DENND4C*). The regulation of differentiating pre-adipocytes by retinoic acid is controlled by *RORA* in OCU1, bearing a crucial role in triglyceride (lipids)/glucose homeostasis and various immune functions. The *RORA* functions are tightly related to hepatosteatosis, obesity and insulin resistance. Besides, *RORA* was identified by genomic studies in Nellore (Cesar *et al.*, 2014) and Chinese Wagyu cattle (Wang *et al.*, 2019) having extreme phenotypes of intramuscular fat composition and marbling. Hence, we presented RORA as the principal candidate gene for further studies. In addition, other genes, *GANC* in OCU17 is involved in the hydrolysis of glycogen and *ST8SIA6* in OCU16 is important in the pathways of oligosaccharide metabolic process and carbohydrate biosynthetic process. *DENND4C* in OCU1 could modulate indirectly the intramuscular fat content through control of glucose transport in response to insulin. However, the specific functions of these genes over the intramuscular fat remain unknown. Further analyses would be needed to corroborate the relationships between these genes (their polymorphisms) and the intramuscular fat content in rabbits.

Conclusions

In conclusion, a large number of genomic regions were identified within each method of selection signatures. A total number of 80 SNPs and 73 genes were detected using selection signatures exceeding cut-off of 1% at least 2 of the methods: XP-CLR, XP-EHH and F_{st}. General biological functions were retrieved from enrichment analysis. However, genes such as *ACER2, PLIN2* (OCU1), *ST8SIA6, VIM* (OCU16), *RORA, GANC* and *PLA2G4B* (OCU17) linked to energy metabolism, carbohydrates metabolism and lipid metabolism were identified as candidate genes to explain the differences in intramuscular fat observed between the divergent lines. The findings of the current study suggest that the intramuscular fat content in rabbits is influenced by a large polygenic component.

Acknowledgements

Table 2 *SNP name (SNP_ID), rabbit chromosome (OCU), cluster (genomic region), SNP physical position in megabase and values for three methods of selection signatures based on detection in at least two methods using a cut-off of 1%*

SNP_ID	OCU	CLUSTER	Physical position	XP-CLR	XP-EHH	F_{st}
Affx-151788669	1	1	34.40	**42.80**	**−2.78**	3.25
Affx-151841835	1		34.42	**46.32**	**−2.72**	4.05
Affx-151948493	1		34.42	**42.35**	**−2.72**	4.05
Affx-151981842	1		34.53	**70.13**	**−2.89**	7.58
Affx-151800050	1		34.55	**74.74**	**−2.88**	8.12
Affx-151888128	1		34.56	**42.29**	**−2.89**	8.10
Affx-151808312	1		34.57	**72.84**	**−2.88**	9.02
Affx-151996305	1		34.59	**76.09**	**−2.87**	9.25
Affx-151796600	1		34.60	**64.85**	**−2.86**	9.33
Affx-151996963	3	2	148.58	**39.42**	**2.41**	−1.34
Affx-151940966	6	3	6.58	**68.34**	**−2.89**	1.05
Affx-151916999	6		6.59	**47.82**	**−2.89**	1.11
Affx-151906393	6		6.60	**80.26**	**−2.88**	1.17
Affx-151850643	6		6.62	**75.42**	**−3.11**	1.62
Affx-152006617	6		6.63	**80.80**	**−3.07**	1.62
Affx-151909107	6		6.65	**53.68**	**−2.98**	1.86
Affx-151858638	7	4	7.85	**62.05**	**−2.48**	0.64
Affx-151988414	7		7.87	**48.46**	**−2.53**	1.11
Affx-151901134	7		7.89	**83.02**	**−2.54**	1.26
Affx-151884578	7		7.89	**80.85**	**−2.50**	1.26
Affx-151968222	7		7.91	**82.74**	**−2.53**	1.26
Affx-151923372	7		7.92	**82.83**	**−2.55**	1.55
Affx-151832398	7		7.95	**60.81**	**−2.41**	2.09
Affx-151887243	7		7.96	**78.92**	**−2.41**	2.09
Affx-151798377	7		7.99	**121.74**	**−2.38**	2.65
Affx-152002624	16	5	44.14	**56.15**	0.57	**16.01**
Affx-151964090	16		44.16	**76.86**	0.58	**16.05**
Affx-151954735	16		44.18	**56.19**	0.57	**16.63**
Affx-152011401	16		44.21	**49.99**	0.58	**16.44**
Affx-151994299	16		44.25	**76.42**	0.58	**16.82**
Affx-151935006	16		44.27	**44.46**	0.58	**16.87**
Affx-151934731	16		44.28	**39.12**	0.58	**16.91**
Affx-151916386	16		44.31	**41.23**	0.58	**17.23**
Affx-151945660	16		44.31	**49.81**	0.58	**17.05**
Affx-151892655	16		44.33	**85.41**	0.58	**16.87**
Affx-151923274	16		44.34	**139.09**	0.57	**16.92**
Affx-151981680	16		44.34	**164.39**	0.57	**16.92**
Affx-151922936	16		44.37	**166.24**	0.57	**16.42**
Affx-152012312	16		44.37	**146.40**	0.85	**16.23**
Affx-151904619	16		44.38	**111.16**	0.85	**16.83**
Affx-151947283	16		44.39	**78.58**	0.85	**16.83**
Affx-151892171	16		44.41	**47.52**	0.85	**17.38**
Affx-151999419	16		44.42	**50.66**	0.71	**17.38**
Affx-151831673	16		44.46	**53.22**	0.74	**16.96**
Affx-151877806	16		44.47	**52.99**	0.74	**17.19**
Affx-151933718	16		44.48	**52.90**	0.74	**17.41**
Affx-151961515	16		44.50	**51.22**	0.73	**17.85**
Affx-152004824	16		44.51	**50.14**	0.73	**17.67**
Affx-152008187	16		44.52	**74.51**	0.73	**17.46**
Affx-151886887	16		44.54	**93.91**	0.74	**17.25**
Affx-151900728	16		44.55	**100.41**	0.32	**17.25**
Affx-151875439	16		44.56	**96.80**	0.31	**17.25**
Affx-151786498	16		44.57	**118.82**	0.30	**16.81**
Affx-151820958	16		44.59	**131.41**	0.32	**16.86**
Affx-151942314	16		44.61	**112.63**	0.31	**16.40**
Affx-151797733	16		44.63	**100.38**	0.30	**15.93**

(Continued)

Table 2 *(Continued)*

SNP_ID	OCU	CLUSTER	Physical position	XP-CLR	XP-EHH	F_{st}
Affx-151854426	17	6	11.32	**47.27**	−1.68	**17.79**
Affx-152017855	17		11.45	**41.36**	−1.71	**20.90**
Affx-151809007	17		11.46	**40.34**	−1.71	**20.65**
Affx-151945077	17		11.49	**44.26**	−1.71	**20.79**
Affx-151827750	17		11.50	**45.09**	−1.71	**21.05**
Affx-151813388	17		11.51	**49.33**	−1.70	**20.79**
Affx-151897106	17		11.51	**60.52**	−1.71	**20.69**
Affx-151970040	17		11.52	**61.40**	−1.69	**20.95**
Affx-151854218	17		11.53	**60.99**	−1.69	**20.69**
Affx-151999939	17		11.54	**58.80**	−1.69	**20.70**
Affx-151872016	17		11.56	**56.37**	−1.69	**20.06**
Affx-151809616	17		11.57	**55.06**	−1.67	**19.42**
Affx-151800782	17		11.58	**54.38**	−1.69	**19.06**
Affx-151992875	17		11.59	**53.07**	−1.64	**18.71**
Affx-151953403	17		11.62	**46.73**	−1.64	**16.22**
Affx-151860917	17	7	29.59	**39.64**	−0.22	**19.14**
Affx-151984545	17		29.64	**39.42**	−0.40	**17.52**
Affx-151841455	17	8	30.42	**80.81**	−0.16	**15.89**
Affx-152009920	17		30.45	**39.82**	−0.27	**16.57**
Affx-151819416	17		30.55	**84.32**	0.06	**16.43**
Affx-151905376	17		30.56	**87.35**	−0.13	**16.43**
Affx-151933923	17		30.57	**83.99**	−0.12	**16.15**
Affx-151909639	17		30.57	**81.33**	0.74	**16.45**
Affx-151912729	17		30.59	**71.87**	0.79	**15.99**

SNP = single nucleotide polymorphism; XP-CLR = cross population – composite likelihood ratio test; XP-EHH = cross population - extended haplotype homozygosity test, these values are normalised; F_{st} = fixation index, these values are normalised.
The bold data and green cells indicate the values exceeding a cut-off of 1% (XP-CLR: 38.94 and XP-EHH: +/− 2.34, F_{st}: 15.81).

Table 3 *Biological processes of highlighted genes identified by at least two methods of selection signatures in rabbits for cut-off of 1%*

Biological Process	Highlighted genes
Insulin signalling pathways	*DENND4C*
Lipid droplets and storage	*VIM, PLIN2, ACER2*
mTORC signalling pathways	*RRAGA, RPS6*
Carbohydrate metabolism process	*RORA, ST8SIA6, GANC*
Lipid metabolic process	*RORA, ACER2, PLA2G4B*
Regulation of adipocytes differentiation	*RORA*
Phospholipase activity	*PLA2G4B*
Processes related to intramuscular fat*	*PLIN2, SLC24A2, RTF1, RORA*

mTOR = the mammalian target of rapamycin; *DENND4C* = DENN domain containing 4C; *VIM* = Vimentin; *PLIN2* = Perilipin 2; *ACER2* = alkaline ceramidase 2; *RRGA* = Ras-related GTP binding A; *RPS6* = ribosomal protein S6; *RORA* = RAR related orphan receptor A; *ST8SIA6* = ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 6; *GANC* = glucosidase alpha neutral C; *PLA2G4B* = phospholipase A2 group IVB; *SLC24A2* = solute-carrier gene family 24 member 2; *RTF1* = RNA Polymerase-Associated Protein RTF1 Homolog.
*Based on genomic and gene expression studies of intramuscular fat.

B. S. Sosa-Madrid 0000-0002-4269-7593
L. Varona 0000-0001-6256-5478
A. Blasco 0000-0003-3558-0810
P. Hernández 0000-0001-7857-7334
C. Casto-Rebollo 0000-0003-4646-345X
N. Ibáñez-Escriche 0000-0002-6221-3576

Declaration of interest

The authors declare that they do not have any conflict of interest.

Ethics statement

The animal manipulations were approved by the Ethical Committee of the Universitat Politècnica de València, according to Council Directives 98/58/EC.

Software and data repository resources

Genomic data of the 9th generation of intramuscular fat lines were deposited in Figshare Repository (https://doi.org/10.6084/m9.figshare.9934058.v1). The remaining data sets used in the current study are available from the corresponding author on reasonable request.

Author's contributions

BSS and LV carried out the main statistical analyses. AB, NIE and PH conceived of the study design and secured substantial funding. PH, BSS and CCR performed the phenotypic data recording,

and collected and processed DNA samples. LV, CCR, AB, PH, BSS and NIE supervised analyses, drafted and reviewed critically the manuscript. All authors read and approved the final manuscript. BSS is a researcher undertaking new genomic research for applied livestock programmes in rabbits. LV is a full professor of animal breeding and genetics focused on the development of new statistical procedures in quantitative and molecular genetics. PH is a full professor of carcass and meat quality. AB is a full professor of animal breeding and genetics. CCR is a doctoral student undertaking new omics research for applied livestock programmes in rabbits. NIE is a senior lecture focusing on aspect of genomic selection and canalisation in pigs and rabbits.

Supplementary material

To view supplementary material for this article, please visit https://doi.org/10.1017/S1751731120001263

References

Aloulou A, Ali Y Ben, Bezzine S, Gargouri Y and Gelb MH 2012. Phospholipases: an overview. In Lipases and phospholipases: methods and protocols (ed. G. Sandoval), pp. 63–85. Humana Press, Totowa, NJ, USA.

Beissinger TM, Rosa GJ, Kaeppler SM, Gianola D and de Leon N 2015. Defining window-boundaries for genomic analyses using smoothing spline techniques. Genetics Selection Evolution 47, 30.

Carneiro M, Albert FW, Afonso S, Pereira RJ, Burbano H, Campos R, Melo-Ferreira J, Blanco-Aguiar JA, Villafuerte R, Nachman MW, Good JM and Ferrand N 2014a. The Genomic Architecture of Population Divergence between Subspecies of the European Rabbit. PLoS Genetics 10, e1003519.

Carneiro M, Rubin CJ, Di Palma F, Albert FW, Alföldi J, Barrio AM, Pielberg G, Rafati N, Sayyab S, Turner-Maier J, Younis S, Afonso S, Aken B, Alves JM, Barrell D, Bolet G, Boucher S, Burbano HA, Campos R, Chang JL, Duranthon V, Fontanesi L, Garreau H, Heiman D, Johnson J, Mage RG, Peng Z, Queney G, Rogel-Gaillard C, Ruffier M, Searle S, Villafuerte R, Xiong A, Young S, Forsberg-Nilsson K, Good JM, Lander ES, Ferrand N, Lindblad-Toh K and Andersson L 2014b. Rabbit genome analysis reveals a polygenic basis for phenotypic change during domestication. Science 345, 1074–1079.

Cesar AS, Regitano LC, Tullio RR, Lanna DP, Nassu RT, Mudado MA, Oliveira PS, do Nascimento ML, Chaves AS, Alencar MM, Sonstegard TS, Garrick DJ, Reecy JM and Coutinho LL 2014. Genome-wide association study for intramuscular fat deposition and composition in Nellore cattle. BMC Genetics 15, 39.

Chen H, Patterson N and Reich D 2010. Population differentiation as a test for selective sweeps. Genome Research 20, 393–402.

Damon M, Wyszynska-Koko J, Vincent A, Hérault F and Lebret B 2012. Comparison of muscle transcriptome between pigs with divergent meat quality phenotypes identifies genes related to muscle metabolism and structure. PLoS ONE 7, e33763.

Gandolfi G, Mazzoni M, Zambonelli P, Lalatta-Costerbosa G, Tronca A, Russo V and Davoli R 2011. Perilipin 1 and perilipin 2 protein localization and gene expression study in skeletal muscles of European cross-breed pigs with different intramuscular fat contents. Meat Science 88, 631–637.

Gol S, Ros-Freixedes R, Zambonelli P, Tor M, Pena RN, Braglia S, Zappaterra M, Estany J and Davoli R 2016. Relationship between perilipin genes polymorphisms and growth, carcass and meat quality traits in pigs. Journal of Animal Breeding and Genetics 133, 24–30.

González-Rodríguez A, Munilla S, Mouresan EF, Cañas-Álvarez JJ, Díaz C, Piedrafita J, Altarriba J, Baro JÁ, Molina A and Varona L 2016. On the performance of tests for the detection of signatures of selection: a case study with the Spanish autochthonous beef cattle populations. Genetics Selection Evolution 48, 81.

Grams V, Wellmann R, Preuß S, Grashorn MA, Kjaer JB, Bessei W and Bennewitz J 2015. Genetic parameters and signatures of selection in two divergent laying hen lines selected for feather pecking behaviour. Genetics Selection Evolution 47, 77.

Gurgul A, Jasielczuk I, Ropka-Molik K, Semik-Gurgul E, Pawlina-Tyszko K, Szmatoła T, Szyndler-Nądza M, Bugno-Poniewierska M, Blicharski T, Szulc K, Skrzypczak E and Krupiński J 2018. A genome-wide detection of selection signatures in conserved and commercial pig breeds maintained in Poland. BMC Genetics 19, 95.

Johansson AM, Pettersson ME, Siegel PB and Carlborg Ö 2010. Genome-wide effects of long-term divergent selection. PLoS Genetics 6, e1001188.

Kim E-S, Ros-Freixedes R, Pena RN, Baas TJ, Estany J and Rothschild MF 2015. Identification of signatures of selection for intramuscular fat and backfat thickness in two Duroc populations. Journal of Animal Science 93, 3292–3302.

Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW and Ma'ayan A 2016. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Research 44, W90–W97.

Li X, Lee C-K, Choi B-H, Kim T-H, Kim J-J and Kim K-S 2010. Quantitative gene expression analysis on chromosome 6 between Korean native pigs and Yorkshire breeds for fat deposition. Genes & Genomics 32, 385–393.

Lillie M, Sheng Z, Honaker CF, Dorshorst BJ, Ashwell CM, Siegel PB and Carlborg Ö 2017. Genome-wide standing variation facilitates long-term response to bidirectional selection for antibody response in chickens. BMC Genomics 18, 99.

Ma H, Zhang S, Zhang K, Zhan H, Peng X, Xie S, Li X, Zhao S and Ma Y 2019. Identifying selection signatures for backfat thickness in Yorkshire pigs highlights new regions affecting fat metabolism. Genes 10, 254.

Mallick S, Gnerre S, Muller P and Reich D 2009. The difficulty of avoiding false positives in genome scans for natural selection. Genome Research 19, 922–933.

Martínez-Álvaro M, Hernández P and Blasco A 2016. Divergent selection on intramuscular fat in rabbits: responses to selection and genetic parameters. Journal of Animal Science 94, 4993–5003.

Mauch E, Servin B, Gilbert H and Dekkers J 2018. Signatures of selection in two independent populations of pigs divergently selected for feed efficiency. Animal Industry Report AS 664, ASL R3274.

Oleksyk TK, Smith MW and O'Brien SJ 2010. Genome-wide scans for footprints of natural selection. Philosophical Transactions of the Royal Society B: Biological Sciences 365, 185–205.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ and Sham PC 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. The American Journal of Human Genetics 81, 559–575.

Qanbari S and Simianer H 2014. Mapping signatures of positive selection in the genome of livestock. Livestock Science 166, 133–143.

Rui L 2014. Energy metabolism in the liver. In Comprehensive physiology, pp. 177–197. John Wiley & Sons, Inc., Hoboken, NJ, USA.

Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie X, Byrne EH, McCarroll SA, Gaudet R, Schaffner SF and Lander ES 2007. Genome-wide detection and characterization of positive selection in human populations. Nature 449, 913–918.

Sargolzaei M, Chesnais JP and Schenkel FS 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genomics 15, 478.

Sosa-Madrid BS, Hernández P, Blasco A, Haley CS, Fontanesi L, Santacreu MA, Pena RN, Navarro P and Ibáñez-Escriche N 2020. Genomic regions influencing intramuscular fat in divergently selected rabbit lines. Animal Genetics 51, 58–69.

Sosa-Madrid BS, Ibañez-Escriche N, Santacreu MA, Varona L and Blasco A 2017. Huellas de selección en un experimento de seleccion divergente para capacidad uterina en conejo. In Proceedings of the XVII Jornadas sobre Producción Animal, 30–31 May 2017, Zaragoza, Spain, pp. 558–560.

Szpiech ZA and Hernandez RD 2014. selscan: an efficient multithreaded program to perform EHH-based scans for positive selection. Molecular Biology and Evolution 31, 2824–2827.

Utsunomiya YT, Pérez O'Brien AM, Sonstegard TS, Van Tassell CP, do Carmo AS, Mészáros G, Sölkner J and Garcia JF 2013. Detecting loci under recent positive selection in dairy and beef cattle by combining different genome-wide scan methods. PLoS ONE 8, e64280.

Walter M, Chen FW, Tamari F, Wang R and Ioannou YA 2009. Endosomal lipid accumulation in NPC1 leads to inhibition of PKC, hypophosphorylation of vimentin and Rab9 entrapment. Biology of the Cell 101, 141–153.

Wang Z, Ma H, Xu L, Zhu B, Liu Y, Bordbar F, Chen Y, Zhang L, Gao X, Gao H, Zhang S, Xu L and Li J 2019. Genome-wide scan identifies selection signatures in Chinese Wagyu cattle using a high-density SNP array. Animals 9, 296.

Wipperman MF, Montrose DC, Gotto AM and Hajjar DP 2019. Mammalian target of rapamycin. The American Journal of Pathology 189, 492–501.

Zomeño C, Hernandez P and Blasco A 2013. Divergent selection for intramuscular fat content in rabbits. I. Direct response to selection. Journal of Animal Science 91, 4526–4531.